Technische Universität München
Fakultät für Mathematik

On the dynamics of stochastic heat equations

Anne Pein

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
genehmigten Dissertation.

Vorsitzende(r):		Prof. Dr. Elisabeth Ullmann
Prüfer der Dissertation:	1.	Prof. Dr. Christian Kuehn
	2.	Prof. Dr. Alexandra Neamţu
	3.	Prof. Dr. Codina Cotar

Die Dissertation wurde am 15.02.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 22.06.2021 angenommen.

Acknowledgements

I would like to thank my advisor, Christian Kuehn, for his support and for giving me the opportunity and freedom to work on my dissertation during the last four years. I am thankful to Alexandra Neamţu for supporting me and for co-examining this thesis. I am grateful for having met Eleanor Archer and I would like to thank her for supporting me. Likewise, I would like to thank Codina Cotar for hosting me at University College London, for her support and for co-examining this thesis. I am thankful to David Croydon for hosting me at Kyoto University. My research stays at UCL and Kyoto University were funded by the German Academic Exchange Service, the Japan Society for the Promotion of Science and TUM Graduate School, for which I am very grateful. Furthermore, I am thankful to the mathematicians Carina Geldhauser, Manuel Gnann and Felix Voigtlaender for their support and to Elisabeth Ullmann for taking the chair of the examination committee.

During my time at TUM, UCL and Kyoto University, I was always surrounded by cordial and inspiring colleagues with whom I shared many meals, discussions and moments of joy. Thank you.

I am grateful for wonderful friends, old and new, here and there. Lastly, I am thankful to my parents and to my sister for their support.

Thank you.

Abstract

The main objective of this thesis is to analyse the dynamical behaviour of different linear and non-linear stochastic heat equations using analytical and probabilistic techniques. We begin with the linear setting of the parabolic Anderson model on a random tree with a stationary random potential, where we characterize the intermittent long-time behaviour of solutions. Then we continue to analyse coupled systems of non-linear stochastic heat equations on continuous spatial domains with time-dependent additive and multiplicative random perturbations given by Wiener processes. Here, we derive the existence of random attractors by means of random dynamical system theory. Finally, we characterize fluctuations around the slow manifold of a linear fast-slow system with a slowly varying parameter perturbed by an infinite-dimensional Wiener process.

Zusammenfassung

In dieser Doktorarbeit analysieren wir das dynamische Verhalten von verschiedenen linearen und nicht-linearen stochastischen Wärmeleitungsgleichungen. Zunächst befassen wir uns mit dem linearen parabolischen Anderson Modell, definiert auf einem zufälligen Baum und gestört durch ein stationäres zufälliges Potential. Wir zeigen, dass sich Lösungen im Laufe der Zeit auf einen einzelnen Knoten lokalisieren. Anschließend beschäftigen wir uns mit gekoppelten Systemen von nicht-linearen stochastischen Wärmeleitungsgleichungen auf kontinuierlichen Räumen mit zeitabhängigen additiven und multiplikativen zufälligen Störungen in Form von Wiener Prozessen. Wir charakterisieren das Langzeitverhalten von Lösungen anhand von zufälligen Attraktoren mit Hilfe der Theorie von zufälligen dynamischen Systemen. Abschließend konzentrieren wir uns auf eine stochastische Wärmeleitungsgleichung mit einem langsam variierenden Parameter und beschreiben Fluktuationen der Lösung um die langsame Mannigfaltigkeit.

Contents

Introduction

In the year 1807 J. B. Fourier (1768-1830) formulated in his seminal manuscript 'Théorie de la Propagation de la Chaleur dans les Solides' (later publsihed in [Fou22]) amongst others the partial differential equation (PDE) describing the process of heat conduction, today known as the *heat equation*. It reads as follows

$$\rho c \frac{\partial T(t,x)}{\partial t} = \nabla \cdot K \nabla T(t,x), \tag{0.1}$$

where $T = T(t,x)$ stands for the temperature at a point $x \in D \subset \mathbb{R}^3$ in a solid body at time $t \geq 0$. The constants K, ρ and c describe the *thermal conductivity*, the *density* and the *specific heat capacity* of the solid, respectively. Equation (0.1) can be derived from a physical conservation law, which states that the change in heat content in D per unit time is equal to the flux of heat through the boundary (in the absence of sinks and sources), together with the laws that (a) the heat flow is proportional to the temperature gradient and that (b) the heat content is proportional to the temperature. A historical perspective on Fourier's work on heat propagation can be found in [Nar99].

Assuming that the thermal conductivity is homogeneous within the body and defining the *thermal diffusivity* $d := \frac{K}{c\rho}$ (also called *diffusion constant*), equation (0.1) reduces to

$$\partial_t u(t,x) = d\Delta u(t,x), \quad (t,x) \in \mathbb{R}^+ \times \mathbb{R}^n, \tag{0.2}$$

where we have used the letter u for the generic unknown function in a partial differential equation and where we consider an n-dimensional spatial domain $\mathbb{R}^n$ instead of the real world space $D \subset \mathbb{R}^3$. Equation (0.2) is the simplest representative of the class of *parabolic partial differential equations* and it forms its conceptual foundation. For every bounded initial condition $u(0,x) = u_0(x)$ equation (0.2) possesses a unique continuous solution given by the convolution of the *heat kernel* with the initial condition, that is

$$u(t,x) = \frac{1}{(4\pi dt)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4dt}} u_0(y) \, \mathrm{d}y. \tag{0.3}$$

Adding a time and/or space dependent suitable function $f(t,x)$ to the equation allows to model sources and sinks of heat, i.e.

$$\partial_t u(t,x) = d\Delta u(t,x) + f(t,x), \quad (t,x) \in \mathbb{R}^+ \times \mathbb{R}^n. \tag{0.4}$$

We refer to (0.4) as an *inhomogeneous heat equation* and f is often referred to as a *forcing term*. The (inhomogeneous) heat equation models not only the conduction of heat in space but it can be used to describe the dynamics of the density of any diffusing material, and thus appears in the analysis of numerous physical, biological or social systems (also often referred to as *diffusion equation*).

Non-linear processes play a major role in natural systems. Thus, to capture the relevant dynamics it is often essential to include non-linear terms in a model. We formulate the *non-linear heat equation* as

$$\partial_t u(t,x) = d\Delta u(t,x) + F(u(t,x)), \quad (t,x) \in \mathbb{R}^+ \times \mathbb{R}^n, \tag{0.5}$$

where $F(u)$ is a suitable, possibly non-linear function in u. Equations of this form are also often called *reaction-diffusion equations* in the applied sciences, as they model the dynamics of concentrations of chemical substances that diffuse in space, while undergoing reactions (which depend on the current local concentration). The interplay or competition between the two central processes, diffusion and non-linear reaction, may lead to complex dynamical phenomena. Due to their importance in applications there is an extensive literature available on this type of equations (see for example the monographs [Smo94, Vol14, GR11, CD17, Per15] amongst many others).

Most real-world physical systems are subject to some kind of *random fluctuations*, which may enter at different 'levels': for example as internal fluctuations, as a random external forcing or as environmental noise. Internal fluctuations are mainly caused by microscopic effects such as molecular collisions and electric fluctuations. These microscopic effects are usually combined into a random perturbation in the dynamics of the macroscopic observables. That is, to do justice to all these effects, one may include random terms into the macroscopic modelling equation of a physical system, leading to *stochastic* or *random partial differential equations* (SPDEs or RPDEs), whose solutions are random objects. Here, one often uses the term random PDE when the randomness appears as a random parameter, while the term SPDE usually refers to a PDE combined with some stochastic process, which requires some form of stochastic integration. Further insight into the modelling aspect of spatially extended physical systems by SPDEs may for example be found in [GOS12]. The beginnings of the mathematical theory of SPDEs can be traced back at least to the early 1970s, with initial works by e.g. E. Cabaña [Cn70], A. Bensoussan and R. Temam [BT72] and É. Pardoux. For a historical perspective on this field we refer to [Zam20].

Let us now take a closer look at the *stochastic heat equation*

$$\begin{aligned} \partial_t u(t,x) &= d\triangle u(t,x) + \sigma(u(t,x))\xi(t,x), \qquad (t,x) \in (0,\infty) \times \mathbb{R}^n, \\ u(0,x) &= u_0(x), \end{aligned} \tag{0.6}$$

where $(\xi(t,x) : t > 0, x \in \mathbb{R}^n)$ is a random field in space and time, $\sigma : \mathbb{R} \to \mathbb{R}$ is a suitable, non-random function and $u_0(x)$ is a non-random initial condition.

For the choices $\sigma(u) \equiv 1$ and $\sigma(u) = u$, equation (0.6) is called the stochastic heat equation with *additive noise* and with *linear multiplicative noise*, respectively. The unabated interest in the stochastic heat equation and its non-linear variants stems not only from its wide-ranging applications in modelling, such as in astrophysics [Jon99] and neurophysiology [Wal81], to name just a few; but also from its connection to *particle systems* [Mue15, CM94] and its relation to the famous *Kardar-Parisi-Zhang equation* for interface growth via the *Cole-Hopf transformation* [Hai13] and to the stochastic *Burger's equation* [CM94]. Note that equation (0.6) can be considered as a multiparameter stochastic equation and the solution is a one-dimensional random field [Wal86], or as the solution to an infinite-dimensional stochastic differential equation and the solution is a stochastic process taking values in an infinite-dimensional function space, see also (0.8) for the standard notation in this interpretation. This latter approach was mainly developed by G. Da Prato and J. Zabczyk [DPZ92] and it is often referred to as the *semigroup approach.* As outlined in the next paragraph, the lack of regularity of random terms greatly influences the concept and analysis of solutions to (0.6).

Let us consider the case with *additive noise*; in this setting equation (0.6) can be regarded as a heat equation with a random inhomogeneity. If the randomness is caused by internal microscopic fluctuations on a very small scale, one may assume that it is completely uncorrelated in space and time. Then, a suitable mathematical model for the noise is given by the so-called *space-time white noise.* Space-time white noise is a mean zero Gaussian random field, which is formally characterised by the correlation function

$$\mathbb{E}[\xi(t,x)\xi(s,y)] = \delta(t-s)\delta(x-y).$$

Note that ξ is a *distribution* instead of a function (generalized Gaussian noise). In particular, its action on space-time test functions f, g has covariance

$$\mathbb{E}[\xi(f)\xi(g)] = \langle f, g\rangle_{L^2(\mathbb{R}^+\times\mathbb{R}^n)}.$$

For $d = 1$, $u(0,x) \equiv 0$, $\sigma \equiv 1$ the solution to (0.6) is given by the space-time convolution of the heat kernel with the random inhomogeneity

$$u(t,x) = \int_0^t \frac{1}{(4\pi|t-s|)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4(t-s)}} \xi(s,y)\, \mathrm{d}y\, \mathrm{d}s. \tag{0.7}$$

It turns out that the solution is a centred Gaussian, whose variance is finite if and only if $n = 1$. In this case, $u(t,x)$ is for fixed x almost surely α-Hölder continuous for any $\alpha < 1/4$ and for fixed t it is almost surely α-Hölder continuous for any $\alpha < 1/2$. For $n > 1$ the solution is not a function but a distribution. In particular in connection with non-linear problems this may cause problems as products of distributions may be ill-defined, i.e. it might not be clear what is actually meant for a distribution to be a solution in this setting. Such equations are called *singular* stochastic equation and dealing with them leads to the highly active fields of *rough paths*, *paracontrolled calculus* and *regularity structures* [Hai14, GIP15]. To avoid such subtleties we will consider more regular types of noise within this thesis. That is, we consider noise which is *spatially coloured*, i.e. spatially correlated. Formally, this may be expressed as

$$\mathbb{E}[\xi(t,x)\xi(s,y)] = \delta(t-s)f(x,y),$$

with some correlation function f, see Chapter 3 for details.

As mentioned above, one may view, just like in PDE theory, a randomly perturbed heat equation as a stochastic ordinary differential equation (SODE) in an infinite-dimensional Hilbert space H. It is in fact possible to define suitable *Hilbert space valued Wiener processes* $(W(t))_{t\geq 0}$, such that the non-linear heat equation with additive noise can be written as

$$\mathrm{d}u = (d\triangle u + F(u))\ \mathrm{d}t + \mathrm{d}W(t). \tag{0.8}$$

and the solution $u(t)$ is an element in H. The SPDE (0.8) can be analysed by means of infinite-dimensional Itô integration theory, as developed in [DPZ92]. We will introduce this approach in more detail in Chapter 3. A major source of research problems arises from asking how the noise influences the presumably complex dynamical behaviour, e.g. pattern formation, of such a non-linear equation. Hereby, it can not only be observed that effects governing the deterministic counterpart are perturbed by the noise, but also new phenomena, so-called *noise-induced phenomena*, which are not present in the deterministic setting, may appear.

Let us also take a brief look at the *heat equation with linear multiplicative noise*,

$$\begin{aligned} \partial_t u(t,x) &= d\triangle u(t,x) + u(t,x)\xi(t,x), \qquad (t,x)\in(0,\infty)\times\mathbb{R}^n, \\ u(0,x) &= u_0(x), \end{aligned} \tag{0.9}$$

where again $(\xi(t,x) : t > 0, x \in \mathbb{R}^n)$ is a random field in space and time. Equation (0.9) is also sometimes called the *parabolic Anderson model* (PAM) in a wider sense. One may differentiate between a stationary case, meaning here that ξ is time-independent, and a non-stationary case, meaning that ξ is time-dependent. In the stationary setting we arrive at,

$$\begin{aligned} \partial_t u(t,x) &= d\triangle u(t,x) + u(t,x)\xi(x), \qquad (t,x)\in(0,\infty)\times\mathbb{R}^n \\ u(0,x) &= u_0(x), \end{aligned} \tag{0.10}$$

and the random field $(\xi(x) : x \in \mathbb{R}^n)$ is called a *random potential*. This equation may be seen as a model for the spread of mass in a static random environment. The solution to the PAM (0.10) is well-known to exhibit an *intermittency effect*, meaning that almost all the solution is asymptotically concentrated in a small number of disjoint regions. More details on this phenomenon can be found in Chapter 2.

In this thesis we will analyse different types of *linear* and *non-linear stochastic heat equations* from different points of view. The overall goal is to seek a better understanding of the *dynamical behaviour* of solutions using various analytic and probabilistic techniques. This work consists of three main results, which can be found in the Chapters 2, 5 and 6. We now give a short summary of each result and an outline of the structure of this thesis.

We begin with a brief introduction to random trees in **Chapter 1**. The main purpose of this chapter is to introduce and characterise critical Galton-Watson trees conditioned to survive with an offspring distribution in the domain of attraction of a stable law, as this will be the spatial domain on which we work in Chapter 2. This chapter is based on joint work with Eleanor Archer.

In **Chapter 2** we analyse the *parabolic Anderson model* with a stationary random potential given by a family of iid random variables on a discrete spatial domain, namely on a *random tree*. The *intermittency phenomenon* that we mentioned above is on $\mathbb{Z}^n$ relatively well understood by now. In particular, the strength of the localisation behaviour depends on the tail decay of the random potential and several different regimes have been identified. We consider in this chapter the heavy-tailed *Pareto potential* that causes a very pronounced intermittency effect. Indeed, on $\mathbb{Z}^n$ it was shown in [KLMS09] that the solution to the PAM with Pareto potential localises eventually almost surely on only two sites and with high probability on one single site.

Motivated by the increasing interest in dynamics on networks, it is an intriguing task to consider the PAM on other graphs different from $\mathbb{Z}^n$. In Chapter 2 we therefore analyse the intermittency behaviour of the PAM on critical Galton-Watson trees conditioned to survive T_∞. This extends the current literature since the underlying graph is now random with *non-uniform volume growth* and *unbounded degree*. We prove that, similar to the $\mathbb{Z}^n$ setting, the PAM with Pareto potential on the tree model T_∞ localises with high probability in one single vertex for time going to infinity. The proof relies on a spectral analysis of the *Anderson Hamiltonian* $\triangle + \xi(\cdot)$ and the representation of the solution by a *Feynman-Kac formula*.

Chapter 2 is based on joint work with Eleanor Archer.

After Chapter 2 we switch over to reaction-diffusion equations on continuous spatial domains and with time-dependent random perturbations for the rest of the thesis. For this, we introduce relevant concepts and results in the fields of SPDEs and random dynamical system theory in **Chapters 3** and **4**, respectively.

In **Chapter 5** we analyse the long-term behaviour of certain systems of stochastic reaction-diffusion equations using the *random dynamical systems approach* (see Chapter 4). More precisely, we derive the existence of *random attractors*, i.e. random invariant compact sets of the phase space towards which solutions of the system evolve. The first type of systems that we consider are so-called *partly dissipative* *reaction-diffusion systems* with *additive noise* (Section 5.2), meaning that they consist of a coupling between a SPDE and a SODE, i.e. they have the general form

$$\begin{aligned} \mathrm{d}u_1 &= (d\Delta u_1 - h(x,u_1) - f(x,u_1,u_2))\,\mathrm{d}t + B_1\,\mathrm{d}W_1, \\ \mathrm{d}u_2 &= (-\sigma(x)u_2 - g(x,u_1,u_2))\,\mathrm{d}t + B_2\,\mathrm{d}W_2, \end{aligned} \tag{0.11}$$

where $W_{1,2}$ are Wiener processes, the σ, f, g, h are given functions, $B_{1,2}$ are suitable operators, $d > 0$ is a parameter and the equation is posed on a bounded open domain $D \subset \mathbb{R}^n$. Systems of this form appear in numerous models in the natural sciences such as the spatial Morris-Lecar model [ML81] in neuroscience, the cubic-quintic Allen-Cahn equation [Kue15a] in elasticity, and the Barkley model [Bar91] for spiral waves used in cardiac dynamics.

As will be outlined in Chapter 4 the existence of a random attractor follows from the existence of a compact absorbing set for the corresponding random dynamical system. To derive the existence of a bounded absorbing set, we impose certain regularity assumptions on the noise and growth conditions on the reaction terms in (0.11). Due to the absence of the regularizing effect of the Laplacian in the second component we have to perform a certain splitting technique for the compactness argument.

In Section 5.3 we perform a similar analysis for system (0.11), however, this time we consider a *multiplicative perturbation* by a Brownian motion in the Stratonovich sense.

In Section 5.5 we consider the *stochastic Field-Noyes system*, again with a *multiplicative perturbation* by a Brownian motion in the Stratonovich sense. This reaction-diffusion system arrives in chemical kinetics and possesses a non-linear coupling between components (which is not covered by the analysis in the previous sections). We show that non-negativity is preserved under the flow and make explicit use of this information in order to derive a random attractor in this setting.

Note that Chapter 5 is based on joint works with Christian Kuehn and Alexandra Neamţu and that results in Section 5.2 were jointly published in [KNP20].

In **Chapter 6** we are concerned with a finer resolution of the dynamical behaviour of stochastic reaction-diffusion equations compared to the existence proofs of random attractors in the previous chapter. Here, we focus on *fast-slow systems*, that is, coupled systems where different components evolve on widely different time scales. Suppose that a deterministic fast-slow ODE system exhibits a *hyperbolic attracting slow manifold*, then there exists an exhaustive theory on how sample paths of the corresponding SODE system (perturbed by Brownian motion) behave close to this manifold. Namely, there exists exponential estimates on the probability that sample paths leave a certain neighbourhood around the manifold. We would like to extend this theory to fast-slow SPDEs of reaction-diffusion type and thereby contribute to a finer characterisation of the dynamical behaviour of these systems. As a first step towards this goal, we consider in Chapter 6 a linear SPDE with a non-autonomous reaction term on a bounded domain. In our approach we use a finite-dimensional *Galerkin approximation* of this equation, which can be treated by the corresponding SODE theory, and then we pass in a suitable way to the limit with infinite modes. By this means, we are able to show that, similar to the finite-dimensional setting, the probability of a sample path to leave a certain neighbourhood around the corresponding deterministic slow manifold is exponentially small in the size of this neighbourhood.

Chapter 6 is based on joint work with Manuel Gnann and Christian Kuehn; and results in Section 6.4 were jointly published in [GKP19].

Finally, notations are listed after Chapter 6 and some classical results are stated in the appendices, to which we may refer throughout the thesis. Note that constants may always change from line to line if their precise value is not relevant. If not mentioned otherwise constants are assumed to be positive and finite.

Chapter 1

Random trees

Random trees are fundamental objects in probability theory. In the following we give a brief overview on this topics, where we restrict ourselves to *discrete trees*. The main purpose of this chapter is to define critical Galton-Watson trees conditioned to survive with an offspring distribution in the domain of attraction of a stable law and to derive several estimates for these objects, which will be needed in the following Chapter 2. This chapter is based on joint work with Eleanor Archer.

1.1 Discrete trees

Discrete, plane, rooted trees are connected graphs with no cycles where one vertex is designated as the root of the tree. They can be defined with the so-called *Ulam-Harris formalism* introduced in [Nev86]. For that, let us define

$$\mathcal{U} := \bigcup_{n\geq 0} \mathbb{N}^n,$$

with the convention that $\mathbb{N}^0 = \{\emptyset\}$. Furthermore, we set for $u = (u_1, ..., u_n) \in \mathcal{U}$, $n \geq 1$, $|u| = n$ and $|\emptyset| = 0$. We also define the concatenation of two elements $u = (u_1, ..., u_n), v = (v_1, ..., v_m) \in \mathcal{U}$ as $uv = (u_1, ..., u_n, v_1, ..., v_m)$. If $u = \emptyset$ respectively $v = \emptyset$ we set $uv = v$ respectively $uv = u$. We call v an *ancestor* of u if there exists $\emptyset \neq w \in \mathcal{U}$ such that $u = vw$. This genealogical relation is denoted as $v \prec u$.

Definition 1.1. A tree T is a subset of $\mathcal{U}$ such that

(i) $\emptyset \in T$,

(ii) If $v \in T$ and $v = uj$ for some $j \in \mathbb{N}$, then also $u \in T$,

(iii) For every $u \in T$ there exists $k_u \in \mathbb{N}_0 \cup \{\infty\}$ such that $uj \in T$ if and only if $1 \leq j \leq k_u$.

The vertex $\emptyset$ is the *root* of the tree T, which we will also often denote as O in the following. The number k_u denotes the number of *offspring* of the vertex u, in particular its *degree* is given by $\deg(u) = k_u + 1$. A vertex $u \in T$ is a *leaf* if $k_u = 0$. Furthermore, for $u \in T$ we define the *subtree* T_u of T as

$$T_u := \{v \in \mathcal{U} : uv \in T\}.$$

We also impose a *lexicographic order* on T, that is, for $u, v \in T$, we say $v < u$ if either $v \prec u$ or $v = wjv'$ and $u = wiu'$ with $j < i$ for some $i, j \in \mathbb{N}$ and $w, u', v' \in \mathcal{U}$. We define the *height* and the *generation size at level* n of a tree T as

$$\mathsf{Height}(T) := \sup\{|u|, u \in T\}, \quad z_n(T) := \#\{u \in T : |u| = n\},$$

where $\#A$ denotes the cardinality of a set A. Finally, let us set $\mathbb{T}$ to be the set of all discrete rooted trees as defined above.

The following figure depicts a finite tree with Ulam-Harris labelling.

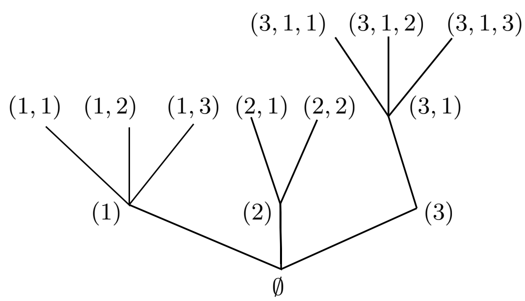

Figure 1.1: Finite tree T with Ulam-Harris labelling, note that $\mathsf{Height}(T) = 3$ and $z_2(T) = 6$.

1.2 Galton-Watson trees

Random trees are trees formed by stochastic processes. In this section we will introduce the canonical example of a random tree, the *Galton-Watson tree*, which describes the genealogy of a *Galton-Watson process*. These processes were primarily studied by the French scientist I.-J. Bienaymé (1796-1878) and the British scientists F. Galton (1822-1911) and H. W. Watson (1827-1903). We refer to [Har63] and [AN12] for detailed presentations of this topic.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space.

Definition 1.2 (Galton-Watson process (GWP))**.** Let $(Z_n)_{n\geq 0}$ be a sequence of integer-valued random variables, recursively defined by

$$Z_n := \sum_{k=1}^{Z_{n-1}} X_{n,k}, \quad n \geq 1,$$

where $\{X_{n,k} : n, k \geq 1\}$ are iid integer-valued random variables with common distribution $(p_n)_{n\geq 0}$ and independent of Z_0. Then $(Z_n)_{n\geq 0}$ is called *Galton-Watson process* (GWP) with *offspring distribution* $(p_n)_{n\geq 0}$ and Z_0 ancestors.

Let X have the offspring distribution $(p_n)_{n\geq 0}$, i.e. $\mathbf{P}(X = n) = p_n$ and let $m := \mathbf{E}[X] = \sum_{n=1}^{\infty} np_n$ be its mean. We call the corresponding Galton-Watson process *sub-critical* if $m < 1$, *critical* if $m = 1$ and *super-critical* if $m > 1$.

Remark 1.3. Every GWP with $Z_0 = k$ ancestors is the sum of k independent GWPs with $Z_0 = 1$ ancestors and the same offspring distribution. We always consider the case $Z_0 = 1$ from now on.

For each $n \geq 0$, the random variable Z_n is interpreted as the size of the nth generation of a given population. The population is called *extinct* at generation n if $Z_n = 0$ and we define the *extinction event*

$$\mathsf{Ext} := \{\exists n \in \mathbb{N} : Z_n = 0\} = \lim_{n\to\infty} \{Z_n = 0\}.$$

Note the following fundamental theorem.

Theorem 1.4 (see e.g. [AN12]). *We consider a GWP with offspring distribution $(p_n)_{n\geq 0}$ and we assume*

$$0 < p_0 \leq p_0 + p_1 < 1. \tag{1.1}$$

If $m \leq 1$ we have $\mathbf{P}(\mathsf{Ext}) = 1$, i.e. the process is almost surely extinct. If $m > 1$ we have $\mathbf{P}(\mathsf{Ext}) = q$ with $0 < q < 1$, i.e. the process has a positive probability to not get extinct (i.e. to survive).

Remark 1.5.

(i) The proof is based on the fact that the extinction probability is a fixed point of the generating function of the corresponding process. The statement can then be derived easily by using elementary properties of the generating function.

(ii) In the trivial cases $p_0 = 0$ and $p_0 = 1$ it holds $\mathbf{P}(\mathsf{Ext}) = 0$ and $\mathbf{P}(\mathsf{Ext}) = 1$, respectively.

(iii) If $0 < p_0 < 1$ and $p_0 + p_1 = 1$ it holds $\mathbf{P}(\mathsf{Ext}) = 1$.

Definition 1.6 (Galton-Watson tree). A *Galton-Watson tree* T with offspring distribution $p = (p_n)_{n\geq 0}$ and root O is a $\mathbb{T}$-valued random variable with law $\mathbf{P}_p$ satisfying the following properties

(i) $\mathbf{P}_p(k_O = n) = p_n$ for all $n \geq 0$.

(ii) Let $v_1, v_2...$ denote the offspring of O, then the subtrees $T_{v_1}, ..., T_{v_n}$ are independent under the conditional probability $\mathbf{P}_p(\cdot|k_O = n)$ and distributed as the original tree T, i.e. with law $\mathbf{P}_p$.

Remark 1.7.

(i) For any probability measure p on $\mathbb{N}_0$ there exists a unique probability measure $\mathbf{P}_p$ on the set of plane trees satisfying the properties in Definition 1.6, see [Nev86].

(ii) Let T be a Galton-Watson tree with root O and law $\mathbf{P}_p$. Let $(Z_n)_{n\geq 0}$ be a Galton-Watson process with offspring distribution p and a single ancestor. Then $(z_n(T))_{n\geq 0}$ and $(Z_n)_{n\geq 0}$ have the same distribution.

(iii) In other words, a Galton-Watson tree T can be associated to a Galton-Watson process $(Z_n)_{n\geq 0}$ with offspring distribution $(p_n)_{n\geq 0}$ and single ancestor. The construction goes as follows: Start with the ancestor (or root) and suppose that individuals in a given generation have offspring independently of the past and of each other according to the distribution $(p_n)_{n\geq 0}$. The vertex set of T is the entire collection of individuals, edges are the parent-offspring bonds and Z_n is the number of individuals in the nth generation of T.

(iv) We will omit the subscript p in $\mathbf{P}_p$.

1.3 Offspring distribution in the domain of attraction of a stable law

In the following we recall the notion of distributions in the *domain of attraction* of a *β-stable distribution*. We refer to [Fel71] for further details.

Definition 1.8 (β-stable distribution)**.** Let μ be a probability distribution on $\mathbb{R}$ and let $X_1, X_2, ...$ be iid random variables with distribution μ. Then we call μ *β-stable with index* $\beta \in (0, 2]$ if there exist a sequence $(b_n)_{n\geq 1}$ such that

$$X_1 + X_2 + ... + X_n \stackrel{\mathcal{D}}{=} n^{1/\beta} X_1 + b_n, \quad \text{for all } n \in \mathbb{N}. \tag{1.2}$$

μ is called *strictly* β-stable if (1.2) holds with $b_n = 0$ for all $n \in \mathbb{N}$.

Remark 1.9.

(i) If $\beta = 2$, then μ is a *Gaussian distribution*.

(ii) If $\beta = 1$, then $b_n = (c^+ - c^-)n \log(n)$, for $n \in \mathbb{N}$, $c^+, c^- \geq 0$. If additionally $c^+ = c^-$, then μ is a *Cauchy distribution*.

It turns out that stable distributions are those that appear as limits of sums of iid random variables.

Definition 1.10 (Domain of attraction)**.** Let μ be a probability distribution on $\mathbb{R}$ that is not concentrated in one point and let Z be a random variable with distribution μ. Let $X, X_1, X_2, ...$ be iid random variables with distribution $\mathbf{P}_X$ and set $S_n := X_1 + X_2 + ... + X_n$. Then $\mathbf{P}_X$ belongs to the *domain of attraction of* μ, denoted as $\mathsf{Dom}(\mu)$, if there exist sequences of real numbers $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ such that

$$\frac{S_n - b_n}{a_n} \to Z \quad \text{in distribution for } n \to \infty.$$

If μ is stable with index $\beta \in (0, 2]$ and one can choose $a_n = n^{1/\beta}$ we say that $\mathbf{P}_X$ lies in the *domain of normal attraction* of μ.

Proposition 1.11 (see [Fel71, VI.1])**.** *Let μ be a probability distribution on $\mathbb{R}$ that is not concentrated in one point. It holds $\mathsf{Dom}(\mu) \neq \emptyset$ if and only if μ is stable. Then $\mu \in \mathsf{Dom}(\mu)$.*

Remark 1.12. If, in the setting of Definition 1.10, $\mathbf{E}[X] < \infty$ and $\mathbf{E}[X^2] < \infty$ then $\mathbf{P}_X$ lies in the domain of normal attraction of the Gaussian distribution ($\beta = 2$) by the classical *central limit theorem.*

A useful characterization is given in the following.

Theorem 1.13 (cf. [Fel71, XVII.5 Corollary 2])**.** *Let $0 < \beta < 2$. The distribution of a positive random variable X belongs to the domain of attraction of a β-stable distribution if and only if the tail probability $\mathbf{P}(X > x)$ varies regularly with exponent $-\beta$ as $x \to \infty$, i.e.*

$$\mathbf{P}(X > x) \sim x^{-\beta} L(x),$$

where $L(x)$ varies slowly (see Definition B.7).

Proposition 1.14 (see [Fel71, XVII.5])**.** *If $\mathbf{P}_X$ lies in the domain of attraction of a stable distribution with index β, then $\mathbf{E}[|X|^\gamma] < \infty$ for all $\gamma \in (0, \beta)$ and $\mathbf{E}[|X|^\gamma] = \infty$ in case $\gamma \geq \beta$ and $\beta < 2$.*

Now, let $(p_n)_{n\geq 0}$ be an offspring distribution with $m = \sum_{n=0}^{\infty} p_n n = 1$ (critical) in the domain of attraction of a stable law μ with index $\beta \in (1, 2)$. We will ignore slowly varying fluctuations in the offspring distribution, that is, we always assume that the function $L(x)$ in Theorem 1.13 is equal to a positive constant. Let $(Z_n)_{n\geq 0}$ be the corresponding *Galton-Watson process* started from $Z_0 = 1$. Define the *non-extinction probability* of the nth generation $q_n := \mathbf{P}(Z_n > 0)$, then by [Sla68, Lemma 2]

$$q_n^{\beta-1} c \sim \frac{1}{(\beta - 1)n}, \quad \text{for some } c > 0. \tag{1.3}$$

In this setting the following asymptotic behaviour can be established for the associated critical Galton-Watson tree T . Note that $V(T)$ stands for the volume of the tree, i.e. the total number of vertices it contains.

Lemma 1.15. *There exist constants $c, c' < \infty$ such that, as $n \to \infty$,*

$$\mathbf{P}(V(T) \geq n) \sim cn^{\frac{-1}{\beta}}, \tag{1.4}$$

$$\mathbf{P}(\mathit{Height}(T) \geq n) \sim c'n^{\frac{-1}{\beta-1}}. \tag{1.5}$$

Proof. The relation (1.4) has been proved in [Kor12, Lemma 1.11] and the relation (1.5) follows from (1.3). □

Some simulations by I. Kortchemski of critical Galton-Watson trees with offspring distributions in the domain of attraction of a β-stable law can be found in Figure 1.2.

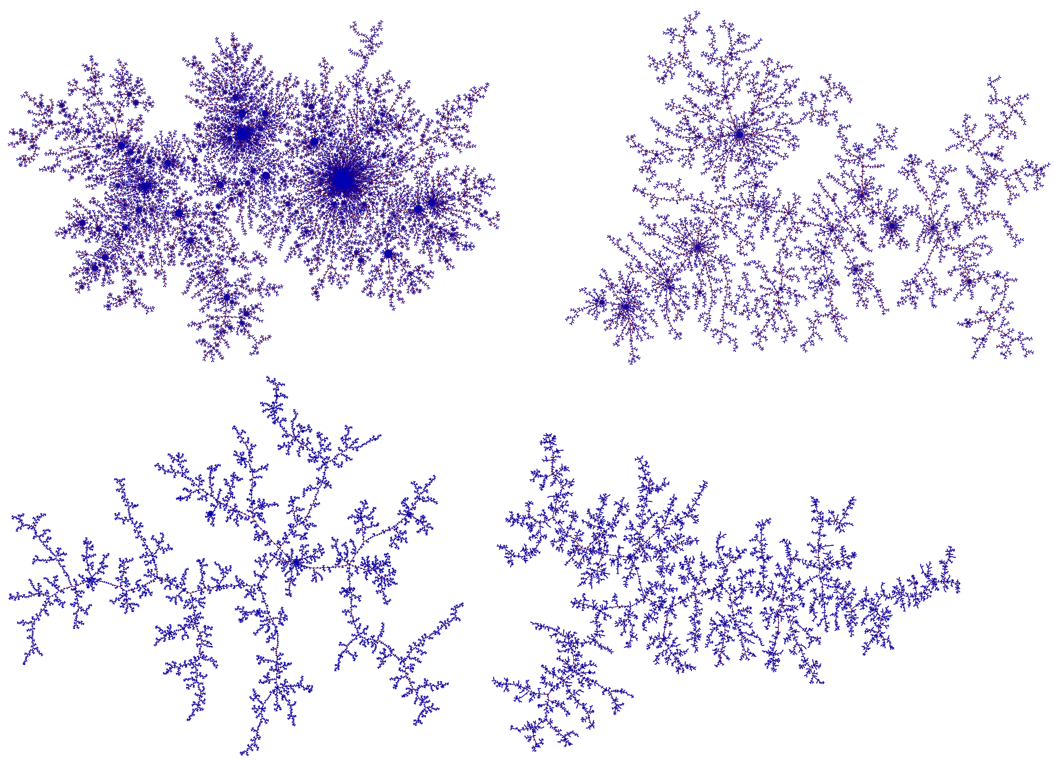

Figure 1.2: Simulations by I. Kortchemski of Galton-Watson trees, conditioned on having a large fixed number of vertices, with a critical offspring distribution in the domain of attraction of an β-stable law with $\beta = 1.1,\ 1.5,\ 1.9,\ 2$ (from top to bottom and left to right), see `https://igor-kortchemski.perso.math.cnrs.fr/images.html`.

Remark 1.16. Recall that $\mathbf{P}(X \geq n) \sim n^{-\beta}$ if and only if $\lim_{n\to\infty} \frac{\mathbf{P}(X\geq n)}{n^{-\beta}} = 1$. In particular, this means that for every $\varepsilon > 0$ there exists an N_ε such that for all

$n \geq N_\varepsilon$

$$n^{-\beta}(1-\varepsilon) \leq \mathbf{P}(X \geq n) \leq (1+\varepsilon)n^{-\beta}.$$

Furthermore, defining

$$c_1 := \min\left\{1-\varepsilon, \min_{0<n<N_\varepsilon}\left\{\mathbf{P}(X \geq n)n^\beta\right\}\right\},$$
$$c_2 := \max\left\{1+\varepsilon, \max_{0<n<N_\varepsilon}\left\{\mathbf{P}(X \geq n)n^\beta\right\}\right\},$$

then $\mathbf{P}(X \geq n) \sim n^{-\beta}$ also implies that for all $n > 0$

$$n^{-\beta}c_1 \leq \mathbf{P}(X \geq n) \leq c_2 n^{-\beta}.$$

We will often use these kinds of bounds in the following.

Furthermore, we will also need the following lemmas later on.

Lemma 1.17 (see [Arc20, Lemma 3.1])**.** *Let $(X_i)_{i\geq 1}$ be iid random variables where $\mathbf{P}(X_1 \geq x) \sim cx^{-\beta}$ for some $\beta \leq 1$. Then*

(i) If $\beta < 1$, then there exists a constant $c' < \infty$ such that for each $n \geq 2$

$$\mathbf{P}\left(\sum_{i=1}^{n} X_i \geq n^{1/\beta}\lambda\right) \leq c'\lambda^{-\beta},$$

as $n \to \infty$.

(ii) If $\beta = 1$, then there exists a constant $c' < \infty$ such that for each $n \geq 2$

$$\mathbf{P}\left(\sum_{i=1}^{n} X_i \geq n^{1/\beta}\lambda\log(n)\right) \leq c'\lambda^{-\beta},$$

as $n \to \infty$.

Lemma 1.18. *Let X be a random variable and $\beta < 1$ such that $\mathbf{P}(X \geq x) \sim c'x^{-\beta}$ as $x \to \infty$. Then there exists $0 < c < \infty$ such that*

$$1 - \mathbf{E}[\exp\{-\theta X\}] \sim c\theta^\beta \text{ as } \theta \to 0.$$

Proof. The statement follows from [Kor04, Theorem 8.2, Section IV]. □

1.4 Kesten's tree

In this subsection we will state Kesten's construction of infinite Galton-Watson trees [Kes86]. Note that in case of a super-critical offspring distribution one can easily define an infinite tree by simply conditioning the corresponding standard Galton-Watson tree on surviving. However, this conditioning does not make sense for sub-critical or critical Galton-Watson trees since their probability to survive is zero. Thus, from now on we assume that $p = (p_n)_{n\geq 0}$ is an offspring distribution with mean $m \leq 1$. We define its size-biased version $p^* = (p_n^*)_{n\geq 0}$ as

$$p_n^* := \frac{np_n}{m}, \quad \text{for } n \geq 0.$$

Definition 1.19 (Kesten's tree, cf. e.g. [AD14a]). A *Kesten's tree* T^* associated to the probability distribution p is a *two-type Galton-Watson tree* distributed as follows:

- Each individual is either of type *normal* or of type *special*.
- The root O of T^* is special, we call it s_0.
- A normal individual produces normal individuals according to p.
- A special individual produces offspring according to the size-biased distribution p^*. If the number of children is finite, one of them is chosen uniformly at random. This individual is of type special, the rest of the produced individuals are of type normal. We denote the special vertex of the nth generation as s_n. If the number of children is infinite, then all children are normal.

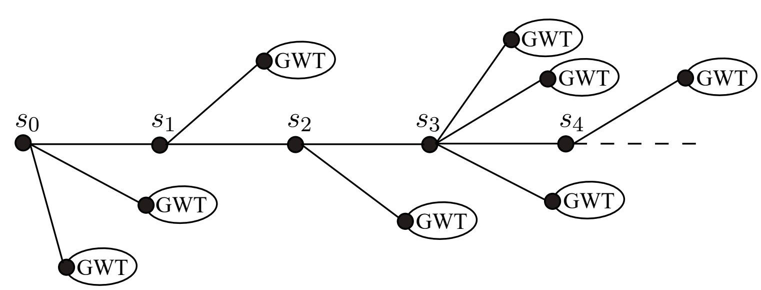

Figure 1.3: Sketch of *Kesten's tree* with critical offspring distribution. Unconditioned critical Galton-Watson trees (GWT's) grow out of the infinite backbone $s_0, s_1, s_2, ...$

Let us now restrict to the case $m = 1$, i.e. critical offspring distributions. Then, $\mathbf{P}$-almost surely the special vertices form a unique one-ended infinite backbone of T^*. In other words $T^* \in \mathbb{T}_\infty$, where $\mathbb{T}_\infty$ denotes the set of plane rooted trees with a single (one-ended) infinite path to infinity. Moreover, each of the subtrees emanating from the children of the backbone vertices are independent unconditioned Galton-Watson trees with offspring distribution $(p_n)_{n\geq 0}$. We let $s_0, s_1, s_2, \ldots$ denote the ordered vertices of the backbone, so that s_0 is the root, and s_n is at distance n from the root, see Figure 1.3.

Let us define

$$\mathbb{T}^{(h)} := \{T \in \mathbb{T} : \mathsf{Height}(T) \leq h\},$$

and the restriction function from $\mathbb{T}$ to $\mathbb{T}^{(h)}$, that is for $T \in \mathbb{T}$

$$r_h(T) := \{u \in T, |u| \leq h\}.$$

Then a sequence of random trees $(T_n)_{n\geq 1}$ converges in distribution with respect to the so-called local topology towards a random tree T, denoted as $T_n \overset{(d)}{\to} T$, if and only if

$$\forall h \in \mathbb{N}, \forall t \in \mathbb{T}^{(h)}, \quad \lim_{n\to\infty} \mathbf{P}(r_h(T_n) = t) = \mathbf{P}(r_h(T) = t).$$

Kesten proved that Galton-Watson trees with critical offspring distribution $(p_n)_{n\geq 0}$ conditioned on reaching at least height n (surviving until generation n) converge in distribution to Kesten's tree with distribution $(p_n)_{n\geq 0}$ for n going to infinity. That is Kesten's tree is the local limit in distribution of such conditioned Galton-Watson trees. More precisely,

Theorem 1.20 (Kesten's theorem, see [Kes86, Jan12])**.** *Let $p = (p_n)_{n\geq 0}$ be a critical offspring distribution that satisfies* (1.1)*. Let T be a Galton-Watson tree with offspring distribution p and T^* Kesten's tree associated to p. Furthermore, for every n let T_n be a random tree distributed as T conditionally on having height at least n. Then*

$$T_n \overset{(d)}{\to} T^*,$$

in the local topology as defined above.

Remark 1.21.

(i) It can be shown that conditioning critical Galton-Watson trees on different events can also lead to convergence in distribution to Kesten's tree, see [AD14b] for details.

(ii) In the sub-critical case $m < 1$, $\mathbf{P}$-almost surely the spine of Kesten's tree is finite and one vertex has infinite degree. Similar local limit theorems can be proved, we refer to [Jan12, AD14a] for further details.

1.5 Critical Galton-Watson trees conditioned to survive T_∞

This section is based on joint work with Eleanor Archer. We consider here the critical Galton-Watson tree conditioned to survive (as defined above) with root O and with an offspring distribution $(p_n)_{n\geq 0}$ that is in the domain of attraction of a β-stable law μ with $\beta \in (1,2]$. Furthermore, for simplicity, we assume that the slowly varying function $L(x)$ in Theorem 1.13 is equal to some constant $c > 0$. We denote this tree as T_∞.

In the following we will denote the number of offspring of a vertex v, i.e. k_v, also as Off(v). Note that the size-biased distribution $(p^*_n)_{n\geq 0}$ belongs to the domain of attraction of the same stable law with index $\beta - 1$.

- In particular, in case $\beta \in (1,2)$, $(p_n)_{n\geq 0}$ has finite mean and infinite variance and for a normal vertex $v \in T_\infty$ we have

$$\mathbf{P}(\mathrm{Off}(v) \geq n) \sim cn^{-\beta}. \tag{1.6}$$

  Furthermore, in that case $(p^*_n)_{n\geq 0}$ has infinite mean and for a special vertex s, we can conclude that

$$\mathbf{P}(\mathrm{Off}(s) \geq n) \sim cn^{-(\beta-1)}. \tag{1.7}$$

- Now, in case $\beta = 2$, $(p_n)_{n\geq 0}$ has finite mean and either infinite variance, in which case for a normal vertex v

$$\mathbf{P}(\mathrm{Off}(v) \geq n) \sim cn^{-2},$$

  or finite variance in which case (from Markov's inequality) we know

$$\mathbf{P}(\mathrm{Off}(v) \geq n) = O(n^{-2}).$$

  Likewise, the distribution $(p^*_n)_{n\geq 0}$ has then either infinite mean and for a special vertex we have

$$\mathbf{P}(\mathrm{Off}(s) \geq n) \sim cn^{-1},$$

  or it has finite mean and

$$\mathbf{P}(\mathrm{Off}(s) \geq n) = O(n^{-1}).$$

We will restrict from now on for the rest of the chapter to the case $\beta \in (1,2)$. That is, equations (1.6) and (1.7) are the relevant identities for the tails of the offspring distribution. This way we avoid the inclusion of certain logarithmic correction terms, which would appear for example when applying Lemma 1.17 (ii) for the size-biased offspring distribution in the $\beta = 2$ setting, and thus we will not have to distinguish different cases throughout the analysis. Furthermore, note that for any $v \neq O$ we have $\deg(v) = \mathrm{Off}(v) + 1$ and therefore, for a normal vertex v, using (1.6), for $n \to \infty$

$$\frac{\mathbf{P}(\deg(v) > n)}{cn^{-\beta}} = \frac{\mathbf{P}(\mathrm{Off}(v) > n-1)}{c(n-1)^{-\beta}} \frac{1}{(1-1/n)^{\beta}} \to 1,$$

that is, the same kind of tail decay holds for the distribution of the degree. Likewise, the tail decay of the degree distribution for a special vertex s is given by (1.7).

In the following we will derive several estimates for this tree model, which we will need in Chapter 2.

1.5.1 Properties of T_∞

Given $r > 0$, we denote by B_r the closed ball of radius r around the root in T_∞ and by $V(B_r)$ the volume of this ball (i.e. the number of vertices it contains). If the ball with radius r in T_∞ is centred at $v \neq O$ it is denoted as $B(v, r)$. Furthermore, we denote by A_r the connected component containing O obtained after removing the vertex s_r from T_∞. We also let $\mathrm{Sp}(r)$ denote the set of offspring of the vertices $s_0, s_1, \ldots, s_{r-1}$ and $V(\mathrm{Sp}(r))$ its volume. Moreover, we define $\mathrm{Sp}^*(r)$ as the set of the offspring of the vertices $s_0, ..., s_{r-1}$ excluding the vertices $s_1, ..., s_r$, again $V(\mathrm{Sp}^*(r))$ denotes the volume of this set. More generally, if $A \subset T_\infty$, we let $V(A)$ denote the number of vertices in A. For the subtree T_v, $v \in T_\infty$, we denote by $B_r^{T_v}$ the ball of radius r around v in T_v. We derive the following volume estimates.

Proposition 1.22. *Let $\beta \in (1, 2)$. Take $\lambda, r > 1$. Then for any $\varepsilon > 0$, there exist constants $0 < C < \infty$ (may be different for each of the following bounds) such that*

(i) $\mathbf{P}\Big(V(B_r) \geq \lambda r^{\frac{\beta}{\beta-1}}\Big) \leq C\lambda^{-(\beta-1-\varepsilon)}$,

(ii) $\mathbf{P}\Big(V(B_r) \geq \lambda r^{\frac{\beta}{\beta-1}}\Big) \geq C\lambda^{-\frac{(\beta-1+\varepsilon)}{2-\beta}}$, *for r sufficiently large,*

(iii) $\mathbf{P}\Big(V(A_r) \geq \lambda r^{\frac{\beta}{\beta-1}}\Big) \leq C\lambda^{\frac{-(\beta-1)}{\beta^2}}$,

(iv) $\mathbf{P}\Big(V(B_r) \leq \lambda^{-1} r^{\frac{\beta}{\beta-1}}\Big) \leq Ce^{-c\lambda^{\frac{\beta-1}{\beta}}}$, *for r, λ sufficiently large.*

Proof. (i) and (ii) are taken from [CK08, Proposition 2.5].

(iii) Due to the construction of T_∞ (recall Subsection 1.4) we can write

$$V(A_r) = r + \sum_{v \in \mathrm{Sp}^*(r)} V(T_v) \leq \sum_{v \in \mathrm{Sp}(r)} V(T_v), \tag{1.8}$$

where $(T_v)_{v\in \mathrm{Sp}(r)}$ is a collection of independent copies of (unconditioned) Galton-Watson trees, so that $\mathbf{P}(V(T_v)\geq n)\sim cn^{\frac{-1}{\beta}}$ for $v\in \mathrm{Sp}(r)$ by Lemma 1.15. Moreover, the degree distribution on the backbone is size-biased, so we have

$$V(\mathrm{Sp}(r))=\sum_{i=0}^{r-1}\mathrm{Off}(s_i),$$

where $\mathrm{Off}(s_i)$ denotes the number of offspring of vertex s_i and recall from (1.7) that $\mathbf{P}(\mathrm{Off}(s_i)\geq x)\sim cx^{-(\beta-1)}$ for $\beta\in(1,2)$. Then, since $r<r^{\frac{\beta}{\beta-1}}\lambda$ for all r, we have by (1.8) that for $\beta\in(1,2)$

$$\begin{aligned}
&\mathbf{P}\Big(V(A_r)\geq r^{\frac{\beta}{\beta-1}}\lambda\Big)\\
&\leq \mathbf{P}\left(\sum_{v\in \mathrm{Sp}(r)}V(T_v)\geq r^{\frac{\beta}{\beta-1}}\lambda\right)\\
&\leq \mathbf{P}\Big(V(\mathrm{Sp}(r))\geq r^{\frac{1}{\beta-1}}\lambda^{\frac{1}{\beta^2}}\Big)\\
&\qquad+\mathbf{P}\left(\sum_{v\in \mathrm{Sp}(r)}V(T_v)\geq r^{\frac{\beta}{\beta-1}}\lambda^{\frac{1}{\beta}}\lambda^{\frac{\beta-1}{\beta}}\,\Big|\,V(\mathrm{Sp}(r))\leq r^{\frac{1}{\beta-1}}\lambda^{\frac{1}{\beta^2}}\right)\\
&\leq C\lambda^{\frac{-(\beta-1)}{\beta^2}},
\end{aligned}$$

where the bounds in the final line follow from Lemma 1.17 (i).

(iv) Assume for convenience that r is even. Again, by the construction of T_∞ the subtrees T_v emanating from each vertex $v\in \mathrm{Sp}^*(r/2)$ are independent Galton-Watson trees with offspring distribution $(p_n)_{n\geq 0}$. Therefore,

$$V(B_r)\geq \sum_{v\in \mathrm{Sp}^*(r/2)}V(B^{T_v}_{r/2}).$$

Moreover, again by Lemma 1.15, for each $v\in \mathrm{Sp}^*(r/2)$ we have $\mathbf{P}(V(T_v)\geq n)\sim cn^{-\frac{1}{\beta}}$. Furthermore,

$$V(\mathrm{Sp}^*(r/2))=\sum_{i=0}^{r/2-1}(\mathrm{Off}(s_i)-1), \tag{1.9}$$

i.e. $V(\mathrm{Sp}^*(r/2))$ is given by the sum of $\frac{1}{2}r$ random variables in the domain of attraction of a $(\beta-1)$-stable law, compare with (iii). Thus for some $\theta>0$, using Markov's inequality and the law of total expectation, we have

$$\begin{aligned}
&\mathbf{P}\Big(V(B_r) \le r^{\frac{\beta}{\beta-1}}\lambda^{-1}\Big) \\
&\le \mathbf{P}\left(\sum_{v\in \mathrm{Sp}^*(r/2)} V(B^{T_v}_{r/2}) \le r^{\frac{\beta}{\beta-1}}\lambda^{-1}\right) \\
&\le \mathbf{E}\left[\exp\left\{-\theta \sum_{v\in \mathrm{Sp}^*(r/2)} V(B^{T_v}_{r/2})\right\}\right] \exp\left\{\theta r^{\frac{\beta}{\beta-1}}\lambda^{-1}\right\} \\
&\le \mathbf{E}\left[\left(\mathbf{E}[\exp\{-\theta V(T_v)\}] + \mathbf{P}\Big(\mathsf{Height}(T_v) \ge \frac{1}{2}r\Big)\right)^{V(\mathrm{Sp}^*(r/2))}\right] \exp\left\{\theta r^{\frac{\beta}{\beta-1}}\lambda^{-1}\right\} \\
&\le \mathbf{E}\left[\left(\mathbf{E}[\exp\{-\theta \tilde{Z}_1\}] + c r^{\frac{-1}{\beta-1}}\right)^{Z_1}\right]^{\frac{1}{2}r} \exp\left\{\theta r^{\frac{\beta}{\beta-1}}\lambda^{-1}\right\},
\end{aligned}$$

where $\tilde{Z}_1$ has $\frac{1}{\beta}$-stable tails and Z_1 has $(\beta-1)$-stable tails. Moreover, it follows for such random variables from Lemma 1.18 that $\mathbf{E}[\exp\{-\theta \tilde{Z}_1\}] \le 1 - c'\theta^{\frac{1}{\beta}}$ and $\mathbf{E}[\exp\{-\phi Z_1\}] \le Ce^{-c\phi^{\beta-1}}$ for all sufficiently small θ, ϕ and some constants $C, c, c' > 0$. Therefore, taking $\theta = r^{\frac{-\beta}{\beta-1}}\lambda$ and assuming $\lambda \ge \left(2\frac{c}{c'}\right)^{\beta}$ yields for r sufficiently large

$$\begin{aligned}
\mathbf{P}\Big(V(B_r) \le r^{\frac{\beta}{\beta-1}}\lambda^{-1}\Big) &\le \mathbf{E}\left[\left(\mathbf{E}[\exp\{-\theta \tilde{Z}_1\}] + c r^{\frac{-1}{\beta-1}}\right)^{Z_1}\right]^{\frac{1}{2}r} \exp\left\{\theta r^{\frac{\beta}{\beta-1}}\lambda^{-1}\right\} \\
&\le \mathbf{E}\left[\left(1 - c' r^{\frac{-1}{\beta-1}}\lambda^{\frac{1}{\beta}} + c r^{\frac{-1}{\beta-1}}\right)^{Z_1}\right]^{\frac{1}{2}r} e \\
&\le \mathbf{E}\left[\left(\exp\left\{-\frac{c'}{2} r^{\frac{-1}{\beta-1}}\lambda^{\frac{1}{\beta}}\right\}\right)^{Z_1}\right]^{\frac{1}{2}r} e \\
&\le e\left(C' e^{-cr^{-1}\lambda^{\frac{\beta-1}{\beta}}}\right)^{\frac{1}{2}r} \\
&\le C'' e^{-c\lambda^{\frac{\beta-1}{\beta}}}.
\end{aligned}$$

□

Corollary 1.23. *Let $\beta \in (1,2)$. $\mathbf{P}$-almost surely for any $\varepsilon > 0$ it holds*

(i)

$$\limsup_{n\to\infty} \frac{V(B_r)}{r^{\frac{\beta}{\beta-1}}(\log r)^{\frac{1+\varepsilon}{\beta-1}}} = 0,$$

(ii)

$$\limsup_{r\to\infty} \frac{V(A_r)}{r^{\frac{\beta}{\beta-1}}(\log r)^{\frac{\beta^2+\varepsilon}{\beta-1}}} = 0,$$

(iii)

$$\liminf_{r\to\infty} \frac{V(B_r)}{r^{\frac{\beta}{\beta-1}}(\log\log r)^{-\frac{\beta}{\beta-1}}} > 0.$$

Proof. We consider the first identity. Let $\delta > 0$ be arbitrary, set $r_n := 2^n$ and $\lambda_n := \frac{\delta}{C}(\log r_n)^{\frac{1+\varepsilon}{\beta-1-\varepsilon}}$ with $C := 2^{\frac{\beta}{\beta-1}+\frac{1+\varepsilon}{\beta-1-\varepsilon}}$. Then we compute using Proposition 1.22 (i) that

$$\sum_{n=0}^{\infty} \mathbf{P}\left(V(B_{r_n}) \geq \lambda_n r_n^{\frac{\beta}{\beta-1}}\right) \leq C' \sum_{n=0}^{\infty} (\log r_n)^{-(1+\varepsilon)} < \infty.$$

Thus by Borel-Cantelli (see Lemma A.1) we can conclude that along the sequence r_n eventually almost surely

$$V(B_{r_n}) < \frac{\delta}{C} r_n^{\frac{\beta}{\beta-1}} (\log r_n)^{\frac{1+\varepsilon}{\beta-1-\varepsilon}} < \delta r_n^{\frac{\beta}{\beta-1}} (\log r_n)^{\frac{1+\varepsilon}{\beta-1-\varepsilon}}.$$

Furthermore, for $r \in [r_n, r_{n+1}]$ eventually almost surely

$$\begin{aligned} V(B_r) < V(B_{r_{n+1}}) &< \frac{\delta}{C} 2^{\frac{\beta}{\beta-1}} r_n^{\frac{\beta}{\beta-1}} (\log r_n + \log(2))^{\frac{1+\varepsilon}{\beta-1-\varepsilon}} \\ &\leq \frac{\delta}{C} 2^{\frac{\beta}{\beta-1}+\frac{1+\varepsilon}{\beta-1-\varepsilon}} r_n^{\frac{\beta}{\beta-1}} (\log r_n)^{\frac{1+\varepsilon}{\beta-1-\varepsilon}} \\ &\leq \delta r^{\frac{\beta}{\beta-1}} (\log r)^{\frac{1+\varepsilon}{\beta-1-\varepsilon}}. \end{aligned}$$

That is, for any $\delta > 0$

$$\frac{V(B_r)}{r^{\frac{\beta}{\beta-1}}(\log r)^{\frac{1+\varepsilon}{\beta-1-\varepsilon}}} < \delta.$$

Hence, the first identity follows for some $\varepsilon' > 0$ that we rename ε. The other two identities of the Corollary follow in a similar manner from (iii) and (iv) of the previous Proposition 1.22. □

Proposition 1.24. *Let Z_r^* denote the size of the rth generation of T_∞. For every $\varepsilon > 0$ there exists a constant c such that for all $r \geq 1, \lambda > 0$*

$$\mathbf{P}\left(Z_r^* \geq \lambda r^{\frac{1}{\beta-1}}\right) \leq c\lambda^{\beta-1-\varepsilon}.$$

Proof. This statement has been proved in [CK08, Proposition 2.2]. □

Let $m > 0$ and let us define for $r > 0$ the set

$$C_r := \left\{ v \in T_\infty : |v| = \left(1 - \frac{1}{(\log r)^m}\right) r, \exists u \in T_\infty : |u| = r, v \prec u \right\}, \tag{1.10}$$

that is, the set of vertices in generation $\left(1 - \frac{1}{(\log r)^m}\right) r$ that have a descendant in generation r. We derive the following volume estimate for this set, which we will need later on.

Lemma 1.25. *Let $k > 0$ and $K := k + \frac{m}{\beta-1}$. For $\varepsilon > 0$ there exists a constant C such that for r sufficiently large*

$$\mathbf{P}\left(V(C_r) \geq (\log r)^K\right) \leq C(\log r)^{-(k-\varepsilon)(\beta-1-\varepsilon)}.$$

Proof. We set $r^* := \left(1 - \frac{1}{(\log r)^m}\right) r$. Let $Z^*_{r^*}$ denote the size of the r^*th generation of T_∞ and let v_i, $i = 1, ..., Z^*_{r^*}$ denote the vertices of this generation. Furthermore, let T_{v_i} denote the subtrees starting at v_i. Then

$$\begin{aligned} V(C_r) &= \sum_{w \in \{v_1, ..., v_{Z^*_{r^*}}\}} \mathbb{1}\left\{ \mathsf{Height}(T_w) \geq \frac{r}{(\log r)^m} \right\} \\ &= 1 + \sum_{w \in \{v_1, ..., v_{Z^*_{r^*}}\} \setminus \{s_{r^*}\}} \mathbb{1}\left\{ \mathsf{Height}(T_w) \geq \frac{r}{(\log r)^m} \right\}, \end{aligned}$$

where we have used that the subtree $T_{s_{r^*}}$ is infinite. Note that the subtrees T_w, $w \in \{v_1, ..., v_{Z^*_{r^*}}\} \setminus \{s_{r^*}\}$, are unconditioned Galton-Watson trees. Recall from Lemma 1.15 that for $v_i \neq s_{r^*}$, $\mathbf{P}(\mathsf{Height}(T_{v_i}) \geq x) \sim c x^{\frac{-1}{\beta-1}}$, i.e. $V(C_r)$ is stochastically dominated by a random variable with distribution $1 + \mathsf{Binom}(Z^*_{r^*} - 1, p_r)$ where $p_r = c' r^{-\frac{1}{\beta-1}} (\log r)^{\frac{m}{\beta-1}}$ for some constant c'. Thus

$$\begin{aligned} &\mathbf{P}(V(C_r) \geq (\log r)^K) \\ &\leq \mathbf{P}(Z^*_{r^*} \geq r^{\frac{1}{\beta-1}} (\log r)^{k-\varepsilon}) + \mathbf{P}\left(X \geq (\log r)^K - 1\right), \end{aligned} \tag{1.11}$$

where $X \sim \mathsf{Binom}(r^{\frac{1}{\beta-1}} (\log r)^{k-\varepsilon} - 1, p_r)$. Applying Proposition 1.24 we obtain for the first term in (1.11) that for $\varepsilon > 0$ there exists a constant c such that

$$\begin{aligned} \mathbf{P}(Z^*_{r^*} \geq r^{\frac{1}{\beta-1}} (\log r)^{k-\varepsilon}) &\leq c \left[(\log r)^{k-\varepsilon} \left(1 - \frac{1}{(\log r)^m}\right)^{-\frac{1}{\beta-1}} \right]^{-(\beta-1-\varepsilon)} \\ &\leq c(\log r)^{-(k-\varepsilon)(\beta-1-\varepsilon)}. \end{aligned}$$

The second term in (1.11) is estimated by a Chernoff bound (see Lemma A.3). Let $\theta > 0$, then

$$\begin{aligned}
&\mathbf{P}(X \geq (\log r)^K - 1) \\
&\leq \frac{\mathbf{E}[e^{\theta X}]}{e^{\theta((\log r)^K - 1)}} \\
&\leq \frac{\exp\left((r^{\frac{1}{\beta-1}}(\log r)^{k-\varepsilon} - 1)c' r^{-\frac{1}{\beta-1}}(\log r)^{\frac{m}{\beta-1}}(e^\theta - 1)\right)}{\exp\left(\theta((\log r)^K - 1)\right)} \\
&\leq \frac{\exp\left(c'(\log r)^{k-\varepsilon+\frac{m}{\beta-1}}(e^\theta - 1)\right)}{\exp(\theta(\log r)^K - \theta)} \\
&\leq \exp(c'(\log r)^{k+\frac{m}{\beta-1}-\varepsilon}e^\theta - \theta(\log r)^K + \theta),
\end{aligned}$$

and since $K = k + \frac{m}{\beta-1}$ the second term in (1.11) can be upper bounded by the first one for r large enough. That is, combining both terms, we deduce that for any $\varepsilon > 0$, there exists $C < \infty$ such that for r large enough

$$\mathbf{P}\left(V(C_r) \geq (\log r)^K\right) \leq C(\log r)^{-(k-\varepsilon)(\beta-1-\varepsilon)}.$$

□

We will also need the following result concerning a uniform vertex in B_r.

Lemma 1.26. *Let $\beta \in (1,2)$. Let v_r be a uniform vertex in B_r.*

(i) Fix some constant $\kappa < \infty$ such that $B(v_r, \kappa) \subset B_r$. For any $\varepsilon > 0$ there exists $c > 0$ (depending on κ) such that

$$\mathbf{P}(V(B(v_r, \kappa)) \geq \lambda) \leq c\lambda^{-(\beta-1-\varepsilon)}.$$

In particular for every function f with $f(r) \to \infty$ as $r \to \infty$ it follows

$$\mathbf{P}(V(B(v_r, \kappa)) \geq f(r)) \to 0, \quad \text{as } r \to \infty.$$

(ii) Let $m \in \mathbb{N}$. For any $\varepsilon > 0$ we have for r sufficiently large

$$\mathbf{P}(\deg(v_r) \geq m) \leq c'' m^{-(\beta-1)} r^{\frac{-(1-\varepsilon)}{\beta-1}} + Ce^{-cr^{\varepsilon/\beta}} + c' m^{-\beta},$$

for some constants c, c', c'', C.

Proof. (i) Let $P^\kappa(v_r)$ denote the ancestor of v_r that is exactly $\kappa \wedge |v_r|$ generations before it. Let $T_{P^\kappa(v_r)}$ be the subtree of T_∞ rooted at $P^\kappa(v_r)$.

- If $P^\kappa(v_r) \neq s_i$ for all $i \in \{0,1,2,...,r\}$. Then $T_{P^\kappa(v_r)}$ is almost surely a finite subtree conditioned on having height at least κ. The offspring distribution along the leftmost path that contains v_r and has maximal height (of all the paths that visit v_r) can be stochastically dominated by the size-biased distribution (e.g. see [GK98, Lemma 2.1]).
- If there exists $i \in \{0,1,...,r\}$ such that $P^\kappa(v_r) = s_i$, then $T_{P^\kappa(v_r)}$ is almost surely infinite with backbone $s_i, s_{i+1}, ...$ and it is conditioned to reach v_r in one of the unconditioned Galton-Watson trees grafted to the T_∞-backbone (if v_r is not on the backbone itself). Again the offspring distribution along the leftmost path that contains v_r and has maximal height (of all the paths that visit v_r) can also be stochastically dominated by the size-biased distribution.

In any case, the volume of $B(v_r,\kappa)$ can be upper bounded by the volume of two balls $B^1_{2\kappa}, B^2_{2\kappa}$ with radius 2κ around the origin of T^1_∞, T^2_∞, where T^1_∞ and T^2_∞ are distributed like T_∞. That is

$$\begin{aligned}\mathbf{P}(V(B(v_r,\kappa)) \geq \lambda) &\leq \mathbf{P}\big(V(B^1_{2\kappa}) + V(B^2_{2\kappa}) \geq \lambda\big) \\ &\leq \mathbf{P}\left(2 \sup_{i=1,2} V(B^i_{2\kappa}) \geq \lambda\right) \\ &\leq 2\mathbf{P}(2V(B_{2\kappa}) \geq \lambda) \leq c\lambda^{-(\beta-1-\varepsilon)},\end{aligned}$$

by Proposition 1.22 (i), where c is some constant that depends on κ.

(ii) We calculate for r sufficiently large

$$\begin{aligned}&\mathbf{P}(\deg(v_r) \geq m) \\ &= \mathbf{P}\left(\deg(v_r) \geq m | \exists i \in \{0,1...,r\} : v_r = s_i\right) \mathbf{P}\left(\exists i \in \{0,1...,r\} : v_r = s_i\right) \\ &\quad + \mathbf{P}\left(\deg(v_r) \geq m | v_r \neq s_i \forall i \in \{0,1,...,r\}\right) \mathbf{P}\left(v_r \neq s_i \forall i \in \{0,1,...,r\}\right) \\ &\leq c' m^{-(\beta-1)} \mathbf{P}\left(\exists i \in \{0,1...,r\} : v_r = s_i | V(B_r) \geq r^{\frac{\beta-\varepsilon}{\beta-1}}\right) \\ &\quad + \mathbf{P}\left(V(B_r) \leq r^{\frac{\beta-\varepsilon}{\beta-1}}\right) + c' m^{-\beta} \\ &\leq c' m^{-(\beta-1)} \frac{r+1}{r^{\frac{(\beta-\varepsilon)}{\beta-1}}} + Ce^{-cr^{\varepsilon/\beta}} + c' m^{-\beta} \\ &= c'' m^{-(\beta-1)} r^{\frac{-(1-\varepsilon)}{\beta-1}} + Ce^{-cr^{\varepsilon/\beta}} + c' m^{-\beta},\end{aligned}$$

where we have used Proposition 1.22 (iv).

□

We also give a bound on the diameter of A_r, which is defined as

$$\mathsf{Diam}(A_r) := \sup\{|u-v| : u,v \in A_r\}.$$

Proposition 1.27. **P**-*almost surely, for any* $\varepsilon > 0$

$$\limsup_{r\to\infty} \frac{\mathsf{Diam}(A_r)}{r(\log r)^{\beta+\varepsilon}} = 0.$$

Proof. By conditioning on the heights of the subtrees grafted to the infinite backbone of T_∞, it follows from the definition of A_r that

$$\begin{aligned}
&\mathbf{P}(\mathsf{Diam}(A_r) \geq 2r\lambda) \\
&\leq \mathbf{P}\left(\sup_{v\in\mathrm{Sp}^*(r)} \mathsf{Height}(T_v) \geq r\lambda\right) \\
&\leq \mathbf{P}\left(\sup_{v\in\mathrm{Sp}^*(r)} \mathsf{Height}(T_v) \geq r\lambda \middle| V(\mathrm{Sp}^*(r)) \leq r^{\frac{1}{\beta-1}}\lambda^p\right) + \mathbf{P}\left(V(\mathrm{Sp}^*(r)) \geq r^{\frac{1}{\beta-1}}\lambda^p\right) \\
&\leq r^{\frac{1}{\beta-1}}\lambda^p \mathbf{P}\left(\mathsf{Height}(T) \geq r\lambda\right) + \mathbf{P}\left(V(\mathrm{Sp}^*(r)) \geq r^{\frac{1}{\beta-1}}\lambda^p\right),
\end{aligned}$$

where T denotes an unconditioned Galton-Watson tree. We invoke Lemma 1.15 and Lemma 1.17 (i) (recall the definition of $\mathrm{Sp}^*(r)$ and (1.9)) to compute further

$$\begin{aligned}
\mathbf{P}(\mathsf{Diam}(A_r) \geq 2r\lambda) &\leq c' r^{\frac{1}{\beta-1}}\lambda^p (r\lambda)^{\frac{-1}{\beta-1}} + c\lambda^{-p(\beta-1)} \\
&\leq c'\lambda^{p-\frac{1}{\beta-1}} + c\lambda^{-p(\beta-1)} \\
&\leq C\lambda^{\frac{-1}{\beta}},
\end{aligned}$$

where we chose $p = \frac{1}{\beta(\beta-1)}$ in the last line. Now fix $\varepsilon > 0$ and set $\hat{\varepsilon} = \frac{\varepsilon}{2}$. Choose $\lambda_r = (\log r)^{\beta+\hat{\varepsilon}}$ and consider the sequence $r_n := 2^n$. Since

$$\sum_{n=0}^{\infty} \mathbf{P}(\mathsf{Diam}(A_{r_n}) \geq 2r_n\lambda_{r_n}) \leq C\sum_{n=0}^{\infty} (\log r_n)^{-(1+\hat{\varepsilon}/\beta)} < \infty,$$

we can infer from Borel-Cantelli that along the sequence $(r_n)_n$ we have eventually almost surely

$$\frac{\mathsf{Diam}(A_{r_n})}{r_n(\log r_n)^{\beta+\hat{\varepsilon}}} \leq 2.$$

For $r \in [r_n, r_{n+1}]$ we can invoke a monotonicity argument, that is there is a $c > 0$ such that eventually almost surely

$$\frac{\mathsf{Diam}(A_r)}{r(\log r)^{\beta+\hat{\varepsilon}}} \leq \frac{\mathsf{Diam}(A_{r_{n+1}})}{r_n(\log r_n)^{\beta+\hat{\varepsilon}}} \leq \frac{\mathsf{Diam}(A_{r_{n+1}})}{\frac{r_{n+1}}{2}(\log \frac{r_{n+1}}{2})^{\beta+\hat{\varepsilon}}} \leq \frac{c}{2}\frac{\mathsf{Diam}(A_{r_{n+1}})}{r_{n+1}(\log r_{n+1})^{\beta+\hat{\varepsilon}}} \leq c.$$

Thus there exists r_0 such that for all $r \geq r_0$ we have

$$\frac{\mathsf{Diam}(A_r)}{r(\log r)^{\beta+\varepsilon}} = \frac{1}{(\log r)^{\hat{\varepsilon}}}\frac{\mathsf{Diam}(A_r)}{r(\log r)^{\beta+\hat{\varepsilon}}} \leq \frac{c}{(\log r)^{\hat{\varepsilon}}};$$

as the right hand side converges to 0 for $r \to \infty$, the statement follows. □

We will also need the following estimate on the sum of the logarithm of the degrees. Let us define Γ_r as the set of all direct paths between two vertices in B_r. Furthermore, for $\gamma \in \Gamma_r$, let $|\gamma|$ denote the length of the path.

Lemma 1.28. *There exist deterministic constants $B, \tilde{B} < \infty$ such that $\mathbf{P}$-almost surely,*

$$\sup_{\gamma \in \Gamma_r} \left\{ \sum_{v \in \gamma} \log(\deg v) - \left(\tilde{B}|\gamma| + B \log r \right) \right\} \leq 0$$

for all sufficiently large r. In particular, since for $\gamma \in \Gamma_r$, $|\gamma| \leq r$, $\mathbf{P}$-almost surely

$$\sup_{\gamma \in \Gamma_r} \sum_{v \in \gamma} \log(\deg v) \leq \tilde{B} r + B \log r,$$

for sufficiently large r.

Proof. Let $\varepsilon > 0$ and let us define $c := \beta - 1 - \varepsilon > 0$ and $B > \frac{1}{c}\left(2\frac{\beta+\varepsilon}{\beta-1} + \varepsilon\right)$. Note that $\mathbf{E}[(\deg v)^c] < \infty$ for all $v \in B_r$ (recall Proposition 1.14). We set $A := \log \mathbf{E}[(\deg v)^c] < \infty$ and $\tilde{B} := \frac{A}{c}$. Also note that the degrees of distinct vertices are independent of each other due to the Galton-Watson structure. We define for $\gamma \in \Gamma_r$, $\lambda_\gamma := \frac{B \log r}{|\gamma|} + \tilde{B}$ and $v_r := r^{\frac{\beta+\varepsilon}{\beta-1}}$. With this, we calculate

$$\begin{aligned}
\mathbf{P}\left(\sum_{v \in \gamma} \log(\deg v) \geq \lambda_\gamma |\gamma| \right) &\leq \mathbf{E}\left[\exp\left(c \sum_{v \in \gamma} \log(\deg v) \right) \right] \exp(-c\lambda_\gamma|\gamma|) \\
&\leq \exp\left(A|\gamma| - c\lambda_\gamma|\gamma|\right) \\
&= r^{-cB}.
\end{aligned}$$

Note that $|\Gamma_r| \leq V(B_r)^2$. Applying a union bound yields

$$\begin{aligned}
&\mathbf{P}\left(\sup_{\gamma \in \Gamma_r} \left\{ \sum_{v \in \gamma} \log(\deg v) - \lambda_\gamma|\gamma| \right\} \geq 0 \right) \\
&= \mathbf{P}\left(\sup_{\gamma \in \Gamma_r} \left\{ \sum_{v \in \gamma} \log(\deg v) - \lambda_\gamma|\gamma| \right\} \geq 0 \,\middle|\, V(B_r) \leq v_r \right) \mathbf{P}(V(B_r) \leq v_r) \\
&\qquad + \mathbf{P}\left(\sup_{\gamma \in \Gamma_r} \left\{ \sum_{v \in \gamma} \log(\deg v) - \lambda_\gamma|\gamma| \right\} \geq 0 \,\middle|\, V(B_r) > v_r \right) \mathbf{P}(V(B_r) > v_r) \\
&\leq v_r^2 r^{-cB} + \mathbf{P}(V(B_r) > v_r) \\
&\leq r^{-\varepsilon} + C r^{-\varepsilon/2} \\
&\leq C' r^{-\varepsilon/2},
\end{aligned}$$

where we have used Proposition 1.22 (i). The statement now follows by applying Borel-Cantelli along the subsequence $r_n = 2^n$, and then applying monotonicity if $r \in [2^n, 2^{n+1}]$. $\square$

1.5.2 Random walk on T_∞

Given a particular realisation of T_∞, we let $X = ((X_m)_{m\geq 0}, \mathbb{P}_v, v \in T_\infty)$ be the continuous time, variable speed random walk on T_∞ started at $v \in T_\infty$. This process, generated by the discrete Laplacian, is strong Markov. In particular, the jump times, started at a vertex $w \in T_\infty$ are exponentially distributed random variables with parameter $\deg(w)$. That is, the mean jump time at a vertex w is $\frac{1}{\deg(w)}$. In the case where X is started from the root O of T_∞, we just write $\mathbb{P}$ for the law of X.

Let $A \subset T_\infty$. We denote the *first exit time* from A as τ_A, i.e.

$$\tau_A := \inf\{t \geq 0 : X_t \notin A\}.$$

Given a vertex $v \in T_\infty$, we let

$$H_v := \inf\{s \geq 0 : X_s = v\},$$

denote the *first hitting time* of v.

Proposition 1.29. *Conditionally on T_∞, we have for $v \in T_\infty$ with $|v| = r$ that*

(i)
$$\mathbb{P}(H_v \leq t) \leq \exp\{-r([\log r - \log t - 1] \vee 0)\},$$

(ii) and for $r > \frac{3}{2}t$ and r sufficiently large

$$\mathbb{P}(H_v \leq t) \geq \left(\prod_{u \prec v} \frac{1}{\deg(u)}\right) \exp\{-r[\log r - \log t]\}.$$

Proof.

(i) Let $O = v_0 \prec v_1 \prec \ldots \prec v_r = v$ denote the ordered ancestors of v. To reach v from v_0, the random walk must first pass through each of the points $v_1, \ldots, v_{r-1}$. When it reaches the point v_i, the time to jump to v_{i+1} will stochastically dominate an $\mathsf{Exp}(1)$ random variable (since it can only reach v_{i+1} through one edge which rings at rate 1). The time delay for the random walk to reach v will thus be greater than $\sum_{u \prec v} E_u$, where $(E_u)_{u \prec v}$ is a sequence of iid $\mathsf{Exp}(1)$ random variables. Hence, using Stirling's formula $n! \geq \sqrt{2\pi} n^{n+1/2} e^{-n}$, for all $n = 1, 2, 3, ...$, and the fact that for $P \sim \mathsf{Poi}(t)$ it holds

$$\mathbb{P}\left(\sum_{u \prec v} E_u \leq t\right) = \mathbb{P}(P \geq r),$$

we obtain

$$\begin{aligned}\mathbb{P}(H_v \le t) \le \mathbb{P}\left(\sum_{u \prec v} E_u \le t\right) &= \mathbb{P}(P \ge r) \\ &= e^{-t}\left(\frac{t^r}{r!} + \sum_{i=r+1}^{\infty} \frac{t^i}{i!}\right) \\ &= e^{-t}\frac{t^r}{r!}\left(1 + \sum_{i=r+1}^{\infty} \frac{t^{i-r} r!}{i!}\right) \\ &\le e^{-t}\frac{t^r}{r!}\left(1 + \sum_{i=1}^{\infty} \frac{t^i}{i!}\right) \\ &= \frac{t^r}{\sqrt{2\pi} r^{r+1/2} e^{-r}} \\ &\le \exp\{-r[\log r - \log t - 1]\}.\end{aligned}$$

(ii) We lower bound the probability of hitting v in time t by the probability of going directly to v in time t, i.e. taking the direct path. Accordingly, let again $O = v_0 \prec v_1 \prec \ldots \prec v_r = v$ denote the ordered ancestors of v, and let $(E_i)_{i=0}^{r-1}$ be independent random variables, where $E_i \sim \mathsf{Exp}(\deg(v_i))$ for each $i = 0, ..., r-1$.

Then, if $3t < 2r$, since all vertices except perhaps v_0 have degree at least 2, we have for all sufficiently large r that

$$\begin{aligned}\mathbb{P}\left(\sum_{i=0}^{r-1} E_i \le t\right) \ge \prod_{i=0}^{r-1} \mathbb{P}\left(E_i \le \frac{t}{r}\right) &= \prod_{i=0}^{r-1}\left(1 - e^{-\frac{t}{r}\deg v_i}\right) \\ &\ge \left(1 - e^{-\frac{t}{r}}\right)\left(1 - e^{-\frac{2t}{r}}\right)^{r-1} \\ &\ge \left(\frac{t}{r}\right)^r \\ &= e^{-r(\log r - \log t)}.\end{aligned}$$

Since $1 - e^{-2x} \ge \frac{73}{67}x$ and $1 - e^{-x} \ge \frac{1}{2}x$ if $x \le \frac{2}{3}$ and for large enough r, $(\frac{73}{67})^{r-1}\frac{1}{2} \ge 1$. Therefore, lower bounding the desired probability by the probability of taking a direct path to v in time less than t, we deduce that

$$\mathbb{P}(H_v \le t) \ge \prod_{i=0}^{r-1} \frac{1}{\deg(v_i)} \exp\{-r[\log r - \log t]\}.$$

□

We can now also derive the following results for the exit times of A_r and B_r.

Corollary 1.30. *Conditional on T_∞, for all $t > 0$*

$$\mathbb{P}(\tau_{A_r} \le t) \le \exp\{-r([\log r - \log t - 1] \vee 0)\}.$$

Proof. Since the random walk can exit A_r only through s_r we have by using Proposition 1.29 that

$$\mathbb{P}(\tau_{A_r} \le t) = \mathbb{P}(H_{s_r} \le t) \le \exp\{-r([\log r - \log t - 1] \vee 0)\}.$$

□

For the exit time of B_r let us define $d := \frac{\beta}{\beta-1}$, $q := \frac{d}{\alpha-d}$, $\alpha > d$ and

$$r(t) := \left(\frac{t}{\log t}\right)^{q+1}, \quad a(t) := \left(\frac{t}{\log t}\right)^{q}.$$

These constants and scaling functions will appear in Chapter 2, where α will be coming from the random potential of the parabolic Anderson model. For now, let us just note these definitions and we continue to prove the following bound for the exit time probability for a ball of radius $r-1$.

Lemma 1.31. *Take any $f, p > 0$. With high $\mathbf{P}$-probability for $t \to \infty$, we have for all $r \in [r(t)(\log\log t)^{-f}, r(t)(\log t)^p]$ that*

$$\mathbb{P}\big(\tau_{B_{r-1}} \le t\big) \le \exp\left\{-r\log\left(\frac{r}{et}\right) + o\left(r(t)\right)\right\}.$$

Proof. Let $\varepsilon > 0$. Let $m > p+1$, $k > \frac{m+\delta}{\beta-1-\varepsilon} + 2\varepsilon$ for $0 < \delta < 1$ and set $K := k + \frac{m}{\beta-1}$. Recall the definition of C_r from (1.10) and by Lemma 1.25 we have for r large enough

$$\mathbf{P}\big(V(C_r) \ge (\log r)^K\big) \le C(\log r)^{-(k-\varepsilon)(\beta-1-\varepsilon)}.$$

Now, since the random walk needs to visit a vertex in C_r before it can exit the ball B_{r-1}, we can lower bound the exit time of the set B_{r-1} by the hitting time of the set C_r. On the high probability event $\{V(C_r) \le (\log r)^K\}$, we can thus compute using Proposition 1.29, for r sufficiently large

$$\begin{aligned}
&\mathbb{P}\big(\tau_{B_{r-1}} \le t\big) \\
&\le \mathbb{P}(\exists v \in C_r : H_v \le t) \\
&\le \sum_{v \in C_r} \mathbb{P}(H_v \le t) \\
&\le V(C_r)\exp\left\{-\left(1 - \frac{1}{(\log r)^m}\right) r \left[\log\left(\frac{r}{et}\right) + \log\left(1 - \frac{1}{(\log r)^m}\right)\right]\right\} \\
&\le (\log r)^K \exp\left\{-\left(1 - \frac{1}{(\log r)^m}\right) r \left[\log\left(\frac{r}{et}\right) - \frac{2}{(\log r)^m}\right]\right\} \\
&\le \exp\left\{K\log\log r - \left(1 - \frac{3}{(\log r)^m}\right) r\log\left(\frac{r}{et}\right)\right\},
\end{aligned} \tag{1.12}$$

where we have used that $\log(1-x) \geq -2x$ for $x < 1/2$. We define $\tilde{r}(t) := r(t)(\log\log t)^{-f}$ and $\tilde{R}(t) := r(t)(\log t)^p$. Now, in order to show that (1.12) holds with high probability simultaneous for all $r \in [\tilde{r}(t), \tilde{R}(t)]$, we define the sequence

$$r_n := \left(1 - \frac{1}{(\log \tilde{R}(t))^m}\right)^n \tilde{R}(t),$$

for $0 \leq n \leq N_t := \lceil \log(\tilde{R}(t))^{m+\delta} \rceil$. Then $r_0 = \tilde{R}(t)$ and for t large enough, using that for any constant c' it holds $\exp(-x^\delta) \leq x^{-c'}$ for x sufficiently large, we have

$$\begin{aligned} r_{N_t} \leq \exp\left(-\frac{\log(\tilde{R}(t))^{m+\delta}}{\log(\tilde{R}(t))^m}\right) \tilde{R}(t) &= \exp(-(\log(\tilde{R}(t)))^\delta)\tilde{R}(t) \\ &\leq \frac{1}{(\log \tilde{R}(t))^{2p}} \tilde{R}(t) \leq C r(t) \log(t)^{-p} \leq \tilde{r}(t). \end{aligned}$$

Then, taking a union bound, we have that

$$\begin{aligned} &\mathbf{P}\Big(\exists n \leq N_t : \mathbb{P}\big(\tau_{B_{r_n - 1}} \leq t\big) \geq \exp\Big\{K(\log\log r_n) - \Big(1 - \frac{3}{(\log r_n)^m}\Big) r_n \log\Big(\frac{r_n}{et}\Big)\Big\}\Big) \\ &\leq \mathbf{P}\big(\exists n \leq N_t : V(C_{r_n}) \geq (\log r_n)^K\big) \\ &\leq \sum_{n \leq N_t} c(\log r_n)^{-(k-\varepsilon)(\beta-1-\varepsilon)} \\ &\leq \sum_{n \leq N_t} c\left[\log(\tilde{R}(t)) + n \log\left(1 - \frac{1}{(\log \tilde{R}(t))^m}\right)\right]^{-(k-\varepsilon)(\beta-1-\varepsilon)} \\ &\leq \sum_{n \leq N_t} c\left[\log(\tilde{R}(t)) - n\frac{2}{(\log \tilde{R}(t))^m}\right]^{-(k-\varepsilon)(\beta-1-\varepsilon)} \\ &\leq \sum_{n \leq 2\log(\tilde{R}(t))^{m+\delta}} c\left[\log \tilde{R}(t)\left(1 - \frac{4}{(\log \tilde{R}(t))^{1-\delta}}\right)\right]^{-(k-\varepsilon)(\beta-1-\varepsilon)} \\ &\leq C\left[\log(\tilde{R}(t))\right]^{m+\delta-(k-\varepsilon)(\beta-1-\varepsilon)}. \end{aligned}$$

Since $k > \frac{m+\delta}{\beta-1-\varepsilon} + 2\varepsilon$ this probability converges to zero as $t \to \infty$, i.e. with high $\mathbf{P}$-probability

$$\begin{aligned} &\mathbb{P}\big(\tau_{B_{r_n - 1}} \leq t\big) \\ &\leq \exp\left\{K(\log\log r_n) - \left(1 - \frac{3}{(\log r_n)^m}\right) r_n \log\left(\frac{r_n}{et}\right)\right\} \quad \text{for all } n \leq N_t. \quad (1.13) \end{aligned}$$

Now, on the event in (1.13), we have for all $r \in [\tilde{r}(t), \tilde{R}(t)]$ with $r \in [r_{n+1}, r_n]$ that

$$\begin{aligned}
&\mathbb{P}\big(\tau_{B_{r-1}} < t\big)\\
&\leq \mathbb{P}\Big(\tau_{B_{r_{n+1}-1}} < t\Big)\\
&\leq \exp\left\{K \log\log r_{n+1} - \left(1 - \frac{3}{(\log r_{n+1})^m}\right) r_{n+1} \log\left(\frac{r_{n+1}}{et}\right)\right\}\\
&\leq \exp\left\{K \log\log r - r_n \left(1 - \frac{1}{(\log \tilde{R}(t))^m}\right) \log\left(\frac{r_{n+1}}{et}\right) + \frac{3r \log\left(\frac{r}{et}\right)}{(\log r_{n+1})^m}\right\}\\
&\leq \exp\left\{K \log\log r - r \log\left(\frac{r}{et}\left(1 - \frac{1}{(\log \tilde{R}(t))^m}\right)\right) + \frac{r \log\left(\frac{r}{et}\right)}{(\log \tilde{R}(t))^m} + \frac{3r \log\left(\frac{r}{et}\right)}{(\log r_{n+1})^m}\right\}\\
&\leq \exp\left\{-r \log\left(\frac{r}{et}\right) + K \log\log r + \frac{2r}{(\log \tilde{R}(t))^m} + \frac{r \log\left(\frac{r}{et}\right)}{(\log \tilde{R}(t))^m} + \frac{3r \log\left(\frac{r}{et}\right)}{(\log r_{n+1})^m}\right\}.
\end{aligned} \tag{1.14}$$

We compute further

$$K \log\log r \leq K \log\log(\tilde{R}(t)) = K \log\log\left(r(t)(\log t)^p\right) = o(r(t)),$$

and

$$\begin{aligned}
\frac{3r \log\left(\frac{r}{et}\right)}{(\log r_{n+1})^m} \leq \frac{3r \log\left(\frac{r}{et}\right)}{(\log r_{N_t})^m} &= \frac{3r \log\left(\frac{r}{et}\right)}{\left[\log\left(\left(1 - \frac{1}{(\log(\tilde{R}(t)))^m}\right)^{N_t} \tilde{R}(t)\right)\right]^m}\\
&\leq \frac{3r \log\left(\frac{r}{et}\right)}{\left[\log\left(\tilde{R}(t)\right) - \frac{2N_t}{(\log(\tilde{R}(t)))^m}\right]^m}\\
&\leq \frac{3r \log\left(\frac{r}{et}\right)}{\left[\log\left(\tilde{R}(t)\right)\left(1 - \frac{4}{\log(\tilde{R}(t))^{1-\delta}}\right)\right]^m}\\
&\leq \frac{Cr \log\left(\frac{r}{et}\right)}{(\log \tilde{R}(t))^m}.
\end{aligned} \tag{1.15}$$

Finally, we compute

$$\begin{aligned}
\frac{r \log\left(\frac{r}{et}\right)}{(\log \tilde{R}(t))^m} \leq \frac{\tilde{R}(t) \log\left(\frac{\tilde{R}(t)}{et}\right)}{(\log \tilde{R}(t))^m} &= \frac{r(t)(\log t)^p \log\left(\frac{r(t)(\log t)^p}{et}\right)}{(\log(r(t)(\log t)^p))^m}\\
&\leq Cr(t) \frac{(\log t)^{p+1}}{\left(\log\left(\left(\frac{t}{\log t}\right)^{q+1}\right)\right)^m}\\
&\leq Cr(t)(\log t)^{p+1-m}\\
&= o(r(t)),
\end{aligned} \tag{1.16}$$

where we have used that $m > p+1$ by assumption. In particular, we can conclude from (1.15) and (1.16) that the last three terms in (1.14) are also of order $o(r(t))$, i.e. with high **P**-probability as $t \to \infty$ we have *for all* $r \in [\tilde{r}(t), \tilde{R}(t)]$

$$\mathbb{P}\big(\tau_{B_{r-1}} \le t\big) \le \exp\left\{-r\log\left(\frac{r}{et}\right) + o\left(r(t)\right)\right\}.$$

□

Chapter 2

The parabolic Anderson model with Pareto potential on critical Galton-Watson trees

We prove that the solution to the parabolic Anderson model with *Pareto potential* on a *critical Galton-Watson tree* conditioned to survive with an offspring distribution in the domain of attraction of a stable law *localises with high probability* in one single vertex for time going to infinity. This chapter is based on joint work with Eleanor Archer.

2.1 Introduction

In a very general form the *parabolic Anderson model* denotes the parabolic initial value problem

$$\begin{aligned} \partial_t u(t,v) &= \triangle u(t,v) + \xi(t,v)u(t,v), \quad && (t,v) \in (0,\infty)\times\mathcal{S}, \\ u(0,v) &= u_0(v), && v \in \mathcal{S}, \end{aligned}$$

where $\mathcal{S}$ is a space equipped with a Laplacian and $(\xi(t,v) : t > 0, v \in \mathcal{S})$ is a possibly time-dependent *random field*. That is, we are dealing with a stochastic heat equation with multiplicative noise, see the Introduction. Here, we are only interested in the *stationary case*, i.e. where the potential is time-independent and furthermore we consider a *discrete space*, say $\mathbb{Z}^d$, and localised initial data. We will thus generally refer to the following problem as the parabolic Anderson model (PAM) on $\mathbb{Z}^d$.

Definition 2.1 (Parabolic Anderson model (PAM) on $\mathbb{Z}^d$).

$$\begin{aligned} \partial_t u(t,z) &= \Delta u(t,z) + \xi(z)u(t,z), \quad && (t,z) \in (0,\infty)\times\mathbb{Z}^d, \\ u(0,z) &= \mathbb{1}\{z=0\}, && z \in \mathbb{Z}^d, \end{aligned} \tag{2.1}$$

where

$$(\Delta f)(z) = \sum_{y \sim z} [f(y) - f(z)], \quad z \in \mathbb{Z}^d, f : \mathbb{Z}^d \to \mathbb{R}, \tag{2.2}$$

denotes the *discrete Laplace operator* and $(\xi(z) : z \in \mathbb{Z}^d)$ is a collection of independent identically distributed random variables, called the *random potential*. The probability measure associated to ξ will be denoted as $\mathcal{P}$ and the corresponding expectation as $\mathcal{E}$.

The PAM is a classical model for the spread of particles in an inhomogeneous random environment, with applications including *population dynamics* and *chemical reactions*, see [CM94] for details. Even though the discrete model on $\mathbb{Z}^d$ may, in some cases, from the modelling point of view only be an approximation to a physical system in continuous space, many qualitative phenomenon of the model are the same on $\mathbb{Z}^d$ and $\mathbb{R}^d$. The PAM can be derived as the evolution equation for the expected number of particles at a given site and time of a *branching random walk in random environment*, see Subsection 2.2.2 for details. A comprehensive survey of results on the model can be found in the book [Kön16]. We also refer to the surveys [GK05, Mör09] for an overview.

One of the most interesting phenomenon that can be observed in the PAM is a localisation behaviour of the solution. That is, for t large, the solution becomes concentrated in a few disjoint islands, outside of which the contribution to the *total mass*

$$U(t) := \sum_{z \in \mathbb{Z}^d} u(t, z),$$

is negligible. This effect is called *intermittency*. The size and number of islands depend on the tail decay of the potential, which leads to several different regimes [vdHKM06]. Note that with a constant, non-random potential, a diffusion process will spread most of its mass at time t over a ball with radius of order $\sqrt{t}$, i.e. it is indeed the randomness of the potential that drives the localisation behaviour in the PAM. We will discuss this phenomenon in Subsection 2.2.1 in more detail. In this chapter, we will focus on a setting with a *Pareto-distributed potential*, that is with polynomially decaying tails, for which a strong intermittency effect can be observed. More precisely, in [KLMS09] it was shown that the solution to (2.1) with Pareto potential localises eventually almost surely at two sites. The authors also derived that with high probability the solution actually localises at one site only for time going to infinity (this result was also proven earlier in the unpublished preprint [KMS06] by different tools).

Motivated by e.g. biological models on network, it is an intriguing task to analyse the PAM and its intermittency behaviour on other graphs besides $\mathbb{Z}^d$. In this chapter we are especially interested in random graphs, or more precisely, *Galton-Watson trees* combined with a Pareto-distributed potential. In order to consider the long-time behaviour of the PAM on a Galton-Watson tree, it is natural to restrict to the case of an infinite tree. The most natural candidate for such a tree would be a super-critical Galton-Watson tree as it has a positive probability to survive (see Theorem 1.4). However, in the case of a Pareto potential the exponential volume growth in a super-critical Galton-Watson tree causes the solution of the PAM to blow up, almost surely. We will therefore consider the PAM on *critical* Galton-Watson trees that are *conditioned to survive*, see Subsection 1.5 for the precise definition of our tree model T_∞. Moreover, we choose an offspring distribution in the *domain of attraction of a β-stable law*, i.e. β is the main parameter controlling the random tree. We note that T_∞ has unbounded degree and non-uniform volume growth, which makes the analysis more delicate compared to the $\mathbb{Z}^d$ case. Furthermore, the random tree introduces another layer of randomness into the model and most results are stated in the joint probability law of the potential and the tree.

In our main Theorem 2.14, we show, in a similar manner as in [KMS06] for the $\mathbb{Z}^d$ case, that the solution of the PAM on T_∞ with Pareto potential localises with high probability in one single vertex as time goes to infinity. This is joint work with Eleanor Archer.

The only other work on the PAM on a random tree, at least that we know of, is given in [dKd20]. There, the authors consider *super-critical Galton-Watson trees* with a bounded degree assumption together with a random potential that has a double exponential tail. In particular, they analyse the *large t asymptotics* of the total mass, which gives information about the localisation and the structure of the localisation set, see Subsection 2.2.1. Apart from that, we know of two other works where the PAM is considered on deterministic graphs different from $\mathbb{Z}^d$, namely on sequences of the *complete graph* [FM90] and the *hypercube* [AGH20] (note that both have unbounded degree in the limit, though to cancel out this effect the Laplace operator was multiplied by a suitable pre-factor in both settings).

The rest of this chapter is structured as follows: In Section 2.2 we will summarize a few aspects of the PAM on $\mathbb{Z}^d$ and provide the corresponding references. We will continue with Section 2.3, where we introduce the PAM on the tree T_∞ (defined in Section 1.5). After deriving estimates on the extremal values of the Pareto potential with parameter α on T_∞, we will establish the existence/non-existence for a non-negative solution of the PAM on T_∞ for $\alpha < \frac{\beta}{\beta-1}$ / $\alpha > \frac{\beta}{\beta-1}$. In the last Subsection 2.3.4 we will state our main theorem and give details on the proof strategy. We continue with Section 2.4, where we will derive estimates for the maximizer of the potential and for the concentration site that will be essential for the main proofs. The aim of Section 2.5 is twofold, firstly to derive a concentration result for the principal eigenfunction of the Hamiltonian appearing in the PAM on a bounded set and secondly to infer a relation between the solution of the PAM and the principal eigenfunction allowing to transfer the concentration result to the solution. We combine the auxiliary results from the previous sections in Section 2.6, where the main theorem about the one point localisation of our model will be proved. Finally, we conclude this chapter with an outlook on future research in Section 2.7.

Note that all results in this chapter are based on joint work with Eleanor Archer.

2.2 The parabolic Anderson model on $\mathbb{Z}^d$

The *parabolic Anderson model* is named after the physicist P. W. Anderson (1923-2020). In particular, it is the parabolic analogue to the Schrödinger equation (on $\mathbb{Z}^d$) in quantum mechanics with the so-called *Anderson Hamiltonian* $H := \Delta + \xi(\cdot)$,

$$i\frac{\partial \psi}{\partial t} = H\psi, \quad \psi(0,z) = \psi_0(z) \in \ell^2(\mathbb{Z}^d);$$

describing the dynamics of the wave function of an electron inside a semiconductor, which is assumed to have random impurities described by the random potential ξ. Anderson's research [And58] revealed that the wave function may concentrate on small domains, which is today known as *Anderson localisation*. The spectral properties of the Anderson Hamiltonian play a central role in the understanding of this phenomenon and a related spectral analysis will also be relevant for proving intermittency effects in the PAM, see Subsection 2.2.1. For further physical background we also refer to [Mol91] and for theoretical background on random operators to [AW15].

The existence of a non-negative solution to (2.1) is established in the following theorem.

Theorem 2.2 (see [CM94, GM90])**.** *Assume that the random potential fulfils*

$$\mathcal{E}\left[\left(\frac{\xi(0) \vee 2}{\log(\xi(0) \vee 2)}\right)^d\right] < \infty. \tag{2.3}$$

Then (2.1) *has a unique non-negative solution.*

Remark 2.3. Since we will be working with a Pareto-distributed potential later on, let us note that if $(\xi(z) : z \in \mathbb{Z}^d)$ is a family of iid random variables with Pareto distribution with parameter $\alpha > 0$, that is

$$\mathcal{P}(\xi(0) < x) = 1 - x^{-\alpha}, \quad x \geq 1,$$

then condition (2.3) translates to $\alpha > d$.

One of the main probabilistic tools for analysing the PAM is the representation of the solution via a *Feynman-Kac formula.*

Definition 2.4 (Feynman-Kac formula)**.** Under condition (2.3) the solution to (2.1) is given by

$$u(t,z) = \mathbb{E}_0\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\} \mathbb{1}\{X_t = z\}\right], \quad (t,z) \in \mathbb{R}^+ \times \mathbb{Z}^d, \tag{2.4}$$

where $(X_s)_{s\in[0,\infty)}$ is a time-continuous random walk on $\mathbb{Z}^d$ with generator Δ starting at $0 \in \mathbb{Z}^d$ under $\mathbb{E}_0$.

The expectation in the Feynman-Kac formula is taken with respect to the random walk and the solution is still random with respect to the potential. This formula already suggests that sites that can be reached quickly by the random walk and that have a large potential value contribute heavily to the mass of the solution.

2.2.1 Intermittency

The parabolic Anderson model (2.1) with iid random potential exhibits a concentration property, called (geometric) *intermittency* [GKM07]. That is, the solution is concentrated for t large, on a few, small, remote islands, which carry most of the total mass $U(t)$. These islands are called *relevant islands* and identifying their number, size and location constitutes the main task in characterising the intermittency phenomenon. It is assumed that the heavier the tails of the distribution of the potential, the smaller the number of relevant islands and the smaller their size. That is, the intermittency effect becomes more pronounced.

The analysis of this effect often relies on a *spectral decomposition* of the solution in a large box, i.e. an expansion with respect to the eigenfunctions of the Anderson Hamiltonian. Let H_B denote the Anderson Hamiltonian on a finite set $B \subset \mathbb{Z}^d$ with Dirichlet boundary conditions. Let $\phi_1, \phi_2, ..., \phi_{|B|}$ be the eigenfunctions for the decreasingly ordered eigenvalues $\lambda_1, ..., \lambda_{|B|}$ of H_B. Note that the eigenfunctions form an orthonormal basis of $\ell^2(B)$. We set $\phi_i(z) = 0$ for $z \in B^c$ and $i = 1, ..., |B|$. Then the solution u_B of (2.1) in B admits the representation

$$u_B(t, \cdot) = \sum_{k=1}^{|B|} e^{t\lambda_k} \phi_k(0)\phi_k(\cdot), \quad t \in (0, \infty). \tag{2.5}$$

Heuristically, intermittency can now be explained as follows: As the *Anderson localisation* predicts, the eigenfunctions of H belonging to eigenvalues close to the boundary of the spectrum are exponentially concentrated in small areas that are randomly located. Assuming that this also holds on the large box B, we expect ϕ_k to be small outside a small neighbourhood around its centre x_k and furthermore, due to the exponential term $e^{t\lambda_k}$, summands in (2.5) with large k become negligible against the leading terms $e^{t\lambda_1}, e^{t\lambda_2}, ...$ Since the concentration centres x_k are also predicted to be far away from each other, the solution u_B is well approximated around x_k by the summand $e^{t\lambda_k}\phi_k(0)\phi_k(x_k + \cdot)$, i.e. u_B has high values on a few (only small values of k relevant), small islands that are well separated in space. See [Kön16] for more heuristics on the intermittency phenomenon. For a rigorous proof the solution is considered on time-dependent boxes B_t that grow for $t \to \infty$ and the spectral analysis is combined with the probabilistic expansion via the Feynman-Kac representation.

Let us review some known results on geometric intermittency in the PAM. In [GKM07] the authors showed that for potentials with *double exponential tails* and heavier tails with finite exponential moments, the contribution to the total mass coming from outside a random number of relevant islands is eventually almost surely negligible. They also proved that for potentials with tails heavier than the double exponential tail (i.e. including Pareto and Weibull distributions) the localisation islands consist of single sites. Later the number of localisation islands were specified for the Pareto and Weibull distribution. In the setting with *Pareto potential* (no finite exponential moments) with parameter $\alpha > d$ it was shown that eventually almost surely the non-negligible contribution to the total mass is coming from at most two single sites [KLMS09] and moreover that with high probability it is indeed only one site. Likewise, in [ST14] localisation in one site with high probability was shown for the setting with the *Weibull potential* (i.e. $\mathcal{P}(\xi(0) > x) \sim \exp(-x^\gamma)$) with parameter $0 < \gamma < 2$. The authors also conjecture that almost surely the solution localises in only two sites. The case with *Weibull potentials* with parameter $\gamma \geq 2$ was settled in [FM14]. Again, the solution is eventually localised with high probability in one single site.

Remark 2.5. The localisation in two points is the strongest form of localisation that can hold almost surely. This is due to the fact that the localisation sites are time-dependent, i.e. there will be a sequence of transition times when the solution 'jumps' from one site to another, meaning that both sites contain non-negligible mass at that time.

A related notion of intermittency is given in terms of large time asymptotics of the moments of the total mass; more precisely, it is expressed as a faster growth rate of higher moments. That is, if all exponential moments $\mathcal{E}[\exp(\lambda\xi(0))]$, $\lambda > 0$, exist, then all moments of the total mass $\mathcal{E}[U(t)^p]$, $t, p > 0$, exist and the PAM is called *intermittent* if

$$\limsup_{t\to\infty} \frac{\mathcal{E}[U(t)^p]^{1/p}}{\mathcal{E}[U(t)^q]^{1/q}} = 0, \qquad \text{for } 0 < p < q. \tag{2.6}$$

This notion was motivated by the statistical physics literature, see for example [ZMRS87]. Furthermore, properties of the relevant islands are on a heuristic level reflected in the asymptotic expansion of $\mathcal{E}[U(t)]$ for large t, cf. [GK05]. For many potentials the large t asymptotics of both the moments of $U(t)$ (annealed setting) and of $U(t)$ itself (almost sure setting, or quenched setting) have been analysed, see for example [BK01,GM90,GM98] and also [vdHKM06]. Note that the definition via (2.6) can not be applied to the PAM with Pareto potential since this distribution does not possess any exponential moments. We therefore apply the more explicit, geometric notion of intermittency from above. For a heuristic relation between these two notions we refer to [GM90] and [BC95, Section 2.4].

2.2.2 Branching random walk in random environment

There exist a close relationship between the PAM and a *branching process in random environment* defined on $\mathbb{Z}^d$, see for example [GM90]. Suppose that at time $t = 0$

there exist a single particle at the site z that starts to move according to the law of a nearest neighbour, time-continuous simple random walk on $\mathbb{Z}^d$. Furthermore, at each site x the particle splits into two particles at a rate $\xi(x)$, where $\xi = (\xi(x) : x \in \mathbb{Z}^d)$ is a family of iid random variables (the *random environment*). Each descendant evolves and splits according to the same laws but independently of each other. Let $N(t, x)$ denote the total number of particles that occupy the site x at time t. Furthermore, for a fixed realization of the random potential ξ we denote the probability law and the expectation over the branching and migration mechanisms as $\mathbb{P}_z$ and $\mathbb{E}_z$, respectively. Then $u(t, z) := \mathbb{E}_z[N(t, 0)]$ satisfies (2.1), i.e. it solves the parabolic Anderson model on $\mathbb{Z}^d$. That is, one may say that the PAM is the *thermodynamic limit* of this particle system. Due to the symmetry of the Anderson operator it holds $u(t, z) = \mathbb{E}_z[N(t, 0)] = \mathbb{E}_0[N(t, z)]$, that is, on a particle level the solution of the PAM can be regarded as the mean number of particles present at site z at time t given that the branching random walk started with a single particle at the origin.

The branching system itself has been the subject of intesive research. Very recently the setting with Pareto potential has been studied in great detail in [OR16, OR17, OR18]. Amongst others, the authors showed that a process that only depends on the random environment ξ, the so-called *lilypad process*, governs many aspects of the system such as the hitting times of sites, the number of particles and the support in a rescaled version. We also refer to the recent survey [Kön20] on branching random walks in random environment.

2.2.3 Continuous space, time-dependent potentials

There are also many interesting research works on the PAM defined on a continuous space and/or with a time-dependent potential. We will mention a few below, however, this is by no means a complete list.

The analysis of intermittency in the PAM on a continuous spatial domain, in particular on $\mathbb{R}^d$, was initiated in [Szn93, Szn98], where Brownian motion among (time-independent) Poisson traps has been analysed. The traps are constructed by setting the potential equal to $-\infty$ in neighbourhoods of sites of a homogeneous Poisson point process. In [CM95] the authors analyse intermittency in the sense of (2.6) for the PAM with a homogeneous ergodic random field ξ on $\mathbb{R}^d$. Gaussian potentials with certain regularity properties were studied in [GK00, GKM00]. Very recently, in [KPvZ20] the PAM with a Gaussian space white noise potential (time-independent) in $\mathbb{R}^2$ has been investigated, in particular, the almost sure large time asymptotic behaviour of the total mass was analysed.

The PAM with time-dependent random fields on $\mathbb{Z}^d$ is subject of the monograph [CM94].

Finally, settings with time-dependent potentials in continuous space have been analysed, for example, in the following works: In [BC95] the authors showed intermittency behaviour for the PAM with space-time white noise in $\mathbb{R}$. The same model is subject of the recent work [CKNP20], where the authors proved spatial ergodicity and a central limit theorem for the solution. In [HHNT15] stochastic heat equations in $\mathbb{R}^d$ with more general multiplicative centred Gaussian noise are analysed. We also refer to the book [Kho14]. Finally, in [CK19] the authors investigate intermittency for the stochastic heat equation with Lévy noise.

2.3 The parabolic Anderson model on T_∞

Let T_∞ be the critical Galton-Watson tree conditioned to survive with an offspring distribution in the domain of attraction of a β-stable law with $\beta \in (1,2)$, as characterized in Section 1.5. We now define the parabolic Anderson model with Pareto potential on this tree. That is, from now on we consider

$$\begin{aligned} \partial_t u(t,v) &= \Delta u(t,v) + \xi(v)u(t,v), \quad (t,v) \in (0,\infty) \times T_\infty, \\ u(0,v) &= \mathbb{1}\{v = O\}, \qquad\qquad\qquad\qquad\quad v \in T_\infty, \end{aligned} \tag{2.7}$$

where O denotes the root of T_∞, $\triangle$ is the discrete Laplacian, i.e.

$$(\Delta f)(v) = \sum_{y \sim v} [f(y) - f(v)], \quad v \in T_\infty, f : T_\infty \to \mathbb{R},$$

and $(\xi(v) : v \in T_\infty)$ is a collection of independent Pareto-distributed random variables with parameter α, i.e. for any $v \in T_\infty$

$$\mathcal{P}(\xi(v) > x) = x^{-\alpha}.$$

Note that we have two sources of randomness in this model, coming from the random potential and the underlying random tree. We sample the tree first, and then independently sample the potential; i.e. the law of the potential $\mathcal{P}$ is defined under the law of the tree $\mathbf{P}$ and results are given in the product law $\mathbf{P} \times \mathcal{P}$.

It turns out that the quotient $\frac{\beta}{\beta-1}$ plays a similar role as the dimension d for the parabolic Anderson model on the d-dimensional lattice. Therefore we set from now on

$$d := \frac{\beta}{\beta - 1}.$$

In the following subsections we will first derive bounds for the maximal value of the potential ξ on T_∞. Subsequently, we will verify the existence of a Feynman-Kac representation of the solution to (2.7) for certain parameter values in Subsections 2.3.2-2.3.3. Then we will state the main localisation theorem and outline the proof strategy in Subsection 2.3.4.

2.3.1 Extremal values of the potential ξ

In the following we will derive upper and lower bounds on the maximal potential ξ in certain subsets of T_∞.

Lemma 2.6. *It holds for any $r, \lambda > 1$*

$$\mathbf{P} \times \mathcal{P}\left(\sup_{v \in A_r} \xi(v) \geq r^{d/\alpha}\lambda\right) \leq C\lambda^{-\frac{\alpha}{1+\beta d}},$$
$$\mathbf{P} \times \mathcal{P}\left(\sup_{v \in B_r} \xi(v) \geq r^{d/\alpha}\lambda\right) \leq C'\lambda^{-\frac{\alpha-\varepsilon}{d}},$$

for some constants C, C'.

Proof. Using Proposition 1.22 (iii), we compute

$$\begin{aligned}
&\mathbf{P} \times \mathcal{P}\left(\exists v \in A_r : \xi(v) > r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}}\lambda\right) \\
&\leq r^{\frac{\beta}{\beta-1}}\lambda^{\frac{\alpha\beta^2}{\beta-1+\beta^2}}\mathbf{P} \times \mathcal{P}\left(\xi(O) > r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}}\lambda\right) + \mathbf{P}\left(V(A_r) \geq r^{\frac{\beta}{\beta-1}}\lambda^{\frac{\alpha\beta^2}{\beta-1+\beta^2}}\right) \\
&< r^{\frac{\beta}{\beta-1}}\lambda^{\frac{\alpha\beta^2}{\beta-1+\beta^2}}r^{-\frac{\beta}{\beta-1}}\lambda^{-\alpha} + C\lambda^{-\frac{\alpha(\beta-1)}{\beta-1+\beta^2}} \\
&= C\lambda^{-\frac{\alpha(\beta-1)}{\beta-1+\beta^2}}.
\end{aligned}$$

The same kind of proof works for B_r. □

Lemma 2.7. $\mathbf{P} \times \mathcal{P}$*-almost surely, for any $\varepsilon > 0$ we have that*

$$\lim_{r\to\infty} \frac{\sup_{v\in A_r} \xi(v)}{r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}}(\log r)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}} = 0 \tag{2.8}$$

and

$$\lim_{r\to\infty} \frac{\sup_{v\in B_r} \xi(v)}{r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}}(\log r)^{\frac{d+\varepsilon}{\alpha}}} = 0. \tag{2.9}$$

Proof. Let $\varepsilon > 0$ be arbitrary and set $c := 2^{\frac{1}{\alpha}\left(2\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}$. Using Lemma 2.6 we deduce for any $\delta > 0$

$$\begin{aligned}
&\mathbf{P} \times \mathcal{P}\left(\exists v \in A_r : \xi(v) > \frac{\delta}{c} r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}}(\log r)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}\right) \\
&\leq C\delta^{-\frac{\alpha}{1+\beta d}}(\log r)^{-\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)\frac{\alpha}{1+\beta^2/(\beta-1)}} \\
&= C\delta^{-\frac{\alpha}{1+\beta d}}(\log r)^{-1-\varepsilon'},
\end{aligned}$$

with $\varepsilon' = \frac{\varepsilon}{1+\beta^2/(\beta-1)} > 0$. This probability is summable along $r_n = 2^n$, that is

$$\begin{aligned}
&\sum_{n=1}^{\infty} \mathbf{P} \times \mathcal{P}\left(\exists v \in A_{r_n} : \xi(v) > \delta r_n^{\frac{1}{\alpha}\frac{\beta}{\beta-1}}(\log r_n)^{\frac{1}{\alpha}\left(\frac{\beta^2+\beta-1+\varepsilon'}{\beta(\beta-1)}+1+\varepsilon'\right)}\right) \\
&< \delta^{-\frac{\alpha}{1+\beta d}} C \sum_{n=1}^{\infty}(n\log 2)^{-1-\varepsilon'} < \infty,
\end{aligned}$$

and thus by Borel-Cantelli eventually $\mathbf{P} \times \mathcal{P}$-almost surely

$$\frac{\sup_{v \in A_{r_n}} \xi(v)}{r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}} (\log r_n)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}} \le \frac{\delta}{c} < \delta.$$

Now, let $r \in [r_n, r_{n+1}]$, n large enough. Since $V(A_r) < V(A_{r_{n+1}})$ we have

$$\begin{aligned}\frac{\sup_{v \in A_r} \xi(v)}{r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}} (\log r)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}} &< \frac{\sup_{v \in A_{r_{n+1}}} \xi(v)}{r_n^{\frac{1}{\alpha}\frac{\beta}{\beta-1}} (\log r_n)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}} \\ &\le 2^{\frac{1}{\alpha}\left(2\frac{\beta}{\beta-1}+1+\varepsilon\right)} \frac{\sup_{v \in A_{r_{n+1}}} \xi(v)}{r_{n+1}^{\frac{1}{\alpha}\frac{\beta}{\beta-1}} (\log r_{n+1})^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}} \\ &< \delta.\end{aligned}$$

Thus,

$$\limsup_{r \to \infty} \frac{\sup_{v \in A_r} \xi(v)}{r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}} (\log r)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}} = 0,$$

and because of non-negativity, the limit is also zero.

The same kind of proof works for B_r. □

2.3.2 Existence of solutions for $\alpha > \frac{\beta}{\beta-1}$

Given a particular realisation of T_∞, let us denote by $X = ((X_m)_{m \ge 0}, \mathbb{P}_v, v \in T_\infty)$ the continuous time, variable speed random walk on T_∞ started at $v \in T_\infty$, as characterised in Subsection 1.5.2. Note again that the law $\mathbb{P}_v$ is defined under the law $\mathbf{P}$ and we write $\mathbb{P}$ for the random walk starting in the root O.

Furthermore define

$$u(t,z) = \mathbb{E}_O\left[\exp\left\{\int_0^t \xi(X_s)\,\mathrm{d}s\right\} \mathbb{1}\{X_t = z\}\right], \quad t > 0,\ z \in T_\infty. \tag{2.10}$$

Heuristically, the expectation on the right hand side considers paths that start at the root and end in z at time t, where in the exponential the potential values that the walker encounters on its path are accumulated and weighted by the times it spends at the respective vertices. A classical result by Gärtner and Molchanov [GM90, Lemma 2.2] yields that if $u(t,z) < \infty$ for all $(t,z) \in (0,\infty) \times T_\infty$, then $u(t,z)$ is a non-negative solution to the Cauchy problem (2.7). Hence, to establish existence of a solution we need to guarantee that the *Feynman-Kac formula* (2.10) does not explode. For that we will compare the likelihood of the random walk exiting the set A_r with the maximum potential it should encounter on the set A_r.

Proposition 2.8. *Take $\alpha > \frac{\beta}{\beta-1}$. Then, $\mathbf{P}\times\mathcal{P}$-almost surely, the Cauchy problem (2.7) possesses a unique non-negative solution $u : (0,\infty)\times T_\infty \to [0,\infty)$ given by the Feynman-Kac formula*

$$u(t,z) = \mathbb{E}_O\left[\exp\left\{\int_0^t \xi(X_s)\ ds\right\}\mathbb{1}\{X_t = z\}\right], \quad t>0,\ z\in T_\infty. \tag{2.11}$$

Proof. In fact we can bound the total mass for all times. Decomposing according to how far the random walk progresses along the backbone up until time t, and then substituting the results of Lemma 2.7 and Corollary 1.30, we see that $\mathbf{P}\times\mathcal{P}$-almost surely,

$$\begin{aligned} U(t) &= \mathbb{E}_O\left[\exp\left\{\int_0^t \xi(X_s)\ \mathrm{d}s\right\}\right] \\ &\le \sum_{r=0}^\infty \mathbb{P}(\tau_{A_r}\le t)\exp\left\{t\sup_{v\in A_r}\xi(v)\right\} \\ &\le C\sum_{r=0}^\infty \exp\{-r[\log r-\log t-1]\}\exp\left\{tr^{\frac{1}{\alpha}\frac{\beta}{\beta-1}}(\log r)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}\right\}, \end{aligned}$$

and for $\alpha > \frac{\beta}{\beta-1}$ the right hand side is finite. *Existence* thus follows from Lemma 2.2 in [GM90].

A potential ξ is called *percolating from below* if for each $k\in\mathbb{R}$ the level set $\{v\in T_\infty : \xi(v)\le k\}$ contains an infinite connected component, see [GM90]. Now, $\mathbf{P}$-almost surely a random potential ξ is non-percolating from below. By Lemma 2.3 in [GM90] this implies the *uniqueness* of the solution. □

2.3.3 Non-existence of solutions for $\alpha < \frac{\beta}{\beta-1}$

For technical reasons let us now consider the sets

$$\begin{aligned} \tilde{T}_\infty &:= \{v\in T_\infty : \deg v\le 4\}, \\ \tilde{A}_r &:= \{v\in A_r : \deg v\le 4\}. \end{aligned}$$

Lemma 2.9. *For a realisation of T_∞, it holds*

$$V(\tilde{A}_{2r}\setminus\tilde{A}_r) \ge \frac{1}{2}V(A_{2r}\setminus A_r).$$

Proof. Let n denote the number of vertices in $A_{2r}\setminus A_r$. Since $A_{2r}\setminus A_r$ is a tree, it contains $n-1$ edges. It therefore follows that

$$\sum_{v\in A_{2r}\setminus A_r}\deg v - 2 = 2(n-1),$$

where we have subtracted 2 from the left-hand side since we have removed one child of s_{2r-1} and the parent of s_r to construct $A_{2r} \setminus A_r$. Then, letting N denote the number of vertices in $A_{2r} \setminus A_r$ with degree at least 4, it follows that

$$4N \leq \sum_{v \in A_{2r} \setminus A_r} \deg v = 2n,$$

so that $N \leq \frac{n}{2}$. This concludes the proof. □

Lemma 2.10. *Let $\lambda \geq 1$. For any $\varepsilon > 0$ there exists a constant $C > 0$ such that for r sufficiently large*

$$\mathbf{P}\left(V(\tilde{A}_{2r} \setminus \tilde{A}_r) \geq r^{\frac{\beta}{\beta-1}} \lambda\right) \geq C \lambda^{-\frac{(\beta-1+\varepsilon)}{2-\beta}}.$$

Proof. By Lemma 2.9 we have that

$$V(\tilde{A}_{2r} \setminus \tilde{A}_r) \geq \frac{1}{2} V(A_{2r} \setminus A_r) \geq \frac{1}{2} V(B_{2r} \setminus A_r),$$

and by construction $V(B_{2r} \setminus A_r)$ has the same distribution as $V(B_r)$. Hence, the statement is a consequence of Proposition 1.22 (ii). □

Now let v_r be the vertex in $\tilde{A}_{2r} \setminus \tilde{A}_r$ such that $\xi(v_r) = \sup_{v \in \tilde{A}_{2r} \setminus \tilde{A}_r} \xi(v)$ (this maximum is almost surely unique for all $r > 0$).

Lemma 2.11. *There exist constants $C_2, C_3 > 0$ such that*

$$\mathbf{P} \times \mathcal{P}(\limsup_{r \to \infty} E_r) = 1,$$

where the event E_r is defined as follows

$$E_r := \left\{ V(\tilde{A}_{2r} \setminus \tilde{A}_r) \geq C_2 r^{\frac{\beta}{\beta-1}}, \mathsf{Diam}(\tilde{A}_{2r} \setminus \tilde{A}_r) \leq C_3 r, \sup_{v \in \tilde{A}_{2r} \setminus \tilde{A}_r} \xi(v) \geq r^{\frac{\beta}{\alpha(\beta-1)}} (\log r)^{1/\alpha} \right\}.$$

Proof. We calculate, using Lemma 2.10, and the elementary relation $P(A \cap B) \geq$

$P(A) - P(B^c)$ for a probability function P and events A, B,

$$\begin{aligned}
&\mathbf{P}\times\mathcal{P}(E_r)\\
&\geq \mathbf{P}\left(V(\tilde{A}_{2r}\setminus\tilde{A}_r)\geq C_2 r^{\frac{\beta}{\beta-1}}, \mathsf{Diam}(\tilde{A}_{2r}\setminus\tilde{A}_r)\leq C_3 r\right)\\
&\quad \mathbf{P}\times\mathcal{P}\Bigg(\sup_{v\in\tilde{A}_{2r}\setminus\tilde{A}_r}\xi(v)\geq r^{\frac{\beta}{\alpha(\beta-1)}}(\log r)^{1/\alpha}\\
&\qquad\qquad \Big| V(\tilde{A}_{2r}\setminus\tilde{A}_r)\geq C_2 r^{\frac{\beta}{\beta-1}}, \mathsf{Diam}(\tilde{A}_{2r}\setminus\tilde{A}_r)\leq C_3 r\Bigg)\\
&\geq \left[\mathbf{P}\left(V(\tilde{A}_{2r}\setminus\tilde{A}_r)\geq C_2 r^{\frac{\beta}{\beta-1}}\right) - \mathbf{P}\Big(\mathsf{Diam}(\tilde{A}_{2r}\setminus\tilde{A}_r)\geq C_3 r\Big)\right]\\
&\quad \left[1-\left(1-r^{-\frac{\beta}{\beta-1}}(\log r)^{-1}\right)^{C_2 r^{\frac{\beta}{\beta-1}}}\right]\\
&\geq \left[C' C_2^{-\frac{\beta-1+\varepsilon}{2-\beta}} - C'' C_3^{\frac{-1}{\beta}}\right]\left[1-e^{-C_2(\log r)^{-1}}\right]\\
&\geq K(\log r)^{-1}
\end{aligned}$$

for sufficiently large r, provided C_2, C_3 are chosen appropriately (i.e. such that the first bracket is positive) and where we have used that $1 - e^{-2x} \geq x$ for x small enough. We have also used that $\mathsf{Diam}(A_r) \sim \mathsf{Diam}(A_{2r}\setminus A_r) \geq \mathsf{Diam}(\tilde{A}_{2r}\setminus\tilde{A}_r)$ and that by the proof of Proposition 1.27 we know

$$\mathbf{P}(\mathsf{Diam}(A_r)\geq 2r\lambda)\leq C\lambda^{\frac{-1}{\beta}}.$$

Now set $r_n := 2^n$, then the events $(E_{r_n})_{n\geq 1}$ are mutually independent and

$$\sum_{n=1}^{\infty}(\mathbf{P}\times\mathcal{P})(E_{r_n}) \geq K\sum_{n=1}^{\infty}(n\log 2)^{-1} = \infty.$$

Then applying the second Borel-Cantelli lemma (see Lemma A.2) we have

$$\mathbf{P}\times\mathcal{P}(\limsup_{r\to\infty} E_r) \geq \mathbf{P}\times\mathcal{P}(\limsup_{n\to\infty} E_{r_n}) = 1,$$

what concludes the proof. □

Proposition 2.12. *Take $\alpha < \frac{\beta}{\beta-1}$. Then, $\mathbf{P}\times\mathcal{P}$-almost surely $u(t,z) = \infty$ for all $z \in T_\infty$ and all $t > 0$.*

Proof. By the same argument as in [GM90, Theorem 2.1b)], it is sufficient to show that

$$\sup_{z\in\tilde{T}_\infty}\mathbb{E}_O\left[e^{\int_0^t \xi(X_s)\mathrm{d}s}\mathbb{1}\{X_t = z\}\right] = \infty. \tag{2.12}$$

To see this, first note that, by reversibility, it follows that for any $w \in \tilde{T}_\infty$,

$$\mathbb{E}_O\left[e^{\int_0^t \xi(X_s)ds}\mathbb{1}\{X_t = w\}\right] \asymp \mathbb{E}_w\left[e^{\int_0^t \xi(X_s)ds}\mathbb{1}\{X_t = O\}\right]. \tag{2.13}$$

Here the restriction to vertices of degree at most 4 ensures that the appropriate comparative constants are uniform over $w \in \tilde{T}_\infty$, and depend only on $\deg O$. Then, for any $v \in T_\infty$

$$\begin{aligned}
u(3t, v) &= \mathbb{E}_O\left[e^{\int_0^{3t} \xi(X_s)\mathrm{d}s}\mathbb{1}\{X_{3t} = v\}\right] \\
&\geq \sup_{w\in\tilde{T}_\infty} \mathbb{E}_O\left[e^{\int_0^{3t} \xi(X_s)\mathrm{d}s}\mathbb{1}\{X_t = w, X_{2t} = O, X_{3t} = v\}\right] \\
&= \sup_{w\in\tilde{T}_\infty} \mathbb{E}_O\left[e^{\int_0^{t} \xi(X_s)\mathrm{d}s}e^{\int_t^{2t} \xi(X_s)\mathrm{d}s}e^{\int_{2t}^{3t} \xi(X_s)\mathrm{d}s}\mathbb{1}\{X_t = w, X_{2t} = O, X_{3t} = v\}\right] \\
&\geq c\left[\sup_{w\in\tilde{T}_\infty} u(t, w)^2\right] u(t, v),
\end{aligned}$$

where we have used independence and (2.13).

Now, to prove (2.12), we let v_r be the point in $\tilde{A}_{2r} \setminus \tilde{A}_r$ maximising the potential and we let r be large enough and $r \geq \frac{3}{2}(t-1)$. Then, by Proposition 1.29 (ii), we have since $|v_r| \geq r \geq \frac{3}{2}(t-1)$ and r sufficiently large that

$$\begin{aligned}
u(t, v_r) &= \mathbb{E}_O\left[e^{\int_0^t \xi(X_s)\mathrm{d}s}\mathbb{1}\{X_t = v_r\}\right] \\
&\geq \exp\{\xi(v_r)\}\mathbb{P}(H_{v_r} \leq t-1, X_s = v_r \forall s \in [H_{v_r}, t]) \\
&\geq \left(\prod_{u\prec v_r} \frac{1}{\deg(u)}\right) e^{-|v_r|[\log|v_r| - \log(t-1)]}e^{-4t}e^{\xi(v_r)} \\
&= \exp\left\{-\sum_{u\prec v_r} \log(\deg u) - |v_r|[\log|v_r| - \log(t-1)] - 4t + \xi(v_r)\right\}.
\end{aligned}$$

Now, on the event E_r it holds $v_r \in B_{(2+C_3)r}$, thus by Lemma 1.28 it holds almost surely for r large enough and some constant $C_1 > 0$

$$\sum_{u\prec v_r} \log(\deg u) \leq \tilde{B}(2 + C_3)r + B\log((2 + C_3)r) \leq C_1 r.$$

Therefore, on the event E_r for r large enough

$$u(t, v_r) \geq \exp\left\{-C_1 r - C_3 r[\log(C_3 r) - \log(t-1)] - 4t + r^{\frac{\beta}{\alpha(\beta-1)}}(\log r)^{1/\alpha}\right\}.$$

If $\frac{\beta}{\beta-1} > \alpha$ this diverges as $r \to \infty$. Thus, since E_r occurs infinitely often $\mathbf{P} \times \mathcal{P}$-almost surely by Lemma 2.11, this establishes (2.12). □

Remark 2.13. If $\alpha = \frac{\beta}{\beta-1}$ the term above converges. However, one might be able to tweak the exponents in order to obtain non-existence in this case as well, at least with positive probability.

2.3.4 Main theorem and proof strategy

Having established the existence of solutions, we are ready to state the main theorem concerning the localisation of the solution in one vertex with high probability for $t \to \infty$.

Theorem 2.14. *Let $\alpha > \frac{\beta}{\beta-1}$ and let u be the unique non-negative solution to (2.7). There exists a process $(\hat{Z}_t)_{t\geq 0}$ with values in T_∞ such that*

$$\frac{u(t,\hat{Z}_t)}{U(t)} \to 1 \text{ in } \mathbf{P} \times \mathcal{P}\text{-probability as } t \to \infty. \tag{2.14}$$

Remark 2.15. The solution can not be localised at one vertex for all large times almost surely. This is clear since the process $\hat{Z}_t$ is not eventually constant, hence, at jump times of the process the solution has to relocalise continuously from one site to another during which it is concentrated in more than one vertex, cf. [KLMS09, Remark 1].

The proof strategy is similar to the one for the localisation result on $\mathbb{Z}^d$ as given in [KMS06]. We will define the process $\hat{Z}_t$ as the maximiser of the following functional

$$\psi_t(z) := \xi(z) - \frac{|z|}{t}\log\left(\frac{|z|}{et}\right), \qquad z \in T_\infty. \tag{2.15}$$

Now, according to the Feynman-Kac representation of the solution, the paths that move quickly to a vertex with a high potential and then stay there for a long time contribute the most to the total mass. Thus the functional (2.15) can be seen as balancing between the 'reward' of visiting a vertex z with potential $\xi(z)$ and the 'cost' of hitting that vertex in time t (which increases with the distance from the root). It is thus plausible that the maximiser of this functional over all vertices in T_∞ will define the localisation process. To prove this, we will split the solution $u(t,z)$ into three parts, that is, we distribute the paths in the Feynman-Kac formula into three categories

(i) those that leave a certain time-dependent ball by time t, they constitute $u_1(t,z)$;

(ii) those that do not leave this ball, but also not visit the 'optimal' site $\hat{Z}_t$, they constitute $u_2(t,z)$;

(iii) and those that do not leave the ball and visit $\hat{Z}_t$, they constitute $u_3(t,z)$.

It will turn out, that $u_1(t,z)$ and $u_2(t,z)$ do not contribute to the total mass for t going to infinity. Heuristically, for $u_1(t,z)$ this is due to the fact that the box is chosen large enough, such that leaving it is too 'costly'; while $u_2(t,z)$ is negligible since the contribution from the second optimal site, i.e. the second maximiser of

(2.15), is negligible compared to the one from $\hat{Z}_t$. It is thereby crucial that the time-dependent radius of the ball is chosen in such a way that $\hat{Z}_t$ is with high probability also the site of the maximal potential within this ball at time t. Finally, by means of a localisation result for the principal eigenfunction of the Anderson Hamiltonian, we will show that $u_3(t,z)$ localises on $\hat{Z}_t$, which concludes the argument.

Before starting with the analysis, let us recall the following definitions, which will appear throughout the following pages, $d := \frac{\beta}{\beta-1}$, $q := \frac{d}{\alpha-d}$ and

$$r(t) := \left(\frac{t}{\log t}\right)^{q+1}, \quad a(t) := \left(\frac{t}{\log t}\right)^{q}. \tag{2.16}$$

Here, $r(t)$ is the distance scale for the localisation site, while $a(t)$ is the scale for the maximum of the random functional (2.15), see Proposition 2.21. These scaling functions are the same as in the setting in $\mathbb{Z}^d$, see [KLMS09].

2.4 The concentration site

2.4.1 Maximizer of the potential ξ

In Section 2.5 we will prove a concentration result for the principal eigenfunction of the Hamiltonian appearing in the PAM. To understand the concentration result for this eigenfunction, we first need to derive some properties of the maximizer of the potential. Let us thus define for a finite set $\Lambda \subset T_\infty$

$$\begin{aligned} Z_\Lambda &:= \arg\max_{z\in\Lambda}\{\xi(z)\}, \\ \tilde{Z}_\Lambda &:= \arg\max_{z\in\Lambda}\{\xi(z) - \deg(z)\}. \end{aligned}$$

Note that these maximizers are almost surely unique. Furthermore, we also define the following gaps

$$\begin{aligned} g_\Lambda &:= \xi(Z_\Lambda) - \max_{z\in\Lambda, z\neq Z_\Lambda}\{\xi(z)\}, \\ \tilde{g}_\Lambda &:= \xi(\tilde{Z}_\Lambda) - \deg(\tilde{Z}_\Lambda) - \max_{z\in\Lambda, z\neq \tilde{Z}_\Lambda}\{\xi(z) - \deg(z)\}. \end{aligned}$$

Henceforth, we will consider $\Lambda = B_r$ and we start by proving an almost sure lower bound for the gap g_{B_r}.

Lemma 2.16. $\mathbf{P} \times \mathcal{P}$*-almost surely, for any* $\varepsilon > 0$ *there exists* $r_\varepsilon < \infty$ *such that*

$$g_{B_r} \geq c' r^{\frac{\beta}{\alpha(\beta-1)}} (\log r)^{\frac{-(1+\varepsilon)}{\alpha}}$$

for all $r \geq r_\varepsilon$. *In particular, almost surely* $g_{B_r} \to \infty$ *as* $r \to \infty$.

Proof. We will start by proving the following: Let $(X_i)_{i=1}^n$ be i.i.d. Pareto random variables with parameter α, and let $m_n = \arg\max_{1\le i\le n} X_i$, $g_n = \sup_{1\le i\le n} X_i - \sup_{1\le i\le n, i\ne m_n} X_i$. Then $\mathcal{P}(g_n \le y) \le e^{-(2y)^{-\alpha}n} + \alpha n y^{-\alpha} e^{-y^{-\alpha}(n-1)}$.

Recall that X_1 has the cumulative distribution function $F(x) = 1 - \mathcal{P}(X_1 \ge x) = 1 - x^{-\alpha}$, and the density $f(x) = \alpha x^{-(\alpha+1)}$. The maximum of $X_1, \dots, X_n$ has density $f_{X_{m_n}}(x) = nf(x)F(x)^{n-1}$ and distribution $F_{X_{m_n}}(x) = (F(x))^n$, so that for y sufficiently large

$$\begin{aligned}
&\mathcal{P}(g_n \le y)\\
&\le \mathcal{P}(X_{m_n} \le 2y) + \mathcal{P}(g_n \le y | X_{m_n} \ge 2y)\\
&= F(2y)^n + \int_{2y}^{\infty} \mathcal{P}(g_n \le y | X_{m_n} = x) f_{X_{m_n}}(x)\mathrm{d}x\\
&= F(2y)^n + \int_{2y}^{\infty} nf(x)F(x)^{n-1}\left[1 - \prod_{1\le i\le n, i\ne m_n} \mathcal{P}(X_i < x-y | X_i < x)\right]\mathrm{d}x\\
&= F(2y)^n + \int_{2y}^{\infty} nf(x)F(x)^{n-1}\left[1 - \left(\frac{F(x-y)}{F(x)}\right)^{n-1}\right]\mathrm{d}x\\
&= (1-(2y)^{-\alpha})^n + \int_{2y}^{\infty} \alpha n x^{-(\alpha+1)}\left[(1-x^{-\alpha})^{n-1} - (1-(x-y)^{-\alpha})^{n-1}\right]\mathrm{d}x,
\end{aligned}$$

defining $h(x) := (1-x^{-\alpha})^{n-1}$ we calculate further for y sufficiently large

$$\begin{aligned}
&\mathcal{P}(g_n \le y)\\
&\le (1-(2y)^{-\alpha})^n + \int_{2y}^{\infty} \alpha n x^{-(\alpha+1)}\left(\int_{x-y}^{x} h'(r)\mathrm{d}r\right)\mathrm{d}x\\
&\le (1-(2y)^{-\alpha})^n + \int_{2y}^{\infty} \alpha n x^{-(\alpha+1)} y h'(x-y)\mathrm{d}x\\
&\le (1-(2y)^{-\alpha})^n + \alpha n y^{-\alpha}[1-(1-y^{-\alpha})^{n-1}]\\
&\le (1-(2y)^{-\alpha})^n + \alpha n y^{-\alpha}(1-y^{-\alpha})^{n-1}\\
&\le e^{-(2y)^{-\alpha}n} + \alpha n y^{-\alpha} e^{-y^{-\alpha}(n-1)},
\end{aligned}$$

where we have used that $h'(x)$ is decreasing for x sufficiently large and that $1-(1-y^{-\alpha})^{n-1} \le (1-y^{-\alpha})^{n-1}$ for y sufficiently large. If $y = \frac{1}{2}\delta n^{\frac{1}{\alpha}}(\log n)^{\frac{-1}{\alpha}}$ for $\delta < 1$, then we have for all $n \ge N_\delta$

$$\mathcal{P}(g_n \le y) \le (1+\alpha 2^{\alpha}\delta^{-\alpha})(\log n) n^{-\delta^{-\alpha}}.$$

The right hand side is summable since $\delta^{-\alpha} > 1$. Therefore, since the volumes of the sets $(B_r)_{r\ge 1}$ are strictly increasing, we deduce by Borel-Cantelli that

$$\mathbf{P} \times \mathcal{P}\left(g_{B_r} \le \frac{1}{3}(V(B_r))^{\frac{1}{\alpha}}(\log(V(B_r))^{\frac{-1}{\alpha}} \text{ i.o.}\right) = 0.$$

Since it also follows from Corollary 1.23 (iii) that there exists $c > 0$ such that

$$\mathbf{P}\Big(V(B_r) \leq cr^{\frac{\beta}{\beta-1}}(\log\log r)^{-\frac{\beta}{\beta-1}} \text{ i.o.}\Big) = 0,$$

we deduce that, $\mathbf{P} \times \mathcal{P}$-almost surely, for any $\varepsilon > 0$ there exists $r_\varepsilon < \infty$ such that

$$g_{B_r} \geq c' r^{\frac{\beta}{\alpha(\beta-1)}} (\log r)^{\frac{-(1+\varepsilon)}{\alpha}}$$

for all $r \geq r_\varepsilon$. □

The following Lemma shows that with high probability the sites Z_{B_r} and $\tilde{Z}_{B_r}$ coincide.

Lemma 2.17. *As $r \to \infty$,*

$$\mathbf{P} \times \mathcal{P}(Z_{B_r} = \tilde{Z}_{B_r}) \to 1.$$

Proof. We instead show that $\mathbf{P} \times \mathcal{P}(\deg Z_{B_r} \leq g_{B_r}) \to 1$. This proves the result, since on this event we have for all $z \in B_r$ with $z \neq Z_{B_r}$ that

$$\begin{aligned}[\xi(Z_{B_r}) - \deg(Z_{B_r})] - [\xi(z) - \deg z] &= [\xi(Z_{B_r}) - \xi(z)] - [\deg(Z_{B_r}) - \deg z] \\ &\geq g_{B_r} - g_{B_r} + 1,\end{aligned}$$

from which it follows that $\tilde{Z}_{B_r} = Z_{B_r}$ (uniquely).

Since the potential is independent of the tree, we have that Z_{B_r} is uniform on the vertices of B_r. By Lemma 1.26 (ii) we have for r sufficiently large

$$\mathbf{P} \times \mathcal{P}(\deg(Z_{B_r}) \geq m) \leq c'' m^{-(\beta-1)} r^{\frac{-(1-\varepsilon)}{\beta-1}} + Ce^{-cr^{\varepsilon/\beta}} + c' m^{-\beta}.$$

In particular, choosing $m_r = c' r^{\frac{\beta-\varepsilon}{\alpha(\beta-1)}}$ we get that for r large enough

$$\mathbf{P} \times \mathcal{P}(\deg(Z_{B_r}) \geq m_r) \leq cr^{-\left(\frac{(\beta-\varepsilon)}{\alpha} + \frac{(1-\varepsilon)}{\beta-1}\right)} + c' r^{\frac{-\beta(\beta-\varepsilon)}{\alpha(\beta-1)}},$$

and by Lemma 2.16 for $r \to \infty$

$$\mathbf{P} \times \mathcal{P}(g_{B_r} \leq m_r) \to 0.$$

We therefore have that

$$\mathbf{P} \times \mathcal{P}(\deg(Z_{B_r}) \geq g_{B_r}) \leq \mathbf{P} \times \mathcal{P}(\deg(Z_{B_r}) \geq m_r) + \mathbf{P} \times \mathcal{P}(g_{B_r} \leq m_r) \to 0.$$

□

With this result we are able to lower bound the gap $\tilde{g}_{B_r}$ by $g_{B_r}/2$ with high probability.

Corollary 2.18. *As $r \to \infty$,*

$$\mathbf{P} \times \mathcal{P}\left(\tilde{g}_{B_r} \geq \frac{g_{B_r}}{2}\right) \to 1. \tag{2.17}$$

Proof. In the proof of Lemma 2.17 we can choose a deterministic sequence $m_r \to \infty$ so that $\mathbf{P} \times \mathcal{P}(g_{B_r} \leq 2m_r) \to 0$ as $r \to \infty$. With $\mathbf{P} \times \mathcal{P}(\deg(Z_{B_r}) \geq m_r) \to 0$, it then follows that

$$\mathbf{P} \times \mathcal{P}(2\deg(Z_{B_r}) \leq g_{B_r}) \to 1.$$

On the event $\{Z_{B_r} = \tilde{Z}_{B_r}\} \cap \{2\deg(Z_{B_r}) \leq g_{B_r}\}$ we have

$$\begin{aligned}\tilde{g}_{B_r} &= \xi(Z_{B_r}) - \deg(Z_{B_r}) - \max_{z \in B_r \setminus \{Z_{B_r}\}} \{\xi(z) - \deg(z)\} \\ &\geq g_{B_r} - \deg(Z_{B_r}) \geq \frac{g_{B_r}}{2},\end{aligned}$$

i.e. (2.17) follows by invoking Lemma 2.17. □

Later on, given a finite $\Lambda \subset T_\infty$, it will be useful to define the following set

$$V_\Lambda^h := \{v \in \Lambda \setminus \{\tilde{Z}_\Lambda\} : \xi(v) > \xi(\tilde{Z}_\Lambda) - \deg \tilde{Z}_\Lambda\}. \tag{2.18}$$

We can show that for B_r this set is empty with high probability.

Lemma 2.19. *As $r \to \infty$,*

$$\mathbf{P} \times \mathcal{P}(V_{B_r}^h = \emptyset) \to 1.$$

Proof. On the event $\{Z_{B_r} = \tilde{Z}_{B_r}\} \cap \{\deg Z_{B_r} \leq g_{B_r}\}$ we have for $z \neq Z_{B_r} = \tilde{Z}_{B_r}$

$$\left[\xi(\tilde{Z}_{B_r}) - \deg(\tilde{Z}_{B_r})\right] - \xi(z) = [\xi(Z_{B_r}) - \xi(z)] - [\deg(Z_{B_r})] \geq g_{B_r} - g_{B_r} = 0.$$

Hence, since $\{\deg Z_{B_r} \leq g_{B_r}\} \subset \{Z_{B_r} = \tilde{Z}_{B_r}\} \cap \{\deg Z_{B_r} \leq g_{B_r}\}$ the statement follows as in Lemma 2.17. □

2.4.2 Maximizer of the functional ψ_t

As motivated earlier the maximizer of the functional (2.15) will turn out to be the localisation process. That is, let us properly define

$$\hat{Z}_t := \arg\max_{z \in T_\infty} \psi_t(z).$$

Note that we will also denote $\hat{Z}_t$ as $\hat{Z}_t^{(1)}$. We also define the second maximizer $\hat{Z}_t^{(2)}$ as

$$\hat{Z}_t^{(2)} := \arg\max_{z \in T_\infty \setminus \{\hat{Z}_t^{(1)}\}} \psi_t(z).$$

Lemma 2.20. *$\mathbf{P}\times\mathcal{P}$-almost surely, $\hat{Z}_t$ is well defined.*

Proof. By Lemma 2.7 for any $\varepsilon>0$ there exists r_0 large enough such that almost surely

$$\sup_{v\in A_r}\xi(v)\le r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}+\varepsilon},\ \text{for all } r\ge r_0.$$

Let t be fixed and let $\varepsilon\in\left(0,1-\frac{1}{\alpha}\frac{\beta}{\beta-1}\right)$, then $\frac{1}{\alpha}\frac{\beta}{\beta-1}+\varepsilon<1$ and in particular there exists $r_1(t)>0$ such that for $r\ge r_1(t)$

$$r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}+\varepsilon}-\frac{r}{t}\log\left(\frac{r}{et}\right)<0,\ \text{for all } r\ge r_1(t).$$

Set $r(t)=\max\{r_0,r_1(t)\}$, then

$$\begin{aligned}\sup_{|z|\ge r(t)}\psi_t(z)&\le\sup_{|z|\ge r(t)}\left[\sup_{v\in A_{|z|}}\xi(v)-\frac{|v|}{t}\log\left(\frac{|v|}{et}\right)\right]\\&\le\sup_{r\ge r(t)}\left[r^{\frac{1}{\alpha}\frac{\beta}{\beta-1}+\varepsilon}-\frac{r}{t}\log\left(\frac{r}{et}\right)\right]<0.\end{aligned}$$

Hence, almost surely, ψ_t is positive only for finitely many z and therefore it attains its maximum. ☐

Proposition 2.21. *With high $\mathbf{P}\times\mathcal{P}$-probability as $t\to\infty$, for any $f>0, 0<g<\frac{fd}{2\alpha}, B>\frac{(q+1)\beta}{\alpha(\beta-1)}$ we have*

(i) $r(t)(\log\log(t))^{-f}\le|\hat{Z}_t|\le r(t)(\log\log t)^B$,

(ii) $a(t)(\log\log(t))^{-g}\le\psi_t(\hat{Z}_t)\le a(t)(\log\log t)^{\frac{Bd}{\alpha}}$.

Proof.

(i) **Upper bound.** Choose $0<\varepsilon<d-\alpha$, $A>\frac{1}{\alpha}\left(\frac{1}{\beta-1-2\varepsilon}+1\right), B>(q+1)A$, set $N_t:=\varepsilon^{-1}\frac{(\varepsilon+2)}{(\beta-1)\log 2}\log\log t, \tilde{N}_t:=\frac{B\log\log\log t}{\log 2}$, and consider a sequence of radii $r_n:=2^n r(t)$. We will show that, with high probability as $t\to\infty$:

(a) $\sup_{v\in B_{r_n}}\xi(v)\le\frac{r_n}{3et}\log\left(\frac{r_n}{3et}\right)$ for all $n>N_t$,

(b) $\sup_{v\in B_{r_n}}\xi(v)\le r_n^{\frac{d}{\alpha}}(\log\log r(t))^A$ for all $\tilde{N}_t\le n\le N_t$.

We deal with case (a) first. First recall from Corollary 1.23 (i) that $\mathbf{P}$-almost surely, $V(B_r)\le r^{\frac{\beta}{\beta-1}}(\log r)^{\frac{1+\varepsilon}{\beta-1}}$ for all sufficiently large r. $\mathbf{P}$-almost surely, we can therefore write for all sufficiently large t that

$$\begin{aligned}
&\mathcal{P}\Big(\exists n \geq N_t, v \in B_{r_n} : \xi(v) \geq \frac{r_n}{3et} \log\Big(\frac{r_n}{3et}\Big)\Big) \\
&\leq \sum_{n \geq N_t} V(B_{r_n}) \mathcal{P}\Big(\xi(O) \geq \frac{r_n}{3et} \log\Big(\frac{r_n}{3et}\Big)\Big) \\
&\leq \sum_{n \geq N_t} r_n^{\frac{\beta}{\beta-1}} (\log r_n)^{\frac{1+\varepsilon}{\beta-1}} \Big(\frac{r_n}{3et}\Big)^{-\alpha} \log\Big(\frac{r_n}{3et}\Big)^{-\alpha} \\
&\leq \sum_{n \geq N_t} C n^{\frac{1+\varepsilon}{\beta-1}} 2^{-n(\alpha - \frac{\beta}{\beta-1} - \varepsilon)} 2^{-n\varepsilon} (\log t)^{\frac{1+\varepsilon}{\beta-1}},
\end{aligned}$$

where $C > 0$ is some constant. Since $n \geq N_t$, we have by our choice of N_t that $2^{n\varepsilon} \geq (\log t)^{\frac{\varepsilon+2}{\beta-1}}$, so that

$$\begin{aligned}
&\mathcal{P}\Big(\exists n \geq N_t, v \in B_{r_n} : \xi(v) \geq \frac{r_n}{3et} \log\Big(\frac{r_n}{3et}\Big)\Big) \\
&\leq \sum_{n \geq N_t} C n^{\frac{1+\varepsilon}{\beta-1}} 2^{-n(\alpha - \frac{\beta}{\beta-1} - \varepsilon)} (\log t)^{\frac{-1}{\beta-1}} \\
&\leq C (\log t)^{\frac{-1}{\beta-1}} \sum_{n \geq 1} n^{\frac{1+\varepsilon}{\beta-1}} 2^{-n(\alpha - \frac{\beta}{\beta-1} - \varepsilon)} \\
&\to 0,
\end{aligned}$$

as $t \to \infty$. This proves (a).

We now turn to (b). By our choice of A, we can compute, using Proposition 1.22 (i),

$$\begin{aligned}
&\mathbf{P} \times \mathcal{P}\Big(\exists n \in [\tilde{N}_t, N_t], v \in B_{r_n} : \xi(v) \geq r_n^{\frac{d}{\alpha}} (\log\log r(t))^A\Big) \\
&\leq \sum_{n \in [\tilde{N}_t, N_t]} \Big[\mathbf{P}\Big(V(B_{r_n}) \geq r_n^d (\log\log r(t))^{\frac{1}{\beta-1-2\varepsilon}}\Big) \\
&\qquad + r_n^d (\log\log r(t))^{\frac{1}{\beta-1-2\varepsilon}} \mathcal{P}\Big(\xi(O) \geq r_n^{\frac{d}{\alpha}} (\log\log (r(t))^A\Big)\Big] \\
&\leq N_t \Big[(\log\log r(t))^{\frac{-(\beta-1-\varepsilon)}{\beta-1-2\varepsilon}} + (\log\log r(t))^{\frac{1}{\beta-1-2\varepsilon}} (\log\log (r(t))^{-A\alpha}\Big] \\
&= \varepsilon^{-1} \frac{(\varepsilon+2)}{(\beta-1)\log 2} \log\log t \Big[(\log\log r(t))^{\frac{-(\beta-1-\varepsilon)}{\beta-1-2\varepsilon}} + (\log\log r(t))^{\frac{1}{\beta-1-2\varepsilon} - A\alpha}\Big] \\
&\to 0,
\end{aligned}$$

as $t \to \infty$. This establishes (b). As a consequence of (a) and (b), we now claim that the following holds with high probability as $t \to \infty$:

(a′) $\sup_{v\in(B_{r_{N_t}})^c}\psi_t(v)\to-\infty$,

(b′) $\sup_{v\in B_{r_{N_t}}\setminus B_{r_{\tilde N_t}}}\psi_t(v)\to-\infty$.

Here $(B_r)^c$ denotes the complement of the ball B_r in T_∞. Note that (a′) is a straightforward deduction from (a). Indeed, if $v\in(B_{r_{N_t}})^c$ with $|v|\in[2^n r(t),2^{n+1}r(t)]$, for $n>N_t$, we have that

$$\begin{aligned}\psi_t(v)\le\sup_{z\in B_{2^{n+1}r(t)}}\xi(z)-\frac{|v|}{t}\log\left(\frac{|v|}{et}\right)&\le\frac{r_{n+1}}{3et}\log\left(\frac{r_{n+1}}{3et}\right)-\frac{|v|}{t}\log\left(\frac{|v|}{et}\right)\\&\le\frac{2|v|}{3et}\log\left(\frac{2|v|}{3et}\right)-\frac{|v|}{t}\log\left(\frac{|v|}{et}\right)\\&\le-\frac{1}{3}\frac{|v|}{t}\log\left(\frac{|v|}{et}\right)\\&\to-\infty,\end{aligned}$$

as $t\to\infty$, which establishes (a′). Point (b) similarly implies that

$$\xi(v)\le\sup_{z\in B_{r_{n+1}}}\xi(z)\le r_{n+1}^{\frac{d}{\alpha}}\left(\log\log r(t)\right)^A\le(2|v|)^{\frac{d}{\alpha}}\left(\log\log r(t)\right)^A,$$

for all $v\in B_{r_{N_t}}\setminus B_{r_{\tilde N_t}}$ with $|v|\in[2^n r(t),2^{n+1}r(t)]$. For each such v, we clearly have $|v|\ge 2^{\tilde N_t}r(t)=(\log\log t)^B r(t)$, so we can write

$$\begin{aligned}&\psi_t(v)\\&\le(2|v|)^{\frac{d}{\alpha}}\left(\log\log r(t)\right)^A-\frac{|v|}{t}\log\left(\frac{|v|}{et}\right)\\&=\frac{|v|}{t}\left(2^{\frac{d}{\alpha}}t|v|^{\frac{d}{\alpha}-1}\left(\log\log r(t)\right)^A-\log\left(\frac{|v|}{et}\right)\right)\\&\le\frac{|v|}{t}\left(2^{\frac{d}{\alpha}}t\left(\frac{t}{\log t}\right)^{-1}(\log\log t)^{B\left(\frac{d}{\alpha}-1\right)}\left(\log\log r(t)\right)^A\right.\\&\qquad\left.-q\log t+(q+1)\log\log t+1\right)\\&=\frac{|v|\log t}{t}\left(2^{\frac{d}{\alpha}}(\log\log t)^{B\left(\frac{-1}{q+1}\right)}\left(\log\log r(t)\right)^A-q+\frac{(q+1)\log\log t}{\log t}+\frac{1}{\log t}\right).\end{aligned}$$

This latter quantity converges to $-\infty$ as $t\to\infty$ by our choice of B. This therefore establishes (b′). To complete the proof of (i) we therefore note that $\psi_t(O)\ge 0$ for all positive t, which means that, on the high probability events considered above, the maximiser of ψ_t inside the ball $B_{2^{\tilde N_t}r(t)}$ is strictly greater than that outside the ball $B_{2^{\tilde N_t}r(t)}$, which implies that $|\hat Z_t|\le 2^{\tilde N_t}r(t)=(\log\log t)^B r(t)$.

Lower bound. Let $f > 0$ and $0 < \varepsilon < \frac{fd}{2\alpha}$. To prove the lower bound, we take $g := \frac{fd}{2\alpha} - \varepsilon$, in particular $f > g$, and we set $f(t) := (\log\log t)^{-f}, g(t) := (\log\log t)^{-g}$. We show that with high probability as $t \to \infty$:

(a) $\sup_{v \in B_{r(t)f(t)}} \psi_t(v) \le r(t)^{d/\alpha} (\log\log t)^{-\frac{fd}{2\alpha}}$,

(b) $\sup_{v \in B_{r(t)} \setminus A_{r(t)g(t)}} \psi_t(v) \ge C r(t)^{d/\alpha} (\log\log t)^{-g}$.

For every $\delta > 0$ with high probability it holds (see Lemma 2.6)

$$\sup_{v \in B_{r(t)f(t)}} \xi(v) \le [r(t)f(t)]^{d/\alpha} (\log\log t)^{\delta},$$

and thus for $\delta = \frac{fd}{2\alpha}$ with high probability

$$\sup_{v \in B_{r(t)f(t)}} \psi_t(v) \le r(t)^{d/\alpha} (\log\log t)^{-\frac{fd}{2\alpha}}.$$

This establishes (a). We now continue to prove (b). Let $r_n := r(t)2^{-n}$ for $1 \le n \le g\log^{(3)}(t)$ be a sequence of decreasing radii (from $r(t)/2$ to $r(t)2^{-g\log^{(3)}(t)} \ge r(t)g(t)$) and let $M := \frac{1}{\frac{\beta-1}{2-\beta}+2\varepsilon}$. Note that the volumes of the sets $B_{2r_n} \setminus A_{r_n}$ are mutually independent and distributed as the volumes of the sets B_{r_n}. Thus by Proposition 1.22 (ii) there exists $c_2 > 0$ such that

$$\begin{aligned}
&\mathbf{P} \times \mathcal{P}\left(\nexists n \le g\log^{(3)}(t) : V(B_{2r_n} \setminus A_{r_n}) \ge r_n^d (\log^{(3)}(t))^M\right) \\
&\le \left[1 - c_2 (\log^{(3)}(t))^{-M\left(\frac{\beta-1}{2-\beta}+\varepsilon\right)}\right]^{g\log^{(3)}(t)} \\
&\le \exp\left\{-c_2 g \log^{(3)}(t) (\log^{(3)}(t))^{-(1-\varepsilon')}\right\} \qquad (2.19) \\
&\to 0 \text{ for } t \to \infty.
\end{aligned}$$

Hence with high probability there exists a $n \le g\log^{(3)}(t)$ such that $V(B_{2r_n} \setminus A_{r_n}) \ge r_n^d (\log^{(3)}(t))^M$. On this event and with this n we calculate further

$$\begin{aligned}
&\mathbf{P} \times \mathcal{P}\left(\nexists v \in B_{2r_n} \setminus A_{r_n} : \xi(v) > k\frac{r_n}{t} \log\left(\frac{r_n}{et}\right)\right) \\
&\le \left[1 - k^{-\alpha} \frac{r_n^{-\alpha}}{t^{-\alpha}} \left(\log\left(\frac{r_n}{et}\right)\right)^{-\alpha}\right]^{r_n^d (\log^{(3)}(t))^M} \\
&\le \exp\left\{-k^{-\alpha} 2^{n(\alpha-d)} r(t)^{d-\alpha} t^{\alpha} (\log^{(3)}(t))^M q^{-\alpha} \log(t)^{-\alpha}\right\} \\
&= \exp\left\{-k^{-\alpha} 2^{n(\alpha-d)} (\log^{(3)}(t))^M q^{-\alpha}\right\} \qquad (2.20) \\
&\to 0 \text{ for } t \to \infty.
\end{aligned}$$

Thus with high probability there exists a $v \in V(B_{2r_n} \setminus A_{r_n})$ such that $\xi(v) > k\frac{r_n}{t}\log\left(\frac{r_n}{et}\right)$. Since $|v| \leq 2r_n$ this yields $\xi(v) > k'\frac{|v|}{t}\log\left(\frac{|v|}{et}\right)$ with $k' = k/3$. It follows further for this v and $k > 3$, using that $|v| \geq r(t)2^{-g\log^{(3)}(t)}$

$$\begin{aligned}\psi_t(v) &= \xi(v) - \frac{|v|}{t}\log\left(\frac{|v|}{et}\right)\\ &\geq (k'-1)\frac{|v|}{t}\log\left(\frac{|v|}{et}\right)\\ &\geq (k'-1)\frac{(\log\log(t))^{-g}r(t)}{t}\log\left(\frac{(\log\log(t))^{-g}r(t)}{et}\right).\end{aligned}$$

Hence, since for $1 \leq n \leq g\log^{(3)}(t)$ we have $B_{2r_n} \setminus A_{r_n} \subset B_{r(t)} \setminus A_{r(t)g(t)}$, it holds for t large enough

$$\begin{aligned}&\sup_{v\in B_{r(t)}\setminus A_{r(t)g(t)}} \psi_t(v)\\ &\geq (k'-1)\left(\frac{t}{\log t}\right)^{q+1}\frac{1}{t(\log\log(t))^g}\log\left(\frac{(\log\log(t))^{-g}r(t)}{et}\right)\\ &\geq (k'-1)\left(\frac{t}{\log t}\right)^{q}\frac{1}{\log(t)(\log\log(t))^g}(q-\tilde{\varepsilon})\log(t)\\ &= (k'-1)r(t)^{d/\alpha}\frac{(q-\tilde{\varepsilon})}{(\log\log(t))^g},\end{aligned} \tag{2.21}$$

i.e. statement (b) follows. From (a) and (b) we conclude that with high probability for $t \to \infty$

$$\sup_{v\in B_{r(t)f(t)}}\psi_t(v) < \sup_{v\in B_{r(t)}\setminus A_{r(t)g(t)}}\psi_t(v)$$

and therefore with high probability for $t \to \infty$

$$|\hat{Z}_t| > r(t)f(t) = r(t)(\log\log t)^{-f}.$$

(ii) It follows from the upper bound in part (i) and with Lemma 2.6 that for any $\delta > 0$, we have with high probability,

$$\psi_t(\hat{Z}_t) \leq \sup_{v\in A_{r(t)(\log\log(t))^B}}\xi(v) \leq r(t)^{\frac{d}{\alpha}}(\log\log t)^{\frac{Bd}{\alpha}+\delta} = a(t)(\log\log t)^{\frac{Bd}{\alpha}+\delta}.$$

Similarly, it follows from the proof of the lower bound in part (i) that with high probability,

$$\begin{aligned}\psi_t(\hat{Z}_t) \geq \sup_{v\in B_{r(t)}\setminus A_{r(t)g(t)}}\psi_t(v) &\geq (k'-1)r(t)^{d/\alpha}\frac{(q-\tilde{\varepsilon})}{(\log\log(t))^g}\\ &= (k'-1)(q-\tilde{\varepsilon})a(t)(\log\log(t))^{-g}.\end{aligned}$$

We can eliminate the constant by adding some sufficiently small δ to g, that is for any $\delta > 0$ we have with high probability

$$\psi_t(\hat{Z}_t) \geq a(t)(\log\log(t))^{-(g+\delta)}.$$

□

For $\varepsilon > 0$ and ε small enough such that $\frac{1+3\varepsilon}{\alpha} < 1 - 2\varepsilon$ let h_t, $t \in (0,\infty)$, be a function such that $(\log(t))^{-(1-2\varepsilon)} < h_t < (\log(t))^{-\frac{1+3\varepsilon}{\alpha}}$. In particular, $\lim_{t\to\infty} h_t = 0$. Furthermore, we define the random radius

$$R_t := |\hat{Z}_t|(1 + h_t).$$

This will be the radius of the ball with respect to which we will split the solution for the localisation argument in Section 2.6 (also recall the proof strategy as outlined in Subsection 2.3.4). The reason why these decay conditions on h_t are necessary will become clear in the proofs of Lemma 2.22 and Proposition 2.31.

We now show that with high probability $\hat{Z}_t$ is equal to $Z_{B_{R_t}}$, i.e. the maximizer of ψ_t is also the maximizer of the potential ξ in the ball with radius R_t.

Lemma 2.22.

$$\lim_{t\to\infty} \mathbf{P} \times \mathcal{P}(\hat{Z}_t = Z_{B_{R_t}}) = 1.$$

Proof. We start by proving that for t sufficiently large

$$\{\hat{Z}_t \neq Z_{B_{R_t}}\} \subset \left\{ g_{B_{R_t}} < \frac{2|\hat{Z}_t|h_t}{t} \log\left(\frac{|\hat{Z}_t|}{et}\right)\right\}.$$

On the event $\{\hat{Z}_t \neq Z_{B_{R_t}}\}$ we calculate

$$\begin{aligned}
g_{B_{R_t}} &= \xi(Z_{B_{R_t}}) - \max_{v\in B_{R_t}\setminus Z_{B_{R_t}}} \xi(v) \\
&\leq \xi(Z_{B_{R_t}}) - \xi(\hat{Z}_t) \\
&= \psi(Z_{B_{R_t}}) + \frac{|Z_{B_{R_t}}|}{t}\log\left(\frac{|Z_{B_{R_t}}|}{et}\right) - \psi(\hat{Z}_t) - \frac{|\hat{Z}_t|}{t}\log\left(\frac{|\hat{Z}_t|}{et}\right) \\
&< \frac{R_t}{t}\log\left(\frac{R_t}{et}\right) - \frac{|\hat{Z}_t|}{t}\log\left(\frac{|\hat{Z}_t|}{et}\right) \\
&= \frac{|\hat{Z}_t|}{t}\left[(1+h_t)\log\left(\frac{|\hat{Z}_t|(1+h_t)}{et}\right) - \log\left(\frac{|\hat{Z}_t|}{et}\right)\right] \\
&\leq \frac{|\hat{Z}_t|}{t}\left[h_t\log\left(\frac{|\hat{Z}_t|}{et}\right) + (1+h_t)\log\left(1+h_t\right)\right] \\
&\leq \frac{2|\hat{Z}_t|}{t}h_t\log\left(\frac{|\hat{Z}_t|}{et}\right),
\end{aligned}$$

where we have used that t is sufficiently large. Setting $r_1(t) := r(t)(\log\log(t))^{-f}$ and $r_2(t) := r(t)(\log\log(t))^B$, for some $B > \frac{q+1}{\alpha(\beta-1)}$, this implies

$$\begin{aligned}
&\{\hat{Z}_t \neq Z_{B_{R_t}}\} \\
&\subset \left\{|\hat{Z}_t| \notin [r_1(t), r_2(t)]\right\} \cup \left\{g_{B_{R_t}} < \frac{2|\hat{Z}_t|h_t}{t} \log\left(\frac{|\hat{Z}_t|}{et}\right), |\hat{Z}_t| \in [r_1(t), r_2(t)]\right\}.
\end{aligned} \tag{2.22}$$

We now show that almost surely for t sufficiently large

$$\left\{|\hat{Z}_t| \in [r_1(t), r_2(t)]\right\} \subset \left\{g_{B_{R_t}} \geq \frac{2|\hat{Z}_t|h_t}{t} \log\left(\frac{|\hat{Z}_t|}{et}\right)\right\}.$$

On the event $\left\{|\hat{Z}_t| \in [r_1(t), r_2(t)]\right\}$ it follows by definition that $R_t \in [r_1(t)(1+h_t), r_2(t)(1+h_t)]$. Now, by Lemma 2.16, almost surely for $\varepsilon > 0$ there exist $t_0 < \infty$ such that for all $t \geq t_0$ and t large enough

$$\begin{aligned}
g_{B_{R_t}} &\geq c' R_t^{d/\alpha} (\log R_t)^{-\frac{1+\varepsilon}{\alpha}} \\
&\geq c' r(t)^{d/\alpha} (\log\log(t))^{-fd/\alpha} (1+h_t)^{d/\alpha} \log\left(\frac{r(t)(1+h_t)}{(\log\log(t))^f}\right)^{-\frac{1+\varepsilon}{\alpha}} \\
&\geq c' \frac{r(t)\log(t)}{t} (\log\log(t))^{-fd/\alpha} [\log r(t) + \log(2)]^{-\frac{1+\varepsilon}{\alpha}} \\
&\geq C \frac{r(t)\log(t)}{t} (\log(t))^{-\frac{1+2\varepsilon}{\alpha}},
\end{aligned} \tag{2.23}$$

for some $C > 0$. Furthermore, on the event $\left\{|\hat{Z}_t| \in [r_1(t), r_2(t)]\right\}$ we also have

$$\begin{aligned}
\frac{2|\hat{Z}_t|h_t}{t} \log\left(\frac{|\hat{Z}_t|}{et}\right) &\leq 2h_t \frac{r(t)(\log\log(t))^B}{t} \log\left(\frac{r(t)(\log\log(t))^B}{et}\right) \\
&\leq 2(q+B) \frac{r(t)\log(t)}{t} (\log\log(t))^B h_t.
\end{aligned} \tag{2.24}$$

Now, we can choose $\varepsilon > 0$ small enough such that $h_t < (\log t)^{-\frac{1+3\varepsilon}{\alpha}}$ by definition of h_t. Hence, on the event $\left\{|\hat{Z}_t| \in [r_1(t), r_2(t)]\right\}$ for t sufficiently large it holds, by (2.23) and (2.24), that

$$g_{B_{R_t}} \geq \frac{2|\hat{Z}_t|h_t}{t} \log\left(\frac{|\hat{Z}_t|}{et}\right).$$

In particular, (2.22) then implies that almost surely there exists t_0 such that for all $t \geq t_0$

$$\left\{|\hat{Z}_t| \in [r_1(t), r_2(t)]\right\} \subset \{\hat{Z}_t = Z_{B_{R_t}}\}.$$

Applying Proposition 2.21 (i) concludes the proof of the statement. □

2.4.3 The gap

In this subsection we prove that the gap between the potential at the 'optimal' site $\hat{Z}_t^{(1)}$ and at the 'second best' site $\hat{Z}_t^{(2)}$ is asymptotically large.

Lemma 2.23. *For any $\varepsilon > 0$ there exist $t_\varepsilon < \infty$ and $\delta_\varepsilon > 0$ such that for all $t \geq t_\varepsilon$*

$$\mathbf{P} \times \mathcal{P}\left(\psi_t(\hat{Z}_t^{(1)}) - \psi_t(\hat{Z}_t^{(2)}) < \delta_\varepsilon a(t)\right) < \varepsilon.$$

Proof. Let us first prove the following claim.
Claim: For any given $\delta > 0$ we can choose $c > 0$ such that $\mathbf{P} \times \mathcal{P}\left(|\hat{Z}_t^{(1)}| < cr(t)\right) \leq \delta$.
Proof of Claim: The proof of this claim follows along the lines of the proof of the lower bound in Proposition 2.21 (i). Let $\delta > 0$ be given. We have for $c < \tilde{c} < 1$ and with $r_n = 2^{-n} r(t)$ for $n \leq \log(\tilde{c}^{-1})$

$$\begin{aligned}\mathbf{P} \times \mathcal{P}\left(|\hat{Z}_t^{(1)}| > cr(t)\right) &\geq \mathbf{P} \times \mathcal{P}\left(\sup_{v \in B_{cr(t)}} \psi_t(v) < \sup_{v \in B_{r(t)} \setminus A_{\tilde{c}r(t)}} \psi_t(v)\right) \\ &\geq 1 - [P_1 + P_2 + P_3]\end{aligned}$$

with

$$\begin{aligned}P_1 &:= \mathbf{P} \times \mathcal{P}\left(\sup_{v \in B_{cr(t)}} \psi_t(v) > C_3 r(t)^{d/\alpha}\right), \\ P_2 &:= \mathbf{P} \times \mathcal{P}\left(\nexists n \leq \log(\tilde{c}^{-1}) : V(B_{2r_n} \setminus A_{r_n}) \geq C_4 r_n^d\right), \\ P_3 &:= \mathbf{P} \times \mathcal{P}\Big(\nexists v \in B_{2r_n} \setminus A_{r_n} : \psi_t(v) > C_3 r(t)^{d/\alpha} \Big| \\ &\qquad\qquad \exists n \leq \log(\tilde{c}^{-1}) : V(B_{2r_n} \setminus A_{r_n}) \geq C_4 r_n^d\Big),\end{aligned}$$

for $C_3, C_4 > 0$. Now, from a similar calculation as for (2.20) (with $k = 4$) we have for $\varepsilon > 0$

$$\begin{aligned}&\mathbf{P} \times \mathcal{P}\left(\nexists v \in B_{2r_n} \setminus A_{r_n} : \xi(v) > 4\frac{r_n}{t} \log\left(\frac{r_n}{et}\right) \Big| \exists n \leq \log(\tilde{c}^{-1}) : V(B_{2r_n} \setminus A_{r_n}) \geq C_4 r_n^d\right) \\ &\leq \exp\left\{-4^{-\alpha} 2^{n(\alpha-d)} (q-\varepsilon)^{-\alpha} C_4\right\}.\end{aligned}$$

We now choose C_4 such that this probability is smaller than $\delta/3$. Furthermore, we set $C_3 := (4/3 - 1)\frac{q-\tilde{\varepsilon}}{\tilde{c}}$ for some $\tilde{\varepsilon} > 0$, and by comparison with (2.21) we can conclude that for this choice of C_4 and C_3 it holds $P_3 \leq \delta/3$. Regarding P_2 we can choose $\tilde{c}$ according to (2.19) such that

$$P_2 \leq \exp\left\{-c_2 \log(\tilde{c}^{-1}) C_4^{-\left(\frac{\beta-1}{2-\beta}+\varepsilon\right)}\right\} < \delta/3.$$

Finally, by Lemma 2.6 we may choose c small enough such that $P_1 \le \delta/3$ and $c < \tilde{c}$. This concludes the argument and proves the claim.

We can now prove the statement of the Lemma. Let $\varepsilon > 0$. We can choose $\eta_\varepsilon > 0$, $\chi_\varepsilon > 0$ and t_ε such that for all $t \ge t_\varepsilon$ we have

$$\frac{\left(1 - \left[\frac{cr(t)}{t} \log\left(\frac{cr(t)}{et}\right)\right]^{-1}\right)^{\alpha} - \eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha} \ge 1 - \varepsilon. \tag{2.25}$$

Furthermore, according to the claim above we can choose c such that

$$\mathbf{P} \times \mathcal{P}\left(|\hat{Z}_t^{(1)}| < cr(t)\right) \le \frac{\eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha}. \tag{2.26}$$

Also note that there exists $0 < \varepsilon' < q$ and $\tau_{\varepsilon'}$ sufficiently large such that for $t \ge \tau_{\varepsilon'}$

$$\begin{aligned} c\frac{r(t)}{t}\log\left(c\frac{r(t)}{et}\right) &= c\frac{a(t)}{\log(t)}\log\left(\frac{ct^q}{e(\log t)^{q+1}}\right) \\ &\ge c\frac{a(t)}{\log(t)}\log\left(t^{q-\varepsilon'}\right) = c(q-\varepsilon')a(t). \end{aligned} \tag{2.27}$$

Now, let $t'_\varepsilon := \max\{t_\varepsilon, \tau_{\varepsilon'}\}$ and set $\delta_\varepsilon := (q-\varepsilon')c\chi_\varepsilon$. Let $z \ge 0$ and $x := \delta_\varepsilon a(t)$. We calculate using (2.26) and the elementary relation $P(A|B) \le P(A) + P(B^c)$ for a probability function P and events A, B,

$$\begin{aligned} &\mathbf{P} \times \mathcal{P}\left(\psi_t(\hat{Z}_t^{(1)}) > x + z \middle| \psi_t(\hat{Z}_t^{(2)}) \in [z-1, z]\right) \\ &\ge \mathbf{P} \times \mathcal{P}\left(\psi_t(\hat{Z}_t^{(1)}) > x + z \middle| \psi_t(\hat{Z}_t^{(2)}) \in [z-1, z], \psi_t(\hat{Z}_t^{(1)}) > z-1, |\hat{Z}_t^{(1)}| > cr(t)\right) \\ &\qquad - \mathbf{P} \times \mathcal{P}(|\hat{Z}_t^{(1)}| \le cr(t)) \\ &\ge \inf_{T \in \mathbb{T}_\infty} \inf_{v \in (B^T_{cr(t)})^c} \mathcal{P}\left(\psi_t(\hat{Z}_t^{(1)}) > x + z \middle| \psi_t(\hat{Z}_t^{(2)}) \in [z-1, z], \psi_t(\hat{Z}_t^{(1)}) > z-1, \hat{Z}_t^{(1)} = v\right) \\ &\qquad - \frac{\eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha}, \end{aligned}$$

where B_r^T denotes the ball of radius r in the tree T.

Thus, we obtain for $t \geq t'_\varepsilon$

$$
\begin{aligned}
&\mathbf{P} \times \mathcal{P}\left(\psi_t(\hat{Z}_t^{(1)}) > x + z \,\middle|\, \psi_t(\hat{Z}_t^{(2)}) \in [z-1, z]\right) \\
&\geq \inf_{T \in \mathbb{T}_\infty} \inf_{v \in (B^T_{cr(t)})^c} \mathcal{P}\left(\psi_t(v) > x + z \,\middle|\, \psi_t(\hat{Z}_t^{(2)}) \in [z-1, z], \psi_t(v) > z - 1\right) - \frac{\eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha} \\
&= \inf_{T \in \mathbb{T}_\infty} \inf_{v \in (B^T_{cr(t)})^c} \mathcal{P}\left(\psi_t(v) > x + z \,\middle|\, \psi_t(v) > z - 1\right) - \frac{\eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha} \\
&= \inf_{T \in \mathbb{T}_\infty} \inf_{v \in (B^T_{cr(t)})^c} \frac{\mathcal{P}\left(\psi_t(v) > x + z\right)}{\mathcal{P}\left(\psi_t(v) > z - 1\right)} - \frac{\eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha} \\
&= \inf_{T \in \mathbb{T}_\infty} \inf_{v \in (B^T_{cr(t)})^c} \frac{\mathcal{P}\left(\xi(v) > x + z + \frac{|v|}{t}\log\left(\frac{|v|}{et}\right)\right)}{\mathcal{P}\left(\xi(v) > z - 1 + \frac{|v|}{t}\log\left(\frac{|v|}{et}\right)\right)} - \frac{\eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha} \\
&= \inf_{T \in \mathbb{T}_\infty} \inf_{v \in (B^T_{cr(t)})^c} \left(\frac{x + z + \frac{|v|}{t}\log\left(\frac{|v|}{et}\right)}{z - 1 + \frac{|v|}{t}\log\left(\frac{|v|}{et}\right)}\right)^{-\alpha} - \frac{\eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha} \\
&\geq \left(\frac{-1 + \frac{cr(t)}{t}\log\left(\frac{cr(t)}{et}\right)}{x + \frac{cr(t)}{t}\log\left(\frac{cr(t)}{et}\right)}\right)^{\alpha} - \frac{\eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha} \\
&\geq \left(\frac{1 - \left[\frac{cr(t)}{t}\log\left(\frac{cr(t)}{et}\right)\right]^{-1}}{1 + x\left[\frac{cr(t)}{t}\log\left(\frac{cr(t)}{et}\right)\right]^{-1}}\right)^{\alpha} - \frac{\eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha},
\end{aligned}
$$

where in the calculation above we have used the fact that

$$
f(z, r) = \left(\frac{z - 1 + \frac{r}{t}\log\left(\frac{r}{et}\right)}{x + z + \frac{r}{t}\log\left(\frac{r}{et}\right)}\right)^{\alpha},
$$

is increasing in z for fixed r and increasing in r for fixed z. By applying (2.27) and (2.25) we obtain

$$
\begin{aligned}
&\mathbf{P} \times \mathcal{P}\left(\psi_t(\hat{Z}_t^{(1)}) > x + z \,\middle|\, \psi_t(\hat{Z}_t^{(2)}) \in [z-1, z]\right) \\
&\geq \frac{\left(1 - \left[\frac{cr(t)}{t}\log\left(\frac{cr(t)}{et}\right)\right]^{-1}\right)^{\alpha} - \eta_\varepsilon}{(1+\chi_\varepsilon)^\alpha} \geq 1 - \varepsilon.
\end{aligned}
$$

It follows that for any $\varepsilon > 0$ there exists $t'_\varepsilon < \infty$ and $\delta_\varepsilon > 0$ such that for all $t \geq t'_\varepsilon$

$$
\mathbf{P} \times \mathcal{P}\left(\psi_t(\hat{Z}_t^{(1)}) - \psi_t(\hat{Z}_t^{(2)}) > \delta_\varepsilon a(t)\right) \geq 1 - \varepsilon,
$$

this concludes the proof. □

From the previous Lemma we can infer the following corollary.

Corollary 2.24. *If $\delta_t \to 0$ as $t \to \infty$, then*

$$\mathbf{P} \times \mathcal{P}\left(\psi_t(\hat{Z}_t^{(1)}) - \psi_t(\hat{Z}_t^{(2)}) < \delta_t a(t)\right) \to 0 \text{ as } t \to \infty.$$

2.5 Spectral results

In this section we derive a concentration result for the principal eigenfunction of the Anderson Hamiltonian restricted to a suitable subset of T_∞. More precisely, take some finite, connected set $\Lambda \subset T_\infty$, and consider equation (2.7) restricted to Λ with zero boundary condition. The corresponding solution u_Λ admits the spectral representation (cf. [MP16, Proposition 3.10])

$$u_\Lambda(t, v) = \sum_{k=1}^{V(\Lambda)} e^{\lambda_\Lambda^{(k)} t} \frac{\phi_\Lambda^{(k)}(O)\phi_\Lambda^{(k)}(v)}{\|\phi_\Lambda^{(k)}\|_2^2}, \quad t > 0, v \in \Lambda, \tag{2.28}$$

where $\lambda_\Lambda^{(1)} \geq \ldots \geq \lambda_\Lambda^{(|\Lambda|)}$ and $\phi^{(1)}, \ldots, \phi_\Lambda^{(|\Lambda|)}$ are the respective eigenvalues and corresponding orthogonal eigenfunctions of the Anderson Hamiltonian restricted to the class of functions supported on Λ, which we will denote as H_Λ, also recall Subsection 2.2.1. As before, we also have the Feynman-Kac representation for the solution on Λ, namely

$$u_\Lambda(t, v) = \mathbb{E}_O\left[\exp\left\{\int_0^t \xi(X_s)\,\mathrm{d}s\right\} \mathbb{1}\{X_t = v, \tau_\Lambda > t\}\right], \quad t > 0, v \in \Lambda.$$

Due to this *dual representation*, it is possible to transfer concentration results for the principal eigenfunction $\phi_\Lambda^{(1)}$ to the solution u_Λ. The objective of this section is therefore to prove such a concentration result. To do this, we will use the following path representation of the principal eigenfunction.

Lemma 2.25. *(see [MP16, Proposition 3.3]). For any $x, y \in \Lambda$,*

$$\frac{\phi_\Lambda^{(1)}(x)}{\phi_\Lambda^{(1)}(y)} = \mathbb{E}_x\left[\exp\left\{\int_0^{H_y} (\xi(X_s) - \lambda_\Lambda^{(1)})\ ds\right\} \mathbb{1}\{H_y < \tau_\Lambda\}\right].$$

In particular, if $\phi_\Lambda^{(1)}$ is normalised so that $\phi_\Lambda^{(1)}(y) = 1$, for some $y \in \Lambda$, then

$$\phi_\Lambda^{(1)}(x) = \mathbb{E}_x\left[\exp\left\{\int_0^{H_y} (\xi(X_s) - \lambda_\Lambda^{(1)})\ ds\right\} \mathbb{1}\{H_y < \tau_\Lambda\}\right]. \tag{2.29}$$

Henceforth, we consider $\Lambda = B_r$ and we assume that $\phi_{B_r}^{(1)}$ is normalised so that $\phi_{B_r}^{(1)}(\tilde{Z}_{B_r}) = 1$. We will show that $\phi_{B_r}^{(1)}$ is concentrated around the vertex $\tilde{Z}_{B_r}$. We first prove a lemma showing that we can restrict to a direct path in (2.29). For this, let us define the set of all paths from $x \in B_r$ to $\tilde{Z}_{B_r}$ as $\Gamma_{x,r}$ and, furthermore, let us denote the direct path from $x \in B_r$ to $\tilde{Z}_{B_r}$, excluding the endpoint $\tilde{Z}_{B_r}$, as $\gamma_{x,r}$. We also introduce the notation $\pi(X_{[0,t]})$, to denote the path that consists of all the sites visited by the random walk $(X_s)_{s\geq 0}$ between times 0 and t.

Lemma 2.26. *For all $x \in B_r$,*

$$\mathbf{P} \times \mathcal{P}\left(\phi^{(1)}_{B_r}(x) \leq \prod_{v\in\gamma_{x,r}} \frac{\deg v}{\tilde{g}_{B_r}}\right) \to 1 \text{ as } r \to \infty.$$

Proof. The proof follows a similar strategy to [KLMS09, Lemma 2.3]. Firstly, it follows from the *Rayleigh-Ritz formula* for the principal eigenvalue of the Anderson Hamiltonian that

$$\begin{aligned}
\lambda^{(1)}_{B_r} &= \sup\{\langle(\xi+\Delta)f, f\rangle_{\ell^2(T_\infty)} : f \in \ell^2(T_\infty), \operatorname{supp}(f) \subset B_r, ||f||_2 = 1\} \\
&\geq \sup_{z\in B_r}\{\langle(\xi+\Delta)\delta_z, \delta_z\rangle_{\ell^2(T_\infty)}\} \\
&= \sup_{z\in B_r}\{\xi(z) - \deg(z)\} \\
&= \xi(\tilde{Z}_{B_r}) - \deg(\tilde{Z}_{B_r}),
\end{aligned}$$

here $\|\cdot\|_2$ denotes the $\ell^2(T_\infty)$ norm. Then, by Lemma 2.25, we have

$$\begin{aligned}
\phi^{(1)}_{B_r}(x) &= \mathbb{E}_x\left[\exp\left\{\int_0^{H_{\tilde{Z}_{B_r}}} (\xi(X_s) - \lambda^{(1)}_{B_r})\, \mathrm{d}s\right\} \mathbb{1}\{H_{\tilde{Z}_{B_r}} < \tau_{B_r}\}\right] \\
&\leq \mathbb{E}_x\left[\exp\left\{\int_0^{H_{\tilde{Z}_{B_r}}} (\xi(X_s) - [\xi(\tilde{Z}_{B_r}) - \deg(\tilde{Z}_{B_r})]\, \mathrm{d}s\right\} \mathbb{1}\{H_{\tilde{Z}_{B_r}} < \tau_{B_r}\}\right].
\end{aligned}$$

Now, let $(T_v)_{v\in B_r}$ denote a set of independent exponentially distributed random variables, each with respective parameter $\deg(v)$, i.e. $T_v \sim \mathsf{Exp}(\deg(v))$. Furthermore, note that every path $\gamma \in \Gamma_{x,r}$ contains the set $\gamma_{x,r}$. Recall the definition of the set $V^h_{B_r}$ in equation (2.18). On the event $\{V^h_{B_r} = \emptyset\}$ we have that for every $v \in B_r \setminus \{\tilde{Z}_{B_r}\}$

$$\frac{\deg v}{\deg v - [\xi(v) + \deg \tilde{Z}_{B_r} - \xi(\tilde{Z}_{B_r})]} \leq 1.$$

Hence, we can compute, on the event $\{V^h_{B_r} = \emptyset\}$, that

$$
\begin{aligned}
&\phi^{(1)}_{B_r}(x) \\
&\leq \mathbb{E}_x\left[\exp\left\{\int_0^{H_{\tilde{Z}_{B_r}}} (\xi(X_s) - [\xi(\tilde{Z}_{B_r}) - \deg \tilde{Z}_{B_r}])\, \mathrm{d}s\right\} \mathbb{1}\{H_{\tilde{Z}_{B_r}} < \tau_{B_r}\}\right] \\
&\leq \sum_{\gamma\in\Gamma_{x,r}} \mathbb{P}_x\left(\pi\left(X_{[0,H_{\tilde{Z}_{B_r}}]}\right) = \gamma\right) \prod_{v\in\gamma\setminus\{\tilde{Z}_{B_r}\}} \mathbb{E}_x\left[\exp\left\{T_v[\xi(v) + \deg \tilde{Z}_{B_r} - \xi(\tilde{Z}_{B_r})]\right\}\right] \\
&\leq \sum_{\gamma\in\Gamma_{x,r}} \mathbb{P}_x\left(\pi\left(X_{[0,H_{\tilde{Z}_{B_r}}]}\right) = \gamma\right) \prod_{v\in\gamma\setminus\{\tilde{Z}_{B_r}\}} \frac{\deg v}{\deg v - [\xi(v) + \deg \tilde{Z}_{B_r} - \xi(\tilde{Z}_{B_r})]} \\
&\leq \prod_{v\in\gamma_{x,r}} \frac{\deg v}{\tilde{g}_{B_r}} \sum_{\gamma\in\Gamma_{x,r}} \mathbb{P}_x\left(\pi\left(X_{[0,H_{\tilde{Z}_{B_r}}]}\right) = \gamma\right) \prod_{v\in\gamma\setminus(\{\tilde{Z}_{B_r}\}\cup\gamma_{x,r})} \frac{\deg v}{\deg v - [\xi(v) + \deg \tilde{Z}_{B_r} - \xi(\tilde{Z}_{B_r})]} \\
&\leq \prod_{v\in\gamma_{x,r}} \frac{\deg v}{\tilde{g}_{B_r}},
\end{aligned}
$$

where we have used that the moment generating function of $X \sim \mathsf{Exp}(\lambda)$ is given by $\mathbb{E}[e^{tX}] = \frac{\lambda}{\lambda - t}$ for $t < \lambda$. The statement now follows by invoking Lemma 2.19. □

We will also need the following Lemma concerning the degrees of vertices in the direct path $\gamma_{x,r}$.

Lemma 2.27. *(i) There exist constants $B, \tilde{B}$ such that for any $C > 0$*

$$\mathbf{P}\times\mathcal{P}\left(\sup_{x\in B_r\setminus B(\tilde{Z}_{B_r},C)} \left\{\sum_{v\in\gamma_{x,r}} \log(\deg v) - [B\log(r) + \tilde{B}|\gamma_{x,r}|]\right\} \leq 0\right) \to 1,$$

as $r\to\infty$.

(ii) For any $\hat{B}, C > 0$

$$\mathbf{P}\times\mathcal{P}\left(\sup_{x\in B(\tilde{Z}_{B_r},C)\cap B_r\setminus\{\tilde{Z}_{B_r}\}} \left\{\sum_{v\in\gamma_{x,r}} \log(\deg v) - \hat{B}\log\log(r)\right\} \leq 0\right) \to 1,$$

as $r\to\infty$.

Proof. We work conditionally on $\tilde{Z}_{B_r}$. Note that this is a uniform vertex in B_r.

(i) The statement follows from Lemma 1.28.

(ii) Let $\varepsilon > 0$ and let us define $c := \beta - 1 - \varepsilon > 0$. Note that $\mathbf{E}[\deg v^c] < \infty$ for all $v \in B_r$. We set $A := \log \mathbf{E}[\deg v^c] < \infty$. Working backwards along the path $\gamma_{x,r}$ starting at $\tilde{Z}_{B_r}$ and ending at x, the degrees of the vertices are independent of each other. Let $x \in B(\tilde{Z}_{B_r}, C) \cap B_r \setminus \{\tilde{Z}_{B_r}\}$. This implies

$1 \le |\gamma_{x,r}| \le C$. We define for $\hat{B} > 0$, $\chi_{x,r} := \frac{\hat{B}\log\log r}{|\gamma_{x,r}|}$ and $f(r) := (\log r)^{c\hat{B}/2}$. We can now compute

$$\begin{aligned}
&\mathbf{P}\left(\sum_{v\in\gamma_{x,r}} \log(\deg v) \ge \chi_{x,r}|\gamma_{x,r}|\right)\\
&\le \mathbf{E}\left[\exp\left(c\sum_{v\in\gamma_{x,r}} \log(\deg v)\right)\right]\exp(-c\chi_{x,r}|\gamma_{x,r}|)\\
&\le \exp\left(A|\gamma_{x,r}| - c\hat{B}\log\log r\right)\\
&\le \exp(AC)\log r^{-c\hat{B}}.
\end{aligned}$$

Applying a union bound yields

$$\begin{aligned}
&\mathbf{P}\left(\sup_{x\in B(\tilde{Z}_{B_r},C)\cap B_r\setminus\tilde{Z}_{B_r}}\left\{\sum_{v\in\gamma_{x,r}} \log(\deg v) - \chi_{x,r}|\gamma_{x,r}|\right\} \ge 0\right)\\
&\le \exp(AC)f(r)\log r^{-c\hat{B}} + \mathbf{P}(V(B(\tilde{Z}_{B_r},C)) > f(r)),\\
&= \exp(AC)(\log r)^{-c\hat{B}/2} + \mathbf{P}(V(B(\tilde{Z}_{B_r},C)) > f(r)).
\end{aligned}$$

Invoking Lemma 1.26 we can conclude that the right hand side converges to 0 for $r \to \infty$.

□

We are now able to show that the principal eigenfunction is concentrated in $\tilde{Z}_{B_r}$ with high probability.

Lemma 2.28.

$$\|\phi^{(1)}_{B_r}\|_2^2 \sum_{z\in B_r\setminus\{\tilde{Z}_{B_r}\}} \phi^{(1)}_{B_r}(z) \to 0 \textit{ in } \mathbf{P}\times\mathcal{P}\textit{-probability as } r\to\infty.$$

Proof. We work conditionally on $\tilde{Z}_{B_r}$. Let $\varepsilon > 0$ and let us define $c := \beta - 1 - \varepsilon > 0$. Let B and $\tilde{B}$ be the constants appearing in Lemma 2.27 (i), that is, in particular it holds $B > \frac{1}{c}\left(2\frac{\beta+\varepsilon}{\beta-1} + \varepsilon\right)$ (see the proof of Lemma 1.28), and we set

$$C_{\alpha,\beta} := \frac{2\alpha(\beta-1)}{\beta}\left(B + \frac{\beta+\varepsilon}{\beta-1} + \varepsilon\right).$$

We split the sum into two parts, which we bound separately in the following,

$$\sum_{x\in B_r\setminus\{\tilde{Z}_{B_r}\}} \phi^{(1)}_{B_r}(x) = \sum_{x\in B_r\setminus B(\tilde{Z}_{B_r},C_{\alpha,\beta})} \phi^{(1)}_{B_r}(x) + \sum_{x\in B(\tilde{Z}_{B_r},C_{\alpha,\beta})\cap B_r\setminus\{\tilde{Z}_{B_r}\}} \phi^{(1)}_{B_r}(x). \tag{2.30}$$

Part 1: In particular, for $x \in B_r \setminus B(\tilde{Z}_{B_r}, C_{\alpha,\beta})$ it holds $|\gamma_{x,r}| \geq C_{\alpha,\beta}$. We set $v_r := r^{\frac{\beta+\varepsilon}{\beta-1}}$. On the event

$$\left\{ \sup_{x \in B_r \setminus B(\tilde{Z}_{B_r}, C_{\alpha,\beta})} \left\{ \sum_{v \in \gamma_{x,r}} \log(\deg v) - [B \log(r) + \tilde{B}|\gamma_{x,r}|] \right\} \leq 0 \right\} \cap \{V(B_r) \leq v_r\}$$
$$\cap \left\{ \tilde{g}_{B_r} \geq \frac{c'}{2} r^{\frac{\beta}{\alpha(\beta-1)}} (\log r)^{\frac{-(1+\varepsilon)}{\alpha}} \right\} \cap \left\{ \phi^{(1)}_{B_r}(x) \leq \prod_{v \in \gamma_{x,r}} \frac{\deg v}{\tilde{g}_{B_r}} \text{ for all } x \in B_r \right\},$$

we calculate for sufficiently large r

$$\begin{aligned}
&\sum_{x \in B_r \setminus B(\tilde{Z}_{B_r}, C_{\alpha,\beta})} \phi^{(1)}_{B_r}(x) \\
&\leq \sum_{x \in B_r \setminus B(\tilde{Z}_{B_r}, C_{\alpha,\beta})} \prod_{v \in \gamma_{x,r}} \frac{\deg v}{\tilde{g}_{B_r}} \\
&= \sum_{x \in B_r \setminus B(\tilde{Z}_{B_r}, C_{\alpha,\beta})} \exp\left(\sum_{v \in \gamma_{x,r}} [\log(\deg(v)) - \log(\tilde{g}_{B_r})] \right) \\
&\leq \sum_{x \in B_r \setminus B(\tilde{Z}_{B_r}, C_{\alpha,\beta})} \exp\left(B \log(r) + \tilde{B}|\gamma_{x,r}| - |\gamma_{x,r}| \left(\log(c'/2) + \frac{\beta \log r}{\alpha(\beta-1)} - \frac{1+\varepsilon}{\alpha} \log\log r \right) \right) \\
&\leq v_r \exp\left(-\log r \left[C_{\alpha,\beta} \frac{\beta}{2\alpha(\beta-1)} - B \right] - C_{\alpha,\beta} \left[\frac{\beta \log r}{2\alpha(\beta-1)} - \frac{1+\varepsilon}{\alpha} \log\log r - \tilde{B} \right] \right) \\
&\leq r^{-\varepsilon} \exp\left(-C_{\alpha,\beta} \left[\frac{\beta}{2\alpha(\beta-1)} \log r - \frac{1+\varepsilon}{\alpha} \log\log r - \tilde{B} \right] \right),
\end{aligned}$$

and the right hand side converges to 0 for $r \to \infty$. Note that we have assumed $c' > 2$, as we can choose c' large enough according to Lemma 2.16.

Part 2: Let $f(r) := \log(r)$ and $\hat{B} > 0$. On the event

$$\left\{ \sup_{x \in B(\tilde{Z}_{B_r}, C_{\alpha,\beta}) \cap B_r \setminus \{\tilde{Z}_{B_r}\}} \left\{ \sum_{v \in \gamma_{x,r}} \log(\deg v) - \hat{B} \log\log r \right\} \leq 0 \right\}$$
$$\cap \{V(B(\tilde{Z}_{B_r}, C_{\alpha,\beta})) \leq f(r)\}$$
$$\cap \left\{ \tilde{g}_{B_r} \geq \frac{c'}{2} r^{\frac{\beta}{\alpha(\beta-1)}} (\log r)^{\frac{-(1+\varepsilon)}{\alpha}} \right\} \cap \left\{ \phi^{(1)}_{B_r}(x) \leq \prod_{v \in \gamma_{x,r}} \frac{\deg v}{\tilde{g}_{B_r}} \text{ for all } x \in B_r \right\},$$

we calculate similar to the first part for sufficiently large r

$$\begin{aligned}
&\sum_{x\in B(\tilde{Z}_{B_r},C_{\alpha,\beta})\cap B_r\setminus\{\tilde{Z}_{B_r}\}} \phi^{(1)}_{B_r}(x) \\
&\le \sum_{x\in B(\tilde{Z}_{B_r},C_{\alpha,\beta})\cap B_r\setminus\{\tilde{Z}_{B_r}\}} \exp\left(\sum_{v\in\gamma_{x,r}}[\log(\deg(v))-\log(\tilde{g}_{B_r})]\right) \\
&\le f(r)\exp\left(\hat{B}\log\log r-\left(\log(c'/2)+\frac{\beta}{\alpha(\beta-1)}\log r-\frac{1+\varepsilon}{\alpha}\log\log r\right)\right) \\
&\le (\log r)^{1+\hat{B}+(1+\varepsilon)/\alpha} r^{-\frac{\beta}{\alpha(\beta-1)}},
\end{aligned}$$

and the right hand side converges to zero as $r\to\infty$. Again, we have assumed $c'>2$.

Thus, invoking Proposition 1.22, Lemma 1.26, Lemma 2.26, Lemma 2.27 and Corollary 2.18 it follows that both parts of the sum (2.30) converge to zero in $\mathbf{P}\times\mathcal{P}$-probability as $r\to\infty$. This completes the proof, since this also shows that for all sufficiently large r

$$\begin{aligned}
\|\phi^{(1)}_{B_r}\|_2^2 = \sum_{x\in B_r}\phi^{(1)}_{B_r}(x)^2 &= \phi^{(1)}_{B_r}(\tilde{Z}_{B_r})^2 + \sum_{x\in B_r\setminus\{\tilde{Z}_{B_r}\}}\phi^{(1)}_{B_r}(x)^2 \\
&\le 1+\left(\sum_{x\in B_r\setminus\{\tilde{Z}_{B_r}\}}\phi^{(1)}_{B_r}(x)\right)^2 \le 2.
\end{aligned}$$

□

As mentioned above, we want to transfer the concentration result for the principal eigenfunction to the solution of the PAM on a suitable set. Let us set

$$u_{B_r,\tilde{Z}_{B_r}}(t,v) := \mathbb{E}_0\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_t=v\}\mathbb{1}\{\tau_{B_r}>t\}\mathbb{1}\{H_{\tilde{Z}_{B_r}}\le t\}\right], \tag{2.31}$$

that is, the contribution to the solution of the PAM, which comes from paths in the Feynman-Kac formula that do not leave B_r and also visit the vertex $\tilde{Z}_{B_r}$. Then the following holds conditionally on T_∞.

Lemma 2.29. *For all $v\in T_\infty$ and $t>0$ we have*

$$u_{B_r,\tilde{Z}_{B_r}}(t,v) \le u_{B_r,\tilde{Z}_{B_r}}(t,\tilde{Z}_{B_r})\|\phi^{(1)}_{B_r}\|_2^2\phi^{(1)}_{B_r}(v). \tag{2.32}$$

The Lemma was proved in [GKM07, Theorem 4.1] (in a more general form) for $\mathbb{Z}^d$ and it works in the same way for T_∞. Nevertheless, for completeness, we will give the proof here as well.

Proof. For notational convenience we set $\Lambda := B_r$. First note that (2.32) follows trivially for $v \notin \Lambda$ as the left hand side is zero in this case. For $v = \tilde{Z}_\Lambda$, (2.32) follows due to the normalisation $\phi_\Lambda^{(1)}(\tilde{Z}_\Lambda) = 1$. Hence, we only need to prove the estimate for $v \in \Lambda \setminus \{\tilde{Z}_\Lambda\}$.

Thus, let $v \in \Lambda \setminus \{\tilde{Z}_\Lambda\}$ and we compute using time reversal and the Markov property of the random walk

$$
\begin{aligned}
&u_{\Lambda,\tilde{Z}_\Lambda}(t,v)\\
&= \mathbb{E}_v\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_t = 0\}\mathbb{1}\{\tau_\Lambda > t\}\mathbb{1}\{H_{\tilde{Z}_\Lambda} \le t\}\right]\\
&= \mathbb{E}_v\left[\exp\left\{\int_0^{H_{\tilde{Z}_\Lambda}} \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{H_{\tilde{Z}_\Lambda} \le t\}\mathbb{1}\{\tau_\Lambda > H_{\tilde{Z}_\Lambda}\}\right.\\
&\qquad \left.\times \mathbb{E}_{\tilde{Z}_\Lambda}\left[\exp\left\{\int_0^{t-u} \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_{t-u} = 0\}\mathbb{1}\{\tau_\Lambda > t-u\}\right]_{u = H_{\tilde{Z}_\Lambda}}\right].
\end{aligned}
\tag{2.33}
$$

Furthermore we have by time reversal and the Markov property at time u

$$
\begin{aligned}
&u_{\Lambda,Z_\Lambda}(t,\tilde{Z}_\Lambda)\\
&= \mathbb{E}_{\tilde{Z}_\Lambda}\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_t = 0\}\mathbb{1}\{\tau_\Lambda > t\}\right]\\
&\ge \mathbb{E}_{\tilde{Z}_\Lambda}\left[\exp\left\{\int_0^u \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_u = \tilde{Z}_\Lambda\}\mathbb{1}\{\tau_\Lambda > t\}\right]\\
&\qquad \times \mathbb{E}_{\tilde{Z}_\Lambda}\left[\exp\left\{\int_0^{t-u} \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_{t-u} = 0\}\mathbb{1}\{\tau_\Lambda > t-u\}\right]\\
&\ge e^{\lambda_\Lambda^{(1)}u}\frac{\phi_\Lambda^{(1)}(\tilde{Z}_\Lambda)\phi_\Lambda^{(1)}(\tilde{Z}_\Lambda)}{\|\phi_\Lambda^{(1)}\|_2^2}\\
&\qquad \times \mathbb{E}_{\tilde{Z}_\Lambda}\left[\exp\left\{\int_0^{t-u} \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_{t-u} = 0\}\mathbb{1}\{\tau_\Lambda > t-u\}\right],
\end{aligned}
\tag{2.34}
$$

where we have used the spectral representation (2.28) in the last inequality. Rearranging (2.34) we obtain a upper bound for the expectation, which we plug into (2.33) for $u = H_{\tilde{Z}_\Lambda}$ (also recall the normalisation $\phi_\Lambda^{(1)}(\tilde{Z}_\Lambda) = 1$). Using (2.29) we can thus conclude

$$
\begin{aligned}
u_{\Lambda,\tilde{Z}_\Lambda}(t,v) &\le \mathbb{E}_v\left[\exp\left\{\int_0^{H_{\tilde{Z}_\Lambda}} \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{H_{\tilde{Z}_\Lambda} \le \tau_\Lambda\}e^{-\lambda_\Lambda^{(1)}H_{\tilde{Z}_\Lambda}}\|\phi_\Lambda^{(1)}\|_2^2 u_{\Lambda,\tilde{Z}_\Lambda}(t,\tilde{Z}_\Lambda)\right]\\
&= \|\phi_\Lambda^{(1)}\|_2^2 u_{\Lambda,\tilde{Z}_\Lambda}(t,\tilde{Z}_\Lambda)\mathbb{E}_v\left[\exp\left\{\int_0^{H_{\tilde{Z}_\Lambda}} (\xi(X_s) - \lambda_\Lambda^{(1)})\mathrm{d}s\right\}\mathbb{1}\{H_{\tilde{Z}_\Lambda} \le \tau_\Lambda\}\right]\\
&= \|\phi_\Lambda^{(1)}\|_2^2 u_{\Lambda,\tilde{Z}_\Lambda}(t,\tilde{Z}_\Lambda)\phi_\Lambda^{(1)}(v),
\end{aligned}
$$

and the proof is complete. □

2.6 Localisation with high probability

As motivated in Subsection 2.3.4 we split $u(t,v)$ into three parts

$$u_1(t,v) = \mathbb{E}_0\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_t = v\}\mathbb{1}\{\tau_{B_{R_t}} \le t\}\right],$$
$$u_2(t,v) = \mathbb{E}_0\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_t = v\}\mathbb{1}\{\tau_{B_{R_t}} > t\}\mathbb{1}\{H_{\hat{Z}_t} > t\}\right],$$
$$u_3(t,v) = \mathbb{E}_0\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{X_t = v\}\mathbb{1}\{\tau_{B_{R_t}} > t\}\mathbb{1}\{H_{\hat{Z}_t} \le t\}\right],$$

then clearly it holds $u(t,v) = u_1(t,v) + u_2(t,v) + u_3(t,v)$. We will show that the contributions to the total mass of each of these terms away from the site $\hat{Z}_t$ vanishes, with high probability as time goes to infinity.

We first give a lower bound for the total mass $U(t)$.

Proposition 2.30. *With high* $\mathbf{P} \times \mathcal{P}$*-probability as* $t \to \infty$ *for every* $\varepsilon > 0$

$$\log U(t) \ge t\psi_t(\hat{Z}_t) + o\left(\frac{ta(t)}{(\log t)^{1-\varepsilon}}\right).$$

Proof. Note that for t large enough we can invoke Proposition 1.29 (ii), thus we have for any $\rho \in (0,1)$ that

$$\begin{aligned}
U(t) &\ge u(t,\hat{Z}_t) \\
&\ge \mathbb{P}\Big(H_{\hat{Z}_t} \le \rho t, X_s = \hat{Z}_t \forall s \in [H_{\hat{Z}_t}, H_{\hat{Z}_t} + (1-\rho)t]\Big)\exp\{\xi(\hat{Z}_t)(1-\rho)t\} \\
&\ge \mathbb{P}\Big(H_{\hat{Z}_t} \le \rho t\Big)\exp\{-t\deg(\hat{Z}_t) + \xi(\hat{Z}_t)(1-\rho)t\} \\
&\ge \exp\left\{-|\hat{Z}_t|\log\left(\frac{|\hat{Z}_t|}{\rho t}\right) - \sum_{u \prec \hat{Z}_t}\log(\deg u) - t\deg(\hat{Z}_t) + \xi(\hat{Z}_t)(1-\rho)t\right\}.
\end{aligned}$$

Note from Lemma 1.28 that there exists a constant C such that $\mathbf{P}$-almost surely, $\sum_{u \prec \hat{Z}_t}\log(\deg u) \le C|\hat{Z}_t|$ for all sufficiently large t. Thus, using Proposition 2.21 (i), we can infer

$$\mathbf{P} \times \mathcal{P}\left(\sum_{u \prec \hat{Z}_t}\log(\deg u) \ge r(t)(\log\log t)^{B+1}\right) \to 0 \quad \text{as } t \to \infty,$$

for $B > \frac{q+1}{\alpha(\beta-1)}$. Similarly, it follows from Lemma 2.22 and the fact that $Z_{B_{R_t}}$ is a uniform vertex in B_{R_t} that $\mathbf{P} \times \mathcal{P}\Big(\deg(\hat{Z}_t) \ge \log\log t\Big) \to 0$ as $t \to \infty$ by invoking

Lemma 1.26. Therefore, setting $\rho = \frac{e}{\log t}$ we deduce that with high probability as $t \to \infty$,

$$\begin{aligned} U(t) &\geq \exp\left\{-|\hat{Z}_t| \log\left(\frac{|\hat{Z}_t|}{\rho t}\right) - r(t)(\log\log t)^{B+1} - t\log\log t + \xi(\hat{Z}_t)(1-\rho)t\right\} \\ &= \exp\left\{t\psi_t(\hat{Z}_t) - |\hat{Z}_t|\log\log t - r(t)(\log\log t)^{B+1} - t\log\log t - \frac{et}{\log t}\xi(\hat{Z}_t)\right\}. \end{aligned}$$

Furthermore, we have by Proposition 2.21 (i) and the proof of (ii) that, with high probability for $t \to \infty$,

$$\begin{aligned} &|\hat{Z}_t|\log\log t + r(t)(\log\log t)^{B+1} + t\log\log t + \frac{et\xi(\hat{Z}_t)}{\log t} \\ &\leq 2r(t)(\log\log t)^{B+1} + a(t)\log\log t + \frac{eta(t)}{\log t}(\log\log t)^{\frac{Bd}{\alpha}+\delta} \\ &= ta(t)\left[\frac{2(\log\log t)^{B+1}}{\log t} + \frac{\log\log t}{t} + e\frac{(\log\log t)^{\frac{Bd}{\alpha}+\delta}}{\log t}\right], \end{aligned}$$

which completes the proof. □

We are now able to prove that the contribution from u_1 to the total mass is asymptotically negligible as $t \to \infty$.

Proposition 2.31.

$$\frac{\sum_{v\in T_\infty} u_1(t,v)}{U(t)} \to 0 \textit{ in } \mathbf{P}\times\mathcal{P}\textit{-probability as } t\to\infty.$$

Proof. Let

$$p > \max\left\{q+1, (q+1)\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)\right\}. \tag{2.35}$$

We first note that it is sufficient to bound the sum over all $v \in B_{r(t)(\log t)^p}$. Indeed, similar to the proof of Proposition 2.8 we have for any $\varepsilon > 0$ that $\mathbf{P}\times\mathcal{P}$-almost surely for t sufficiently large

$$\begin{aligned} &\sum_{v\in(B_{r(t)(\log t)^p})^c} u_1(t,v) \\ &\leq \sum_{r\geq r(t)(\log t)^p} \mathbb{P}(\tau_{A_r}\leq t)\exp\{t\sup_{v\in A_r}\xi(v)\} \\ &\leq \sum_{r\geq r(t)(\log t)^p} \exp\left\{-r\log\left(\frac{r}{et}\right)\right\}\exp\left\{tr^{\frac{d}{\alpha}}(\log r)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}\right\}. \end{aligned}$$

Now note that for $r = r(t)(\log t)^p$ and t sufficiently large we have

$$\begin{aligned} r\log\left(\frac{r}{et}\right) &= r(t)(\log t)^p \log\left(\frac{r(t)(\log t)^p}{et}\right) \\ &= r(t)(\log t)^p \log(t^q(\log t)^{p-q-1}/e) \\ &\geq qr(t)(\log t)^{p+1}, \end{aligned}$$

and, using that $\frac{d}{\alpha} + \frac{1}{q+1} = 1$,

$$\begin{aligned} tr^{\frac{d}{\alpha}}(\log r)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)} &= r(t)(\log t)^{1+\frac{pd}{\alpha}}(\log(r(t)(\log t)^p))^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)} \\ &\leq Cr(t)(\log t)^{1+\frac{pd}{\alpha}+\frac{p}{q+1}-\delta} \\ &= Cr(t)(\log t)^{1+p-\delta}, \end{aligned}$$

for some $\delta > 0$ (since we have a strict inequality in (2.35)). In particular this means that for t large enough

$$tr^{\frac{d}{\alpha}}(\log r)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)} < \frac{1}{2}r\log\left(\frac{r}{et}\right). \tag{2.36}$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}r}\left\{-\frac{1}{2}r\log\left(\frac{r}{et}\right) + tr^{\frac{d}{\alpha}}(\log r)^{\frac{1}{\alpha}\left(\frac{\beta^2}{\beta-1}+1+\varepsilon\right)}\right\} < 0,$$

for r large enough, (2.36) actually holds for all $r \geq r(t)(\log t)^p$ and t sufficiently large. Thus we have for t sufficiently large

$$\sum_{v\in(B_{r(t)(\log t)^p})^c} u_1(t,v) \leq \sum_{r\geq r(t)(\log t)^p} \exp\left\{-\frac{1}{2}r\log\left(\frac{r}{et}\right)\right\}.$$

Now, since for every $r \geq r(t)$ and t large enough

$$\begin{aligned} \exp\left\{-\frac{1}{2}(r+1)\log\left(\frac{r+1}{et}\right)\right\} &\leq \exp\left\{-\frac{1}{2}\log\left(\frac{r}{et}\right)\right\}\exp\left\{-\frac{1}{2}r\log\left(\frac{r}{et}\right)\right\} \\ &\leq \frac{1}{1+\frac{1}{2}\log\left(\frac{r(t)}{et}\right)}\exp\left\{-\frac{1}{2}r\log\left(\frac{r}{et}\right)\right\} \\ &\leq \frac{1}{2}\exp\left\{-\frac{1}{2}r\log\left(\frac{r}{et}\right)\right\}, \end{aligned}$$

we obtain

$$\begin{aligned}\sum_{v\in(B_{r(t)(\log t)^p})^c} u_1(t,v) &\le \sum_{r\ge r(t)} \exp\left\{-\frac{1}{2} r\log\left(\frac{r}{et}\right)\right\}\\ &\le \exp\left\{-\frac{1}{2} r(t)\log\left(\frac{r(t)}{et}\right)\right\}\sum_{r\ge 0}\frac{1}{2^r}\\ &\le 2\exp\left\{-\frac{1}{2} r(t)\log\left(\frac{r(t)}{et}\right)\right\}\to 0,\end{aligned}$$

as $t\to\infty$. Since $U(t)\ge 1$ eventually almost surely, this shows that the contribution of u_1 outside the ball of radius $r(t)(\log t)^p$ is asymptotically negligible as $t\to\infty$.

We next prove that, with high probability as $t\to\infty$,

$$\begin{aligned}&\log\left(\sum_{v\in B_{r(t)(\log t)^p}} u_1(t,v)\right)\\ &\le \max\left\{t\psi_t(\hat{Z}_t^{(2)}), t\xi_t(\hat{Z}_t^{(1)}) - R_t\log\left(\frac{R_t}{et}\right)\right\} + o\left(\frac{ta(t)}{\log t}\right).\end{aligned} \tag{2.37}$$

To do this, first note that by definition of u_1, Proposition 2.21 and Lemma 1.31, we have with high probability that

$$\begin{aligned}&\log\left(\sum_{v\in B_{r(t)(\log t)^p}} u_1(t,v)\right)\\ &\le \log\left(\sum_{r\in[R_t, r(t)(\log t)^p]} \mathbb{E}\left[\exp\left\{t\sup_{v\in B_r}\xi(v)\right\}\mathbb{1}\left\{\sup_{s\le t}|X_s| = r\right\}\right]\right)\\ &\le \log\left(\sum_{r\in[R_t, r(t)(\log t)^p]} \exp\left\{t\sup_{v\in B_r}\xi(v)\right\}\mathbb{P}\big(\tau_{B_{r-1}}\le t\big)\right)\\ &\le \max_{r\in[R_t, r(t)(\log t)^p]}\left\{t\sup_{v\in B_r}\xi(v) + \log\left(\mathbb{P}\big(\tau_{B_{r-1}}\le t\big)\right)\right\} + \log\left(r(t)(\log t)^p\right)\\ &\le t\max_{r\in[R_t, r(t)(\log t)^p]}\left\{\sup_{v\in B_r}\xi(v) - \frac{r}{t}\log\left(\frac{r}{et}\right)\right\} + o\left(\frac{ta(t)}{\log t}\right),\end{aligned}$$

where we have used that for positive real numbers $x_1,...,x_n$ it holds

$$\log\left(\sum_{i=1}^n x_i\right)\le \max_{i=1,\dots,n}\{\log(x_i)\} + \log(n).$$

Let $\hat{r}(t) \in [R_t, r(t)(\log t)^p]$ denote the radius for which the maximum is attained, and $\hat{v}(t) = \arg\max\{\xi(v) : v \in B_{\hat{r}(t)}\}$. If $\hat{v}(t) = \hat{Z}_t^{(1)}$, then

$$\begin{aligned} t \max_{r\in[R_t,r(t)(\log t)^p]} \left\{ \sup_{v\in B_r} \xi(v) - \frac{r}{t}\log\left(\frac{r}{et}\right)\right\} &= t\xi(\hat{Z}_t^{(1)}) - \hat{r}(t)\log\left(\frac{\hat{r}(t)}{et}\right) \\ &\le t\xi(\hat{Z}_t^{(1)}) - R_t\log\left(\frac{R_t}{et}\right). \end{aligned}$$

Otherwise, i.e. if $\hat{v}(t) \neq \hat{Z}_t^{(1)}$, we have

$$\begin{aligned} t \max_{r\in[R_t,r(t)(\log t)^p]} \left\{ \sup_{v\in B_r} \xi(v) - \frac{r}{t}\log\left(\frac{r}{et}\right)\right\} &= t\xi(\hat{v}(t)) - \hat{r}(t)\log\left(\frac{\hat{r}(t)}{et}\right) \\ &\le t\xi(\hat{v}(t)) - |\hat{v}(t)|\log\left(\frac{|\hat{v}(t)|}{et}\right) \\ &= t\psi_t(\hat{v}(t)) \\ &\le t\psi_t(\hat{Z}_t^{(2)}). \end{aligned}$$

This establishes (2.37). To complete the proof, we invoke Proposition 2.30 and note that, with high probability as $t \to \infty$ for $\varepsilon > 0$

$$\begin{aligned} &\frac{\sum_{v\in T_\infty} u_1(t,v)}{U(t)} \\ &\le \exp\left\{\max\left\{t\psi_t(\hat{Z}_t^{(2)}), t\xi(\hat{Z}_t^{(1)}) - R_t\log\left(\frac{R_t}{et}\right)\right\} - t\psi_t(\hat{Z}_t^{(1)}) + o\left(\frac{ta(t)}{(\log t)^{1-\varepsilon}}\right)\right\} \\ &= \exp\left\{\max\left\{t\left(\psi_t(\hat{Z}_t^{(2)}) - \psi_t(\hat{Z}_t^{(1)})\right), |\hat{Z}_t^{(1)}|\log\left(\frac{|\hat{Z}_t^{(1)}|}{et}\right) - R_t\log\left(\frac{R_t}{et}\right)\right\}\right. \\ &\qquad \left. + o\left(\frac{ta(t)}{(\log t)^{1-\varepsilon}}\right)\right\}. \end{aligned}$$

Now, recall from Corollary 2.24 that for any function η_t with $\eta_t \to 0$ as $t \to \infty$, we have that for $t \to \infty$

$$\mathbf{P} \times \mathcal{P}\left(\psi_t(\hat{Z}_t^{(1)}) - \psi(\hat{Z}_t^{(2)}) < \eta_t a(t)\right) \to 0.$$

In particular, choosing $\eta_t = (\log t)^{-\frac{1}{2}}$ and choosing $\varepsilon < \frac{1}{2}$ gives that

$$t\left(\psi_t(\hat{Z}_t^{(2)}) - \psi_t(\hat{Z}_t^{(1)})\right) + o\left(\frac{ta(t)}{(\log t)^{1-\varepsilon}}\right) \to -\infty$$

in probability as $t \to \infty$.

To deal with the second term, note that since $R_t = |\hat{Z}_t^{(1)}|(1+h_t)$ and since we can choose ε small enough such that $h_t \geq (\log t)^{-(1-2\varepsilon)}$, we have by Proposition 2.21 that, with high probability, for t sufficiently large

$$\begin{aligned}
&R_t \log\left(\frac{R_t}{et}\right) - |\hat{Z}_t^{(1)}| \log\left(\frac{|\hat{Z}_t^{(1)}|}{et}\right) + o\left(\frac{ta(t)}{(\log t)^{1-\varepsilon}}\right) \\
&\geq |\hat{Z}_t^{(1)}| h_t \log\left(\frac{|\hat{Z}_t^{(1)}|}{et}\right) + o\left(\frac{ta(t)}{(\log t)^{1-\varepsilon}}\right) \\
&\geq C|\hat{Z}_t^{(1)}| h_t \log t + o\left(\frac{ta(t)}{(\log t)^{1-\varepsilon}}\right) \\
&\geq Cta(t)\frac{(\log\log t)^{-f}}{(\log t)^{1-2\varepsilon}} + o\left(\frac{ta(t)}{(\log t)^{1-\varepsilon}}\right),
\end{aligned}$$

which diverges as $t \to \infty$. This completes the proof. □

Likewise, the contribution from u_2 to the total mass is asymptotically negligible as $t \to \infty$, as we will show in the following proposition.

Proposition 2.32.

$$\frac{\sum_{v\in T_\infty} u_2(t,v)}{U(t)} \to 0 \quad \textit{in } \mathbf{P}\times\mathcal{P}\textit{-probability as } t\to\infty. \tag{2.38}$$

Proof. We split the sum into two parts, that is, for $f > 0$,

$$\begin{aligned}
&\sum_{v\in T_\infty} u_2(t,v) \\
&\leq \mathbb{E}\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\} \mathbb{1}\{H_{\hat{Z}_t^{(1)}} > t\}\mathbb{1}\{\tau_{B_{R_t}} > t\}\right] \\
&\leq \sum_{r\leq R_t} \mathbb{E}\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\} \mathbb{1}\{H_{\hat{Z}_t^{(1)}} > t\}\mathbb{1}\{\sup_{s\leq t}|X_s| = r\}\right] \\
&= \sum_{r\leq r(t)(\log\log(t))^{-f}} \mathbb{E}\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\} \mathbb{1}\{H_{\hat{Z}_t^{(1)}} > t\}\mathbb{1}\{\sup_{s\leq t}|X_s| = r\}\right] \\
&\quad + \sum_{r(t)(\log\log(t))^{-f}\leq r\leq R_t} \mathbb{E}\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\} \mathbb{1}\{H_{\hat{Z}_t^{(1)}} > t\}\mathbb{1}\{\sup_{s\leq t}|X_s| = r\}\right].
\end{aligned} \tag{2.39}$$

For the first sum in (2.39), we note that by Lemma 2.6 with high $\mathbf{P}\times\mathcal{P}$-probability as $t \to \infty$ it holds

$$\log\left(\sum_{r \le r(t)(\log\log(t))^{-f}} \mathbb{E}\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\} \mathbb{1}\{H_{\hat{Z}_t^{(1)}} > t\}\mathbb{1}\{\sup_{s\le t}|X_s| = r\}\right]\right)$$
$$\le \log\left(\sum_{r \le r(t)(\log\log(t))^{-f}} \exp\left\{t \sup_{v\in B_r\setminus\{\hat{Z}_t^{(1)}\}} \xi(v)\right\}\right)$$
$$\le t \sup_{v\in B_{r(t)(\log\log(t))^{-f}}} \xi(v) + \log(r(t)(\log\log(t))^{-f})$$
$$= o\left(\frac{ta(t)}{(\log\log t)^c}\right),$$

for some $\frac{fd}{\alpha} > c > 0$. Furthermore, we know by Proposition 2.21 that with high $\mathbf{P} \times \mathcal{P}$-probability for $t \to \infty$, for $B > \frac{q+1}{\alpha(\beta-1)}$

$$R_t = |\hat{Z}_t^{(1)}|(1+h_t) \le r(t)(\log\log(t))^B(1+h_t) \le r(t)(\log(t))^p,$$

for some $p > 0$. Thus, for the second sum in (2.39) we can invoke Lemma 1.31, that is, with high $\mathbf{P} \times \mathcal{P}$-probability as $t \to \infty$

$$\log\left(\sum_{r(t)(\log\log(t))^{-f} \le r \le R_t} \mathbb{E}\left[\exp\left\{t \sup_{v\in B_r\setminus\{\hat{Z}_t^{(1)}\}} \xi(v)\right\} \mathbb{1}\{\sup_{s\le t}|X_s| = r\}\right]\right)$$
$$\le \log\left(\sum_{r(t)(\log\log(t))^{-f} \le r \le r(t)(\log t)^p} \exp\left\{t \sup_{v\in B_r\setminus\{\hat{Z}_t^{(1)}\}} \xi(v)\right\} \mathbb{P}(\tau_{B_{r-1}} \le t)\right)$$
$$\le \max_{\substack{r\in[r(t)(\log\log(t))^{-f},\\ r(t)(\log t)^p]}} \left\{t \sup_{v\in B_r\setminus\{\hat{Z}_t^{(1)}\}} \xi(v) + \log(\mathbb{P}(\tau_{B_{r-1}} \le t))\right\} + \log(r(t)(\log(t))^p)$$
$$\le \max_{\substack{r\in[r(t)(\log\log(t))^{-f},\\ r(t)(\log t)^p]}} \left\{t \sup_{v\in B_r\setminus\{\hat{Z}_t^{(1)}\}} \xi(v) - r\log\left(\frac{r}{et}\right) + o(r(t))\right\} + \log(r(t)(\log(t))^p)$$
$$\le t \max_{r\in[r(t)(\log\log(t))^{-f}, r(t)(\log t)^p]} \left\{\sup_{v\in B_r\setminus\{\hat{Z}_t^{(1)}\}} \xi(v) - \frac{r}{t}\log\left(\frac{r}{et}\right)\right\} + o(r(t)).$$

As argued in Proposition 2.31 we can conclude that

$$t \max_{r\in[r(t)(\log\log(t))^{-f}, r(t)(\log t)^p]} \left\{\sup_{v\in B_r\setminus\{\hat{Z}_t^{(1)}\}} \xi(v) - \frac{r}{t}\log\left(\frac{r}{et}\right)\right\} \le t\psi_t(\hat{Z}_t^{(2)}).$$

Hence, we have established that with high $\mathbf{P} \times \mathcal{P}$-probability for $t \to \infty$

$$\log\left(\sum_{v\in T_\infty} u_2(t,v)\right) \le t\psi_t(\hat{Z}_t^{(2)}) + o\left(\frac{ta(t)}{(\log\log t)^c}\right). \tag{2.40}$$

As outlined in Proposition 2.31, we can conclude from (2.40) by using Proposition 2.30 and Corollary 2.24 (with $\eta_t = (\log\log t)^{-c/2}$) that in $\mathbf{P} \times \mathcal{P}$-probability for $t \to \infty$

$$\frac{\sum_{v\in T_\infty} u_2(t,v)}{U(t)} \to 0.$$

This concludes the proof. $\square$

Finally, we show that u_3 localises in $\hat{Z}_t = \hat{Z}_t^{(1)}$ with high $\mathbf{P} \times \mathcal{P}$-probability as $t \to \infty$. For this, the spectral results that we derived in Section 2.5 become relevant.

Proposition 2.33.

$$\frac{\sum_{v\in T_\infty\setminus\{\hat{Z}_t\}} u_3(t,v)}{U(t)} \to 0 \quad \textit{in } \mathbf{P} \times \mathcal{P}\textit{-probability as } t \to \infty.$$

Proof. Note that on the event $\{\hat{Z}_t = \tilde{Z}_{B_{R_t}}\}$, u_3 is of the form (2.31) with $r = R_t$. Thus, by Lemma 2.29 it follows

$$u_3(t,v) \le u_3(t,\hat{Z}_t)\|\phi^{(1)}_{B_{R_t}}\|_2^2 \phi^{(1)}_{B_{R_t}}(v).$$

Hence,

$$\begin{aligned}\frac{\sum_{v\in T_\infty\setminus\{\hat{Z}_t\}} u_3(t,v)}{U(t)} &\le \frac{\sum_{v\in T_\infty\setminus\{\hat{Z}_t\}} u_3(t,\hat{Z}_t)\|\phi^{(1)}_{B_{R_t}}\|_2^2 \phi^{(1)}_{B_{R_t}}(v)}{u_3(t,\hat{Z}_t)} \\ &= \|\phi^{(1)}_{B_{R_t}}\|_2^2 \sum_{v\in T_\infty\setminus\{\hat{Z}_t\}} \phi^{(1)}_{B_{R_t}}(v),\end{aligned}$$

i.e. the statement of the proposition follows by Lemma 2.28 and Lemma 2.22, Lemma 2.17. $\square$

Proof of Theorem 2.14. As $u(t,v) = u_1(t,v) + u_2(t,v) + u_3(t,v)$, the statement of Theorem 2.14 is a direct consequence of Propositions 2.31, 2.32 and 2.33. $\square$

2.7 Outlook

In this chapter we showed that the parabolic Anderson model with Pareto potential on T_∞ localises with high probability in one vertex for t going to infinity. We conjecture that, just like in $\mathbb{Z}^d$ [KLMS09], our model will almost surely localise in two vertices. The almost sure analysis is more delicate in the following sense. In order to deal with rare events in an almost sure setting, a more complex random functional ψ_t needs to be defined. In particular, it will be crucial to make the role of the degree of the vertices explicit in ψ_t, because the larger the degree of a vertex the smaller is the probability that the random walk remains at that site for a long time. Furthermore, the proof of Theorem 2.14 relied heavily on the fact that the gap between $\psi_t(\hat{Z}_t^{(1)})$ and $\psi_t(\hat{Z}_t^{(2)})$ is with high probability asymptotically large (Lemma 2.23). Such a result can not hold almost surely, recall Remark 2.5. It is therefore crucial to consider the first three maximizers of a suitable functional for the almost sure setting and show that the gap between $\psi_t(\hat{Z}_t^{(2)})$ and $\psi_t(\hat{Z}_t^{(3)})$ is eventually almost surely large. This will lead to a finer splitting of the solution with respect to certain sets of paths in the Feynman-Kac representation. This is work in progress together with Eleanor Archer.

There are many other aspects of the model investigated above that we would like to analyse in the future. One questions concerns the so-called *ageing*. A system is aeging if the time span it stays in a certain state increases with time, i.e. one can tell the age of the system by observing it at the present time. In the context of the parabolic Anderson model we talk about ageing if the frequency with which the localisation site $\hat{Z}_t$ is changing is decreasing over time. For the PAM on $\mathbb{Z}^d$ with Pareto potential this was investigated in [MOS11], where the authors showed a linearly increasing dependency between the periods in which the solution nearly stays constant and time.

Another objective of future research is to consider different potentials for the PAM on T_∞. In particular, we would like to investigate whether, as in the $\mathbb{Z}^d$ case, also the Weibull distribution leads to one point localisation with high probability. An even further step would be to consider potentials with lighter tails, where localisation is less pronounced and the intermittency analysis becomes much more subtle.

Although the results of this chapter are restricted to critical Galton-Watson trees and in certain proofs we take advantage of the tree structure, we anticipate that the PAM should exhibit similar behaviour on other critical random graphs, such as the Uniform Infinite Planar Triangulation, critical Erdős-Rényi graphs, and the critical configuration model.

Chapter 3

Stochastic partial differential equations

The purpose of this chapter is twofold. Firstly, to give a relatively broad insight into the field of SPDEs, and secondly, to state certain specific results in this area that will be needed later on. For this reason the level of detail may very throughout the chapter. The presentation is by no means exhaustive and the interested reader is referred to the monographs [DPZ92, LR15, GM10, DW14, Cho14] and the references therein.

3.1 Stochastic analysis

3.1.1 Wiener processes

One of the most widely used stochastic process to incorporate noise in differential equations is the *Wiener process*, named after the American mathematician N. Wiener (1894 - 1964). Figure 3.1 shows the familiar image of a sample path of a real-valued Wiener process, also called *Brownian motion*. This process is defined as follows.

Definition 3.1 (Brownian motion)**.** A real-valued stochastic process $(B(t))_{t\geq 0}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called *Brownian motion* or *real-valued Wiener process*, if the following three conditions are fulfilled

(i) $B(0) = 0$,

(ii) B has $\mathbb{P}$-almost surely *continuous trajectories*,

(iii) For $t_0 = 0 < t_1 < ... < t_n$, the increments $B(t_i) - B(t_{i-1})$, $1 \leq i \leq n$, are independent with law $\mathcal{N}(0, t_i - t_{i-1})$.

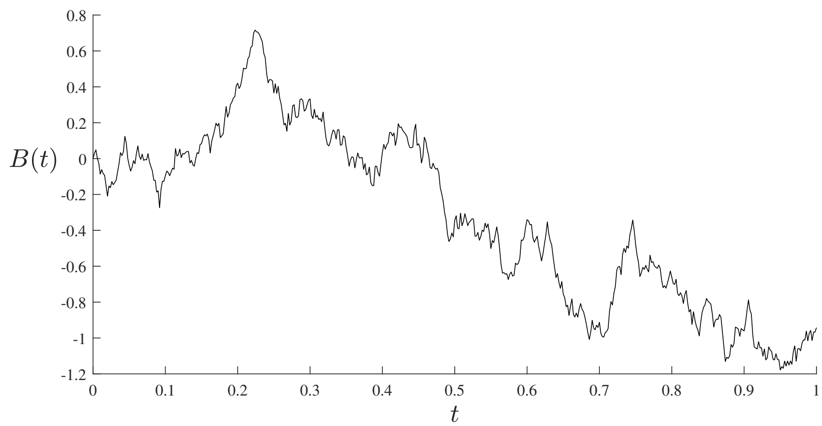

Figure 3.1: Sample path of a real-valued Brownian motion.

Remark 3.2. Without condition (ii) the process is sometimes called *pre-Brownian motion.* In fact, by *Kolmogorov's theorem* every pre-Brownian motion has a modification whose sample paths are continuous, and even locally Hölder continuous with exponent $\frac{1}{2} - \delta$ for every $\delta \in \left(0, \frac{1}{2}\right)$ (see Lemma A.5).

As outlined in the Introduction, we will often consider SPDEs of evolutionary type whose state space is an infinite-dimensional function space. Hence, the infinite-dimensional extension of the real-valued Wiener process, as defined below, will be a central object. For a detailed overview on this topic, we refer to [DQS11].

In the following let $(H, \|\cdot\|_H)$ be a separable Hilbert space with scalar product $\langle\cdot,\cdot\rangle_H$ and let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ be a filtered complete probability space.

Definition 3.3 (Isonormal process, see [vN08])**.** An *H-isonormal process* $\mathcal{W}$ is a family $\mathcal{W} = \{\mathcal{W}(h) : h \in H\}$ of real-valued random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ such that

(i) For every $h \in H$ the random variable $\mathcal{W}(h)$ has Gaussian law with mean zero.

(ii) $\mathbb{E}[\mathcal{W}(h_1)\mathcal{W}(h_2)] = \langle h_1, h_2\rangle_H$ for every $h_1, h_2 \in H$.

Note that all isonormal processes are linear.

Example 3.4 (Brownian motion)**.** *Let $H = L^2([0,T])$ and $\mathcal{W}$ be the associated H-isonormal process. Then $B(t) := \mathcal{W}(\mathbf{1}_{[0,t]})$ defines a real-valued Brownian motion on $[0,T]$.*

Definition 3.5 (H-cylindrical Brownian motion)**.** Set $H_T := L^2([0,T];H)$. A H_T-isonormal process is called a *H-cylindrical Brownian motion on* $[0,T]$.

Example 3.6 (Spatially homogeneous Gaussian noise, cf. [DQS11, Section 2.2]). *Let U be the completion of $C_0^\infty(\mathbb{R}^d)$ with respect to the semi-norm $\|\cdot\|_U$ associated to the semi-inner product*

$$\langle \varphi_1, \varphi_2 \rangle_U = \int_{\mathbb{R}^d} \Lambda(\mathrm{d}x)(\varphi_1 * \tilde{\varphi}_2)(x), \quad \varphi_1, \varphi_2 \in C_0^\infty(\mathbb{R}^d),$$

where $*$ denotes convolution, $\tilde{\varphi}_2(x) = \varphi_2(-x)$ and Λ is a symmetric, non-negative definite measure on $\mathbb{R}^d$. U is a separable Hilbert space. Then $\mathcal{W}$, the U_T-isonormal process or U-cylindrical Brownian motion on $[0,T]$, is called white in time, spatially homogeneous Gaussian noise. In particular, if $\Lambda(\mathrm{d}x) = f(x)\mathrm{d}x$ then, for $h_1, h_2 \in U_T$,

$$\begin{aligned}\mathbb{E}\left[\mathcal{W}(h_1)\mathcal{W}(h_2)\right] &= \int_0^T \langle h_1(t), h_2(t) \rangle_U \ \mathrm{d}t \\ &= \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h_1(t,x) f(x-y) h_2(t,y) \ \mathrm{d}x \ \mathrm{d}y \ \mathrm{d}t,\end{aligned}$$

what explains the notion spatially homogeneous (or translation invariant). This is a type of spatially coloured/correlated noise.

Definition 3.7 (Cylindrical Wiener process, see [DQS11]). Let $Q \in L(H)$ be symmetric (self-adjoint) and non-negative definite. A *cylindrical Wiener process* on H is a family of real-valued random variables $W_H := \{W_H(t)h : t \geq 0, h \in H\}$ such that

(i) $(W_H(t)h)_{t \in \mathbb{R}^+}$ is a real-valued Brownian motion for all $h \in H$.

(ii) For all $t_1, t_2 \in \mathbb{R}^+$ and $h_1, h_2 \in H$ we have

$$\mathbb{E}[W_H(t_1)h_1 \ W_H(t_2)h_2] = \min\{t_1, t_2\} \langle Q h_1, h_2 \rangle_H.$$

Q is called the *covariance operator*. In case $Q = \mathsf{Id}_H$, we call W_H a *standard cylindrical Wiener process*.

To every H-cylindrical Brownian motion, we can associate a standard cylindrical Wiener process according to the following proposition.

Proposition 3.8 (see [DQS11, Proposition 2.5]). *Let $\mathcal{W}_H$ be a H-cylindrical Brownian motion. For $t \in [0,T]$ and $h \in H$ we set*

$$W_H(t)h := \mathcal{W}_H(\mathbf{1}_{[0,t]} \otimes h),$$

where $\otimes$ denotes the tensor product, as we can identify H_T with $L^2([0,T]) \otimes H$. Then $\{W_H(t)h : t \in [0,T], h \in H\}$ is a standard cylindrical Wiener process on H.

Remark 3.9. In particular, to every spatially homogeneous Gaussian noise that is white in time (Example 3.6) we can associate a cylindrical Wiener process in a particular Hilbert space.

Example 3.10 (Space-time white noise)**.** *Let $H := L^2(D)$ with $D \subset \mathbb{R}^d$ open. A standard cylindrical Wiener process on H provides the mathematical model for space-time white noise. In particular, we have*

$$\mathbb{E}\left[W_H(t_1)h_1 \; W_H(t_2)h_2\right] = \min\{t_1, t_2\}\langle h_1, h_2\rangle_{L^2(D)}.$$

Loosely speaking, space-time white noise on D is the time derivative of a standard cylindrical Wiener process on H. We refer to [PZ07, Section 7.1.2] for details.

Remark 3.11. Let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of H and $(\beta_k)_{k\in\mathbb{N}}$ be a sequence of independent Brownian motions. Then

$$W_H(t)h := \sum_{k=1}^{\infty} \langle h, e_k\rangle_H \beta_k(t), \quad \text{for } h \in H,$$

defines a cylindrical Wiener process on H. However, the series

$$\sum_{k=1}^{\infty} \beta_k(t) e_k,$$

does not necessarily converge in $L^2((\Omega, \mathcal{F}, \mathbb{P}); H)$, i.e. it does not necessarily describe a genuine H-valued Gaussian process.

The following definition gives a straightforward extension of the real-valued Wiener process to H-valued Wiener processes. This characterization is based on trace-class operators; see Appendix C.2 for details on nuclear and Hilbert-Schmidt operators.

Definition 3.12 (Q-Wiener process)**.** Let $Q \in L(H)$ be non-negative definite, symmetric and of trace class. A H-valued stochastic process $(W_Q(t))_{t\in[0,T]}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called *Q-Wiener process* if

(i) $W_Q(0) = 0$.

(ii) W_Q has $\mathbb{P}$-almost surely continuous trajectories.

(iii) The increments $W_Q(t) - W_Q(s)$ are independent and have Gaussian laws with mean zero and covariance operator $(t-s)Q$, for all $0 \le s \le t \le T$. This means, that for any $h \in H$ and $0 \le s \le t \le T$, the real-valuesd random variable $\langle W_Q(t) - W_Q(s), h\rangle_H$ is Gaussian with mean zero and variance $(t-s)\langle Qh, h\rangle_H$.

Proposition 3.13 (Representation of a Q-Wiener process, see [LR15, Proposition 2.1.6])**.** *Let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of H where $Qe_k = \lambda_k e_k$, $k \in \mathbb{N}$. Then a H-valued stochastic process $(W_Q(t))_{t\in[0,T]}$, is a Q-Wiener process if and only if*

$$W_Q(t) = \sum_{k\in\mathbb{N}} \sqrt{\lambda_k}\beta_k(t)e_k, \quad t \in [0,T], \tag{3.1}$$

with independent Brownian motions β_k, $k \in \{n \in \mathbb{N} : \lambda_n > 0\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. That is, the series converges in $L^2((\Omega, \mathcal{F}, \mathbb{P}), H)$ for every $t \in [0,T]$. The series converges in $L^2((\Omega, \mathcal{F}, \mathbb{P}); C([0,T], H))$, i.e.

$$\lim_{n\to\infty} \mathbb{E}\left(\sup_{t\in[0,T]} \left\| \sum_{k=1}^{n} \sqrt{\lambda_k}\beta_k(t)e_k - W_Q(t)\right\|_H^2\right) = 0.$$

In particular, this means that the process has a $\mathbb{P}$-a.s. continuous version.

Definition 3.14. A Q-Wiener process $(W_Q(t))_{t\in[0,T]}$ is called a Q-Wiener process with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$, if W_Q is adapted to the filtration and $W_Q(t) - W_Q(s)$ is independent of $\mathcal{F}_s$ for all $0 \le s \le t \le T$.

Remark 3.15. The covariance of W_Q is given as

$$\begin{aligned}\mathbb{E}[\langle W_Q(t), W_Q(s)\rangle_H] &= \mathbb{E}\left[\sum_{k\in\mathbb{N}} \lambda_k \beta_k(t)\beta_k(s)\langle e_k, e_k\rangle_H\right] \\ &= \sum_{k\in\mathbb{N}} \lambda_k \mathbb{E}[\beta_k(t)\beta_k(s)] = \min\{s,t\}\mathsf{Tr}\ Q < \infty,\end{aligned}$$

for $s, t \in [0,T]$. In particular, the variance is given by $\mathbb{E}[\|W_Q(t)\|_H^2] = t\mathsf{Tr}\ Q$. The operator Q is the *covariance operator* (at time $t = 1$).

By Proposition 3.13 a H-valued Q-Wiener process can be represented as the following series

$$W_Q(t) = \sum_{k\in\mathbb{N}} \beta_k(t)a_k,$$

where $(\beta_k)_{k\in\mathbb{N}}$ are independent real-valued Wiener processes and $(a_k)_{k\in\mathbb{N}}$ is an orthonormal basis of $Q^{1/2}(H) =: H_0$. This series converges in $L^2((\Omega, \mathcal{F}, \mathbb{P}), H)$ because the embedding operator $H_0 \hookrightarrow H$ is Hilbert-Schmidt. We will denote the space of all Hilbert-Schmidt operators between two separable Hilbert spaces U and K as $(L_2(U,K), \|\cdot\|_{L_2(U,K)})$, see Appendix C.2 for a precise definition. There is a natural way to associate to every H-valued Q-Wiener process a cylindrical Wiener process on H. More precisely, for any $h \in H$, $t \ge 0$, we set

$$W_H(t)h := \langle W_Q(t), h\rangle_H, \tag{3.2}$$

Then $\{W_H(t)h : t \geq 0, h \in H\}$ is a cylindrical Wiener process on H with covariance operator Q.

However, not every cylindrical Wiener process can be associated to a Q-Wiener process. More precisely, recall from Remark 3.11 that a cylindrical Wiener process on H can be defined via

$$W_H(t)h = \sum_{k=1}^{\infty} \langle h, e_k \rangle_H \beta_k(t), \qquad \text{for } h \in H.$$

However, the series

$$\sum_{k=1}^{\infty} \beta_k(t) e_k, \tag{3.3}$$

might not converge in $L^2((\Omega, \mathcal{F}, \mathbb{P}); H)$. In particular, the following theorem holds.

Theorem 3.16 (see [MP14, p.177])**.** *Let H be a separable Hilbert space and W_H a cylindrical Wiener process on H with covariance Q. Then the following are equivalent*

(i) W_H is associated to a Q-Wiener process $(W_Q(t))_{t\geq 0}$, in the sense of (3.2)*.*

(ii) For any $t \geq 0$, $h \mapsto W_H(t)h$ defines a Hilbert-Schmidt operator from H to $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

(iii) $\mathsf{Tr}\, Q < \infty$.

Nevertheless, there always exist a larger Hilbert space H' and a Hilbert-Schmidt embedding operator $J : H \hookrightarrow H'$, such that the series (3.3) converges in $L^2((\Omega, \mathcal{F}, \mathbb{P}), H')$, see [DPZ92, Section 4.1.2] or [LR15, Section 2.5] for more details. In particular, this means that if $B \in L(H', H)$ is Hilbert Schmidt, then $(BW_H(t))_{t\geq 0}$ defines a H-valued $Q = BB^*$-Wiener process.

Remark 3.17.

(i) To define Wiener processes on the whole of $\mathbb{R}_+$ we simply concatenate independent copies of Wiener processes on intervals $[0, T]$.

(ii) Assume that $(W_1(t))_{t\geq 0}$ and $(W_2(t))_{t\geq 0}$ are two independent Q-Wiener processes. Then

$$W(t) = \begin{cases} W_1(t) & , t \geq 0 \\ W_2(-t) & , t < 0 \end{cases}$$

is a *two-sided* Q-Wiener process $(W(t))_{t\in\mathbb{R}}$ that vanishes at zero.

(iii) In the following, we will often omit the subscripts for W_Q or W_H and simply denote a Q-Wiener or a cylindrical Wiener process as W. Furthermore, by referring to a cylindrical Wiener process in the following we always mean the standard cylindrical Wiener process.

3.1.2 Stochastic integration in Hilbert spaces

The stochastic integral with respect to a H-valued Q-Wiener process is defined as the natural generalization of the *Itô integral* with respect to a real-valued Brownian motion, as for example introduced in [Oks03]. In this subsection we will give a rough outline of the construction and we refer the reader to the monographs [LR15, GM10, DPZ92] for excellent, in-depth presentations of the topic.

As before, let $(H, \|\cdot\|_H)$ denote a separable Hilbert space and let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space.

Itô integral with respect to Q-Wiener processes

The construction of the *Itô integral* with respect to a H-valued Q-Wiener process $(W_Q(t))_{t\in[0,T]}$ (adapted to the normal filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$) begins with the definition for so-called *elementary processes*. Let $(K, \|\cdot\|_K)$ denote another separable Hilbert space. An elementary, $L(H,K)$-valued process is of the form

$$\Psi(t,\omega) = \psi(\omega)\mathbf{1}_0(t) + \sum_{j=0}^{n-1} \psi_j(\omega)\mathbf{1}_{(t_j,t_{j+1}]}(t),$$

with $0 = t_1 < ... < t_n = T$, $n \in \mathbb{N}$, where ψ, ψ_j, $j = 0, ..., n-1$, are $\mathcal{F}_0$-, $\mathcal{F}_{t_j}$-measurable $L(H,K)$-valued random variables that take only a finite number of values and $\mathbf{1}_A(t)$ denotes the indicator function on a subset $A \subset \mathbb{R}^+$. Let $\mathcal{E}(L(H,K))$ denote the set of all $L(H,K)$-valued elementary processes. For $\Psi \in \mathcal{E}(L(H,K))$ the Itô integral is defined as

$$\begin{aligned}\mathrm{Int}(\Psi)(t) &:= \int_0^t \Psi(s)\mathrm{d}W_Q(s) \\ &= \sum_{j=0}^{n-1} \psi_j \left(W_Q\left(\min\{t, t_{j+1}\}\right) - W_Q\left(\min\{t, t_j\}\right)\right), \quad \text{for } t \in [0,T].\end{aligned}$$

$\mathrm{Int}(\Psi)(t)$ is a continuous, square-integrable martingale with respect to the filtration $\mathcal{F}$ (see [LR15, Proposition 2.3.2]). We denote by $\mathcal{M}_T^2(K)$ the space of K-valued *continuous, square integrable martingales* $M(t)$, $t \in [0,T]$, which, equipped with the norm

$$\|M\|_{\mathcal{M}_T^2(K)} := \sup_{t\in[0,T]} \left(\mathbb{E}\left[\|M(t)\|_K^2\right]\right)^{1/2} = \left(\mathbb{E}\left[\|M(T)\|_K^2\right]\right)^{1/2}, \tag{3.4}$$

becomes a Banach space [LR15, Proposition 2.2.9]. Note that the second equality in (3.4) follows since $\|M(t)\|_K^2$ is a sub-martingale. Furthermore, calculating the $\mathcal{M}_T^2(K)$ norm of $\mathrm{Int}(\Psi)$ yields the so-called *Itô isometry* [LR15, Proposition 2.3.5]

$$\|\mathrm{Int}(\Psi)\|_{\mathcal{M}_T^2(K)}^2 = \mathbb{E}\left[\int_0^T \|\Psi(s)Q^{1/2}\|_{L_2(H,K)}^2 \mathrm{d}s\right] =: \|\Psi\|_T^2, \tag{3.5}$$

where we recall that $(L_2(H,K), \|\cdot\|_{L_2(H,K)})$ denotes the space of all Hilbert-Schmidt operators from H to K. Note that $\|\cdot\|_T$ is only a semi-norm on $\mathcal{E}(L(H,K))$ and in order to obtain a norm we switch to equivalence classes of elementary processes without changing the notation. Then,

$$\mathrm{Int} : (\mathcal{E}(L(H,K)), \|\cdot\|_T) \to (\mathcal{M}_T^2(K), \|\cdot\|_{\mathcal{M}_T^2(K)}),$$

is an isometric transformation.

In a second step the *abstract completion* $\overline{\mathcal{E}}(L(H,K))$ of $\mathcal{E}(L(H,K))$ is characterised. More precisely, $\overline{\mathcal{E}}(L(H,K))$ can be explicitly represented as

$$\mathcal{N}_W^2(0,T;K) := \{\Psi : [0,T] \times \Omega \to L_2(H_0,K) : \Psi \text{ is predictable and } \|\Psi\|_T < \infty\}.$$

Here, H_0 denotes the separable Hilbert space $H_0 = Q^{1/2}(H)$ with scalar product

$$\langle u_0, v_0 \rangle_{H_0} := \langle Q^{-1/2}u_0, Q^{-1/2}v_0 \rangle_H,$$

see [LR15, Chapter 2] for details. Note that with this definition the $\|\cdot\|_T$-norm can be written as

$$\|\Psi\|_T^2 = \mathbb{E}\left[\int_0^T \|\Psi(s)\|_{L_2(H_0,K)}^2 \mathrm{d}s\right].$$

Furthermore, *predictable* means here that Ψ is $\mathcal{P}_T/\mathcal{B}(L_2(H_0,K))$ measurable where $\mathcal{P}_T$ is defined as

$$\mathcal{P}_T := \sigma\left(\{(s,t] \times F_s : 0 \le s < t \le T, F_s \in \mathcal{F}_s\} \cup \{\{0\} \times F_0 : F_0 \in \mathcal{F}_0\}\right),$$

where $\sigma(S)$ denotes the σ-algebra generated by the family of sets S. Now, as $\mathcal{E}(L(H,K))$ is dense in $\overline{\mathcal{E}}(L(H,K))$, there is a unique and isometric extension of the Itô integral to $\mathcal{N}_W^2(0,T;H)$. That is,

$$\mathrm{Int} : (\mathcal{N}_W^2(0,T;H), \|\cdot\|_T) \to (\mathcal{M}_T^2(K), \|\cdot\|_{\mathcal{M}_T^2(K)}).$$

is linear and isometric. Furthermore, we have for $\Psi \in \mathcal{N}_W^2(0,T;H)$ that

$$\mathbb{E}[\mathrm{Int}(\Psi)(t)] = 0, \quad t \in [0,T].$$

Finally, by a *localisation procedure* the definition of the Itô integral can be further extended to the set

$$\mathcal{N}_W(0,T;K) := \{\Psi : [0,T] \times \Omega \to L_2(H_0,K) : \Psi \text{ is predictable with } \mathbb{P}\left(\int_0^T \|\Psi(s)\|_{L_2(H_0,K)}^2 \mathrm{d}s < \infty\right) = 1\right\},$$

which is called the set of *stochastically integrable processes*. Note that for $\Psi \in \mathcal{N}_W(0,T;H)$ the integral $\mathrm{Int}(\Psi)$ defined by this procedure is only a continuous K-valued *local martingale*.

Itô integral with respect to cylindrical Wiener processes

Let $W_H := \{W_H(t)h : t \in [0,T], h \in H\}$ be a standard cylindrical Wiener process on H. According to Subsection 3.1.1 there exists another Hilbert space H' and a Hilbert-Schmidt embedding $H \hookrightarrow H'$ such that

$$W_H(t) = \sum_{k=1}^{\infty} \beta_k(t) e_k, \tag{3.6}$$

where $(\beta_k(t))_{k\geq 1}$ are independent Brownian motions, $(e_k)_{k\geq 1}$ is an orthonormal basis of H and the series converges in $L^2((\Omega, \mathcal{F}, \mathbb{P}), H')$. Let J be the corresponding Hilbert-Schmidt embedding operator. In particular, (3.6) is a H'-valued $Q_1 := JJ^*$-Wiener process. Now, for predictable $\Psi : [0,T] \times \Omega \to L_2(H,K)$ with

$$\mathbb{P}\left(\int_0^T \|\Psi(s)\|^2_{L_2(H,K)} \mathrm{d}s < \infty \right) = 1,$$

the stochastic integral with respect to W_H is defined as

$$\int_0^t \Psi(s) \mathrm{d}W_H(s) := \int_0^t \Psi(s) \circ J^{-1} \mathrm{d}W_H(s), \qquad t \in [0,T]. \tag{3.7}$$

In particular, since

$$\Psi \in L_2(H,K) \Leftrightarrow \Psi \circ J^{-1} \in L_2(Q_1^{1/2}(H'), K),$$

the right hand side of (3.7) is well defined as described above, see [LR15, Section 2.5] for further details. Consequently, the Itô isometry holds, that is,

$$\mathbb{E}\left[\left\| \int_0^t \Psi(s) \ \mathrm{d}W_H(s) \right\|_K^2 \right] = \mathbb{E}\left[\int_0^t \|\Psi(s)\|^2_{L_2(H,K)} \ \mathrm{d}s \right], \qquad \text{for } t \in [0,T]. \tag{3.8}$$

Itô formula

An indispensable tool for stochastic calculus in infinite dimension is the *Itô formula*, which can be regarded as the stochastic analogon to the *fundamental theorem of calculus.*

Definition 3.18 (Itô process)**.** Let $(W_Q(t))_t$ be a H-valued Q-Wiener process and $\Psi \in \mathcal{N}_W(0,T;K)$ a stochastically integrable process. Furthermore, let $\Phi : \Omega \times [0,T] \to K$ be $\mathbb{P}$-a.s. Bochner integrable and $\mathcal{F}$-measurable. Finally, let $X(0)$ be a $\mathcal{F}_0$-measurable K-valued random variable. Then

$$X(t) = X(0) + \int_0^t \Phi(s) \mathrm{d}s + \int_0^t \Psi(s) \mathrm{d}W_Q(s), \quad t \in [0,T],$$

is well-defined and called *Itô process.*

Theorem 3.19 (Itô formula, see [DPZ92, Theorem 4.32])**.** *Let $X : \Omega \times [0,T] \to K$ be an Itô process and $F : [0,T] \times K \to \mathbb{R}$ continuous. Furthermore, let the Fréchet derivatives F_t, F_x, F_{xx} be continuous and bounded on bounded subsets of $[0,T] \times K$. Then the following Itô formula holds $\mathbb{P}$-a.s. for all $t \in [0,T]$*

$$\begin{aligned} F(t, X(t)) = F(0, X(0)) &+ \int_0^t \langle F_x(s, X(s)), \Psi(s) \mathrm{d}W_Q(s) \rangle_K \\ &+ \int_0^t [F_t(s, X(s)) + \langle F_x(s, X(s)), \Phi(s) \rangle \\ &\quad + \frac{1}{2} \mathit{Tr} \left[F_{xx}(s, X(s))(\Psi(s)Q^{1/2})(\Psi(s)Q^{1/2})^* \right] \Big] \, \mathrm{d}s. \end{aligned}$$

Remark 3.20. A similar Itô formula can be proved for cylindrical Wiener processes, see [GM10].

Stochastic Fubini theorem

Let again $(H, \|\cdot\|_H)$ and $(K, \|\cdot\|_K)$ be separable Hilbert spaces.

Theorem 3.21 (Stochastic Fubini theorem, [DPZ92, Theorem 4.33])**.** *Let $(E, \mathcal{E})$ be a measurable space and μ be a finite positive measure on this space. Let $(W_Q(t))_t$ be a H-valued Q-Wiener process on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$. Furthermore, assume that*

$$\begin{aligned} \Phi : ([0,T] \times \Omega \times E, \mathcal{P}_T \times \mathcal{E}) &\to (L_2(H_0, K), \mathcal{B}(L_2(H_0, K))), \\ (t, \omega, x) &\mapsto \Phi(t, \omega, x), \end{aligned}$$

is a measurable mapping. Here, we have again defined $H_0 := Q^{1/2}H$. If

$$\int_E \|\Phi(\cdot, \cdot, x)\|_T \mu(\mathrm{d}x) < \infty,$$

then $\mathbb{P}$-a.s.

$$\int_E \left[\int_0^T \Phi(t, x) \; \mathrm{d}W_Q(t) \right] \; \mu(\mathrm{d}x) = \int_0^T \left[\int_E \Phi(t, x) \; \mu(\mathrm{d}x) \right] \; \mathrm{d}W_Q(t).$$

Stratonovich integral

We recall that in the SODE setting the definition of the stochastic integral of a suitable integrand f with respect to a Brownian motion depends on the choice of the approximation via elementary functions. More precisely, when choosing $f(t_j)$ as the value for the interval $[t_j, t_{j+1}]$, the construction leads to the *Itô integral*; whereas when choosing the midpoint, i.e. $f((t_j + t_{j+1})/2)$, this leads to the *Stratonovich integral*, see [Oks03] for details. Compared to Itô integrals, Stratonovich integrals are not martingales, however, they have the advantage that they obey the ordinary chain rule and no second order terms appear as in the Itô formula.

The same holds true when considering Stratonovich integrals of Hilbert space valued integrands. That is, let H be a separable Hilbert space, $(B(t))_{t\geq 0}$ be a real-valued Wiener process and let $\Phi(t,\cdot) : H \to H$ be a Fréchet differentiable mapping and $(u(t))_{t\geq 0}$ be some H-valued process. Then the transformation between the Stratonovich integral (where the differential is denoted as $\circ\, \mathrm{d}B(s)$) and the Itô integral reads as follows

$$\int_0^t \Phi(s,u(s)) \circ \mathrm{d}B(s) = \int_0^t \Phi(s,u(s))\ \mathrm{d}B(s) + \frac{1}{2}\int_0^t \Phi(s,u(s))\Phi_u(s,u(s))\ \mathrm{d}s, \quad (3.9)$$

where Φ_u is the Fréchet derivative of Φ with respect to u, see [DW14, Section 4.5]. Hence, by adding or subtracting a *correction term* one can transform a Stratonovich integral into an Itô integral and vice versa. In particular, if the integrand does not depend on u, we observe that the correction term vanishes. Thus, for stochastic differential equations with additive noise terms there is no difference in interpreting the stochastic differential in the Stratonovich or Itô sense.

3.2 Semi-linear stochastic partial differential equations

Let us fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We are interested in semi-linear Itô SPDEs. More precisely, let $(H, \|\cdot\|_H)$ and $(K, \|\cdot\|_K)$ be two separable Hilbert spaces. Then, we will consider equations of the form

$$\begin{aligned} \mathrm{d}u(t) &= [Au(t) + F(t,u(t))]\ \mathrm{d}t + B(t,u(t))\ \mathrm{d}W(t), \quad t \in [0,T], \\ u(0) &= u_0, \end{aligned} \quad (3.10)$$

under the following assumptions

Assumptions 3.22.

(i) $A : \mathcal{D}(A) \subset H \to H$ is the generator of a C_0-semigroup $(S(t))_{t\geq 0}$ of linear operators on H (see Appendix C.3),

(ii) W is a standard cylindrical Wiener process on K,

(iii) $F : [0,T] \times H \to H$ is $\mathcal{B}([0,T]) \otimes \mathcal{B}(H)/\mathcal{B}(H)$-measurable,

(iv) $B : [0,T] \times H \to L_2(K,H)$ is $\mathcal{B}([0,T]) \otimes \mathcal{B}(H)/\mathcal{B}(L_2(K,H))$ measurable,

(v) u_0 is H-valued and $\mathcal{F}_0$-measurable, where $(\mathcal{F}_t)_{t\geq 0}$ is the filtration generated by the Wiener process.

3.2.1 Solution concepts

There are several different notions of solutions to equation (3.10), see for example [Tap13] for an overview.

Definition 3.23 (Strong solution)**.** A $\mathcal{D}(A)$-valued predictable process $u(t)$, $t \in [0,T]$, is called *strong solution* of equation (3.10) if for all $t \in [0,T]$

$$u(t) = u_0 + \int_0^t [Au(s) + F(s,u(s))] \, \mathrm{d}s + \int_0^t B(s,u(s)) \, \mathrm{d}W(s), \quad \mathbb{P}-\text{a.s.}$$

In particular, the integrals on the right hand side have to be well-defined, that is

$$\mathbb{P}\left(\int_0^T \|u(s)\|_H + \|Au(s)\|_H + \|F(s,u(s))\|_H + \|B(s,u(s))\|^2_{L_2(K,H)} \, \mathrm{d}s < \infty\right) = 1.$$

Definition 3.24 (Weak solution)**.** A H-valued predictable process $u(t)$, $t \in [0,T]$, is called *weak solution* of equation (3.10) if for every test function $\xi \in \mathcal{D}(A^*)$ and every $t \in [0,T]$

$$\begin{aligned}\langle u(t), \xi\rangle = \langle u_0, \xi\rangle &+ \int_0^t \langle u(s), A^*\xi\rangle + \langle F(s,u(s)), \xi\rangle \, \mathrm{d}s \\ &+ \int_0^t \langle B(s,u(s)), \xi\rangle \, \mathrm{d}W(s), \quad \mathbb{P}-\text{a.s.}\end{aligned}$$

Note that $(A^*, \mathcal{D}(A^*))$ denotes the adjoint of $(A, \mathcal{D}(A))$ on H and H is identified with its dual. Again, it is assumed that

$$\mathbb{P}\left(\int_0^T \|u(s)\|_H + \|F(s,u(s))\|_H + \|B(s,u(s))\|^2_{L_2(K,H)} \, \mathrm{d}s < \infty\right) = 1.$$

Definition 3.25 (Mild solution)**.** A H-valued predictable process $u(t)$, $t \in [0,T]$, is called *mild solution* of equation (3.10) if for every $t \in [0,T]$ *Duhamel's formula* holds, i.e. $\mathbb{P}-$a.s.

$$u(t) = S(t)u_0 + \int_0^t S(t-s)F(s,u(s)) \, \mathrm{d}s + \int_0^t S(t-s)B(s,u(s)) \, \mathrm{d}W(s),$$

and

$$\mathbb{P}\left(\int_0^T \|u(s)\|_H + \|F(s,u(s))\|_H + \|B(s,u(s))\|^2_{L_2(K,H)} \, \mathrm{d}s < \infty\right) = 1.$$

Definition 3.26 (Stochastic convolution)**.** The stochastic integral in the mild formulation is called *stochastic convolution.*

Every strong solution is also a weak solution. Furthermore, we have the following relation between weak and mild solutions.

Lemma 3.27 (see [Tap13, Prop. 5.9, 5.11])**.** *Every weak solution to* (3.10) *is a mild solution to* (3.10)*. Every mild solution u of* (3.10) *for which*

$$\mathbb{E}\left[\int_0^T \|B(s,u(s))\|^2_{L_2(K,H)}\ \mathrm{d}s\right] < \infty,$$

is also a weak solution to (3.10)*.*

Remark 3.28.

(i) For more details on how the different concepts are related we refer to [Tap13, DPZ92, GM10].

(ii) For non-linear operators A there exists also the widely used concept of *variational solutions*, see [LR15] for details.

3.2.2 Existence of mild solutions

Let us briefly state the main existence result of mild solutions to SPDEs of the form (3.10). We will need the following assumptions.

Assumptions 3.29. There exists a constant $C > 0$ such that for all $u, v \in H$, $t \in [0,T]$ and almost all $\omega \in \Omega$ it holds

(i)

$$\|F(t,\omega,u) - F(t,\omega,v)\|_H + \|B(t,\omega,u) - B(t,\omega,v)\|_{L_2(K,H)} \leq C\|u-v\|_H,$$

(ii)

$$\|F(t,\omega,u)\|^2_H + \|B(t,\omega,u)\|^2_{L_2(K,H)} \leq C^2(1 + \|u\|_H).$$

Theorem 3.30 (see [DPZ92, Thm. 7.2])**.** *Let Assumptions 3.22 and Assumptions 3.29 hold. Then problem* (3.10) *possesses a unique (up to equivalence) mild solution u. Moreover, u has a continuous modification.*

3.2.3 Stochastic convolution

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let H, K be Hilbert spaces, A be an operator that generates a C_0 semigroup $(S(t))_{t\geq 0}$ on H and let $B \in L_2(K,H)$. Furthermore, let $(W(t))_{t\geq 0}$ denote a standard cylindrical Wiener process on K, adapted to the normal filtration $(\mathcal{F}_t)_{t\geq 0}$. The solution to the following linear equation is called *Ornstein-Uhlenbeck process*

$$\mathrm{d}u = Au\ \mathrm{d}t + B\ \mathrm{d}W, \tag{3.11}$$

and with initial condition $u(0) = 0$ it is simply given by the *stochastic convolution*

$$u(t) = \int_0^t S(t-s)B\ \mathrm{d}W(s), \qquad t \in [0,T]. \tag{3.12}$$

Note that a useful tool to work with the stochastic convolution is given by the following *integration by parts formula*

$$\begin{aligned} u(t) &= \int_0^t S(t-s)B \, \mathrm{d}W(s) \\ &= BW(t) + A\int_0^t S(t-s)BW(s) \, \mathrm{d}s \\ &= -A\int_0^t S(t-s)B(W(t)-W(s)) \, \mathrm{d}s + S(t)BW(t). \end{aligned}$$

Space and time regularity of Ornstein-Uhlenbeck processes on Hilbert spaces are closely linked. More precisely, interpreted as a evolution process on an infinite-dimensional space, the time regularity depends on the functional space under consideration. In the following we state some well-known regularity results for stochastic convolutions.

Proposition 3.31 (see [DP12, Prop. 2.2 & 2.3])**.** *In the setting described above, $u(t)$ is a Gaussian random variable with mean zero and covariance operator Q_t, where*

$$Q_t x := \int_0^t S(s)BB^*S^*(s)x \, \mathrm{d}s, \quad \textit{for } x \in H, \tag{3.13}$$

where $(S^*(t))_{t\geq 0}$ is the semigroup generated by A^*. Furthermore, the process $u(\cdot)$ is adapted and mean square continuous, i.e.

$$\sup_{t\in[0,T]} \mathbb{E}(\|u(t)\|_H^2) < \infty.$$

One can actually show a stronger result, namely that $u(\cdot)$ is $\mathbb{P}$-almost surely continuous.

Proposition 3.32 (cf. [DPZ92, Thm 5.11])**.** *In the setting described above, u has almost surely continuous sample-paths in H and for $p > 0$*

$$\mathbb{E}\left(\sup_{t\in[0,1]} \|u(t)\|_H^p\right) < \infty. \tag{3.14}$$

In case A generates an *analytic semigroup* stronger regularity results can be proved. In particular, sample paths are Hölder continuous, as stated in the following proposition.

Proposition 3.33 (see [DPZ92, Theorem 5.15])**.** *Assume that, in the setting above, the semigroup $(S(t))_{t\geq 0}$ is analytic and let $\alpha \in (0, 1/2)$. Then for arbitrary $\delta \in (0, \alpha)$ the sample paths of u are in $C^\delta([0,T]; H)$.*

Moreover, we will need the following result later on in Chapter 5, cf. [DPZ92, Theorem 5.25].

Proposition 3.34. *Consider equation* (3.11) *with* $H = L^2(D)$, *where* D *is bounded. Assume that the spectrum of* A *lies entirely in the open left half plane and that* A *generates an analytic semigroup* $(S(t))_{t\geq 0} = (\exp(tA))_{t\geq 0}$ *on* $L^p(D)$, *for any* $p \geq 1$, *and that* $D((-A)^\gamma)$ *can be identified with the Sobolev space* $W^{2\gamma,p}(D)$, *for any* $\gamma > 0$. *Furthermore, let the eigenfunctions* $\{e_k\}_{k=1}^\infty$ *of* A *form a complete orthonormal system of* H *with* $|e_k(x)|^2 < \kappa$ *for some* $\kappa > 0$ *and for all* k *and* $x \in D$. *Let* $B \in L(K, H)$ *and assume that the operator* Q_t *defined in* (3.13) *is of trace class. Set* $Q = BB^*$. *Furthermore, let* $K = H$ *and assume that there exist sequences* $\{\lambda_k\}_{k=1}^\infty$ *and* $\{\delta_k\}_{k=1}^\infty$ *such that*

$$Ae_k = -\lambda_k e_k, \qquad \text{and} \qquad Qe_k = \delta_k e_k, \quad k \in \mathbb{N}.$$

Finally, we assume that there exists $\alpha \in \left(0, \frac{1}{2}\right)$ *such that*

$$\sum_{k=1}^\infty \delta_k \lambda_k^{2\alpha+1} < \infty. \tag{3.15}$$

Then

$$\mathbb{E}\left(\sup_{t\in[0,1]} \|u(t)\|_p^p\right) < \infty.$$

Proof. For completeness we give the proof here. Note that using the well-known *factorization formula* (see [DPZ92, Theorem 5.10]), we have for $(t, x) \in [0, T] \times D$ and $\alpha \in (0, 1/2)$

$$u(t) = \frac{\sin(\pi\alpha)}{\pi} \int_0^t \exp\left((t-\tau)A\right)(t-\tau)^{\alpha-1} y(\tau)\, \mathrm{d}\tau,$$

with

$$\begin{aligned} y(\tau, x) &= \int_0^\tau \exp\left((\tau - s)A\right)(\tau - s)^{-\alpha} B\, \mathrm{d}W(s, x) \\ &= \sum_{k=1}^\infty \int_0^\tau \exp\left((\tau - s)A\right)(\tau - s)^{-\alpha} B e_k(x) d\beta_k(s) \\ &= \sum_{k=1}^\infty \int_0^\tau \exp\left(-(\tau - s)\lambda_k\right)(\tau - s)^{-\alpha} \sqrt{\delta_k} e_k(x) d\beta_k(s), \end{aligned}$$

where we have used the formal representation $W(s, x) = \sum_{k=1}^\infty \beta_k(s) e_k(x)$ of the cylindrical Wiener process, with $\{\beta_k\}_{k=1}^\infty$ being a sequence of mutually independent real-valued Brownian motions. $y(\tau, x)$ is a real-valued Gaussian random variable with mean zero and variance

$$\begin{aligned}
&\mathbb{E}\left[|y(\tau,x)|^2\right] \\
&= \mathbb{E}\left[\sum_{k=1}^{\infty}\left(\int_0^{\tau} \exp\left(-(\tau-s)\lambda_k\right)(\tau-s)^{-\alpha}\sqrt{\delta_k}\ \mathrm{d}\beta_k(s)\right)^2 |e_k(x)|^2\right] \\
&= \sum_{k=1}^{\infty}\delta_k|e_k(x)|^2\mathbb{E}\left[\left(\int_0^{\tau} \exp\left(-(\tau-s)\lambda_k\right)(\tau-s)^{-\alpha}\ \mathrm{d}\beta_k(s)\right)^2\right] \\
&= \sum_{k=1}^{\infty}\delta_k|e_k(x)|^2\int_0^{\tau}\exp\left(-2s\lambda_k\right)s^{-2\alpha}\ \mathrm{d}s,
\end{aligned}$$

where we have used Parseval's identity and the Itô isometry. Our assumption on the boundedness of the eigenfunctions $\{e_k\}_{k=1}^{\infty}$ yields together with (3.15) that

$$\begin{aligned}
\mathbb{E}\left[|y(\tau,x)|^2\right] &< \sum_{k=1}^{\infty}\delta_k\kappa^2\int_0^{\infty}\exp\left(-2s\lambda_k\right)s^{-2\alpha}\ \mathrm{d}s \\
&= \kappa^2 2^{2\alpha-1}\Gamma(1-2\alpha)\sum_{k=1}^{\infty}\delta_k\lambda_k^{2\alpha-1} < \infty.
\end{aligned}$$

Hence, $\mathbb{E}\left[|y(\tau,x)|^{2m}\right] \leq C_m$ for $C_m > 0$ and every $m \geq 1$ (note that all odd moments of a Gaussian random variable are zero). Thus we have

$$\mathbb{E}\left[\int_0^T\int_D |y(\tau,x)|^{2m}\mathrm{d}x\mathrm{d}\tau\right] \leq TC_m|D|,$$

i.e., in particular for all $p \geq 1$ we have $y \in L^p([0,T]\times D)$ $\mathbb{P}$-a.s.. Now, since A generates an analytic semigroup and its spectrum lies entirely in the open left half plane and by the assumed identification of the interpolation spaces with Sobolev spaces, we can conclude using Lemma C.21 that

$$\|\exp(tA)x\|_{W^{\gamma,p}} \leq Ct^{-\gamma/2}e^{-\delta t}\|x\|_p \leq Ct^{-\gamma/2}\|x\|_p, \qquad \text{for } x \in L^p(D).$$

Choosing $\gamma = \alpha$ we observe

$$\begin{aligned}
&\|u(t)\|_{W^{\gamma,p}(D)} \\
&\leq \frac{\sin(\pi\alpha)}{\pi}\int_0^t (t-\tau)^{\alpha-1}\|\exp\left((t-\tau)A\right)y(\tau,\cdot)\|_{W^{\gamma,p}(D)}\ \mathrm{d}\tau \\
&\leq C\frac{\sin(\pi\alpha)}{\pi}\int_0^t (t-\tau)^{\alpha-1}(t-\tau)^{-\alpha/2}\|y(\tau,\cdot)\|_p\ \mathrm{d}\tau \\
&\leq C\sup_{\tau\in[0,t]}\|y(\tau,\cdot)\|_p\int_0^t (t-\tau)^{\alpha/2-1}\ \mathrm{d}\tau,
\end{aligned}$$

and thus

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\in[0,1]}\|u(t)\|_p\right)\\
&\le C\ \mathbb{E}\left(\sup_{t\in[0,1]}\sup_{\tau\in[0,t]}\|y(\tau,\cdot)\|_p\right)\int_0^1 \tau^{\alpha/2-1}\ \mathrm{d}\tau\\
&= C\ \mathbb{E}\left(\sup_{\tau\in[0,1]}\|y(\tau,\cdot)\|_p\right).
\end{aligned}$$

Now, the right hand side is finite as all moments of $y(\tau,x)$ are bounded uniformly in x,τ, see above. Due to embedding of Lebesgue spaces on a bounded domain we have that

$$\mathbb{E}\left(\sup_{t\in[0,1]}\|u(t)\|_p\right)<\infty \quad \text{implies} \quad \mathbb{E}\left(\sup_{t\in[0,1]}\|u(t)\|_p^p\right)<\infty.$$

□

Remark 3.35. Alternatively, one can analyse the regularity of the stochastic convolution (3.12) using the integration by parts formula as stated above.

3.3 Galerkin approximation of SPDEs

We refer to [LPS14] for a comprehensive presentation of numerical methods for solving SPDEs. Here, we only recall the *spectral Galerkin approximation*, as this will play a role later on.

Let us consider the special case of a semi-linear evolution equation on a Hilbert space H with additive noise of the form

$$\begin{aligned}
\mathrm{d}u(t) &= [Au(t)+F(u(t))]\ \mathrm{d}t+\sigma\ \mathrm{d}W(t), \quad t\in[0,T],\\
u(0) &= u_0,
\end{aligned} \tag{3.16}$$

where $\sigma>0$ and

(i) $A:\mathcal{D}(A)\subset H\to H$ has a complete orthonormal set of eigenfunctions $(\phi_j)_{j\in\mathbb{N}}$ and corresponding eigenvalues $\lambda_j>0$, where $\lambda_{j+1}\ge\lambda_j$,

(ii) W is a H-valued Q-Wiener process, where the eigenfunctions of the covariance operator are given by $(\phi_j)_{j\in\mathbb{N}}$,

(iii) $F:H\to H$ is Lipschitz continuous,

(iv) $u_0\in H$.

For the *spectral Galerkin approximation* of a SPDE we define a finite-dimensional subspace $V_m := \text{span}\{\phi_1, ..., \phi_m\}$ that is spanned by the first m eigenfunctions of A. Furthermore, we define the *orthogonal projection* $P_m : H \to V_m$. In particular, for $u \in H$ [LPS14, Lemma 1.41]

$$P_m u = \sum_{j=1}^{m} \hat{u}_j \phi_j, \quad \hat{u}_j := \langle u, \phi_j \rangle,$$

and

$$\|P_m u\|_H \leq \|u\|_H, \quad \|u - P_m u\|_H \to 0, \text{ as } m \to \infty.$$

Then, we seek the so-called Galerkin approximation $u_m(t) \in V_m$ that satisfies the equation

$$\mathrm{d}u_m(t) = [A_m u_m(t) + P_m F(u_m)]\, \mathrm{d}t + \sigma P_m\, \mathrm{d}W(t), \tag{3.17}$$

with $A_m := P_m A$ and with initial data $u_m^0 := P_m u_0$.

Now, (3.17) is an SDE on a finite-dimensional space and one can apply for example the *Euler-Maruyama method* with a stepsize Δt to find an approximation $u_{m,n}$ of $u_m(t_n)$ with $t_n := n\Delta t$. That is starting with u_m^0 we consider the iteration

$$u_{m,n+1} = u_{m,n} - \Delta t A_m u_{m,n+1} + \Delta t P_m F(u_{m,n}) + \sigma P_m \Delta W_n,$$

with $\Delta W_n := \int_{t_n}^{t_{n+1}} \mathrm{d}W(s)$.

Remark 3.36. In more general settings with multiplicative noise and where the eigenfunctions of Q and A do not coincide the approximation becomes more involved, see [LPS14, Section 10.6].

Chapter 4

Random dynamics

4.1 Random dynamical systems

In this section we describe the theory of *random dynamical systems*; for a comprehensive reference book on this topic we refer to [Arn13].

4.1.1 Metric dynamical systems

We begin with the concept of a metric dynamical system, which serves as a generalized model of noise and its time-evolution.

Definition 4.1 (Metric dynamical system (MDS))**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\theta = \{\theta_t : \Omega \to \Omega\}_{t \in \mathbb{R}}$ be a family of $\mathbb{P}$-preserving transformations (meaning that $\theta_t \mathbb{P} = \mathbb{P}$ for all $t \in \mathbb{R}$), which satisfy for all $t, s \in \mathbb{R}$ that

(i) $(t, \omega) \mapsto \theta_t \omega$ is $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}, \mathcal{F})$-measurable,

(ii) $\theta_0 = \mathrm{Id}_\Omega$,

(iii) $\theta_{t+s} = \theta_t \theta_s = \theta_t \circ \theta_s$ (group property).

Then $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ is called a *metric dynamical system* (MDS).

A MDS is called *ergodic* if invariant sets $F \in \mathcal{F}$, that is sets satisfying $\theta_t F = F$ for all $t \in \mathbb{R}$, have either full or zero measure, i.e. $\mathbb{P}(F) \in \{0, 1\}$.

Example 4.2 (MDS associated to a Q-Wiener process)**.** *Consider the following canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$: For a separable Hilbert space H set $\Omega := C_0(\mathbb{R}; H)$; i.e. the sample space consists of all continuous functions on $\mathbb{R}$ with values in H that vanish at the origin. Furthermore, let $\mathcal{F} := \mathcal{B}(\Omega)$ be the Borel σ-algebra induced by the compact-open topology on Ω. As probability measure $\mathbb{P}$ on $\mathcal{F}$, we choose the Wiener measure induced by the trace-class covariance operator on H. Furthermore, let us introduce the Wiener shift on $C_0(\mathbb{R}; H)$, defined by*

$$\theta_t \omega(\cdot) := \omega(\cdot + t) - \omega(t). \tag{4.1}$$

For $A \in \mathcal{F}$, $s,t \in \mathbb{R}$ we verify

$$\begin{aligned}\theta_t\mathbb{P}(\{\omega : \omega(s) \in A\}) &= \mathbb{P}(\{\omega : \theta_t\omega(s) \in A\}) \\ &= \mathbb{P}(\{\omega : \omega(s+t) - \omega(t) \in A\}) = \mathbb{P}(\{\omega : \omega(s) \in A\}),\end{aligned}$$

i.e. θ_t is a $\mathbb{P}$-preserving transformation. Furthermore,

$$\theta_0\omega(s) = \omega(s) - \omega(0) = \omega(s),$$

and

$$\theta_{t+s}\omega(r) = \omega(r+t+s) - \omega(t+s) = \theta_t(\omega(r+s) - \omega(s)) = \theta_t(\theta_s\omega(r)),$$

for $t,s,r \in \mathbb{R}$. Thus the quadruple $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ defines a MDS. By using Kolmogorov's 0-1 law one can show that the MDS is ergodic.

For an ergodic MDS the following form of *Birkhoff's ergodic theorem* holds, see for instance [Wal00] for a proof.

Theorem 4.3 (Birkhoff's ergodic theorem)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be an ergodic MDS. Let $X : \Omega \to \mathbb{R}$ be an integrable random variable. Then there exists a θ-invariant set $\Omega' \in \mathcal{F}$ of full $\mathbb{P}$-measure such that for every $\omega \in \Omega'$*

$$\lim_{t\to\pm\infty} \frac{1}{t}\int_0^t X(\theta_s\omega)\ \mathrm{d}s = \mathbb{E}X.$$

4.1.2 Temperedness

A key concept in the theory of random dynamical systems is that of temperedness, which we are going to introduce now.

Definition 4.4 (Tempered random variable)**.** Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be a MDS. A random variable $X : \Omega \to [0,\infty)$ is called *tempered (from above)*, if there exists a set of full $\mathbb{P}$-measure Ω' such that for every $\omega \in \Omega'$ it holds

$$\lim_{t\to\pm\infty} \frac{\log^+ X(\theta_t\omega)}{|t|} = 0, \tag{4.2}$$

where $\log^+(x) = \max\{0, \log(x)\}$ for $x \geq 0$. A random variable $X : \Omega \to (0,\infty]$ is called *tempered from below* if X^{-1} is tempered.

Remark 4.5.

(i) Property (4.2) is equivalent to

$$\lim_{t\to\pm\infty} e^{-c|t|}X(\theta_t\omega) = 0, \quad \text{for any } c > 0. \tag{4.3}$$

(ii) If θ is ergodic, then the only alternative to (4.2) is

$$\limsup_{t\to\pm\infty} \frac{\log^+ X(\theta_t\omega)}{|t|} = \infty,$$

i.e., the stationary random process $X(\theta_t\omega)$ either grows sub-exponentially or blows up.

(iii) Note that for a tempered random variable the set Ω' of full $\mathbb{P}$-measure on which (4.2) holds is θ invariant.

We will also make use of the following sufficient condition for the temperedness of a positive random variable.

Proposition 4.6 (see [Arn13, Proposition 4.1.3])**.** *Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be a MDS and let $X : \Omega \to [0, \infty)$ be a random variable. If*

$$\mathbb{E}\left(\sup_{t\in[0,1]} X(\theta_t\omega)\right) < \infty,$$

then X is tempered.

Proof. For completeness we give the proof here. Let $\varepsilon > 0$ be arbitrary and define for $n \in \mathbb{N}$ the event

$$E_n := \left\{\sup_{n\le t\le n+1} \frac{\log^+ X(\theta_t\omega)}{t} > \varepsilon\right\}.$$

Then, by using the $\mathbb{P}$-invariance of the MDS, we compute

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}(E_n) &\le \sum_{n=1}^{\infty} \mathbb{P}\left(\sup_{n\le t\le n+1} \log^+ X(\theta_t\omega) > \varepsilon n\right) \\
&= \sum_{n=1}^{\infty} \theta_{-n}\mathbb{P}\left(\sup_{n\le t\le n+1} \log^+ X(\theta_t\omega) > \varepsilon n\right) \\
&\le \sum_{n=1}^{\infty} \mathbb{P}\left(\frac{1}{\varepsilon}\sup_{0\le t\le 1} \log^+ X(\theta_t\omega) > n\right) \\
&\le \int_0^\infty \mathbb{P}\left(\frac{1}{\varepsilon}\sup_{0\le t\le 1} \log^+ X(\theta_t\omega) > x\right)\,\mathrm{d}x \\
&\le \frac{1}{\varepsilon}\mathbb{E}\left(\sup_{0\le t\le 1} \log^+ X(\theta_t\omega)\right) \le \frac{1}{\varepsilon}\mathbb{E}\left(\sup_{0\le t\le 1} X(\theta_t\omega)\right) < \infty.
\end{aligned}$$

Hence, by Borel Cantelli for all $\varepsilon > 0$

$$\mathbb{P}\left(\frac{\log^+ X(\theta_t\omega)}{t} > \varepsilon \text{ i.o.}\right) = 0$$

and thus the statement for $t \to \infty$ follows. A similar argument holds for $t \to -\infty$. This concludes the proof. $\square$

We also note that products and sums of tempered random variables are tempered random variables as well. More precisely, it holds

Lemma 4.7 (see [Arn13, 4.1.2 Lemma]). *Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be a MDS. The set of tempered random variables $X : \Omega \to [0, \infty)$ is a commutative ring with unit element.*

We will use the Hausdorff semi-distance to measure the distance between two sets.

Definition 4.8 (Hausdorff semi-distance). Let (X, d) be a metric space, then the *Hausdorff semi-distance* between two non-empty subsets $\mathcal{A}, \mathcal{B} \in 2^X$ is given by

$$\operatorname{dist}(\mathcal{A}, \mathcal{B}) := \sup_{v \in \mathcal{A}} \inf_{\tilde{v} \in \mathcal{B}} d(v, \tilde{v})$$

and the distance between $v \in X$ and $\mathcal{B} \in 2^X$ is given by $\operatorname{dist}(v, \mathcal{B}) := \operatorname{dist}(\{v\}, \mathcal{B})$.

Let V be a separable Banach space with norm $\| \cdot \|$.

Definition 4.9 (Random set). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A set-valued map $\mathcal{K} : \Omega \to 2^V$ is called a *random set* on V if for all $v \in V$ the map

$$\omega \mapsto \operatorname{dist}(v, \mathcal{K}(\omega))$$

is measurable. A random set $\mathcal{K}$ is called closed or compact if for every $\omega \in \Omega$ the set $\mathcal{K}(\omega)$ is closed or compact.

Definition 4.10 (Tempered set). Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be a MDS. A bounded random set $\mathcal{A}$ is called tempered with respect to θ provided that for all $\omega \in \Omega$

$$\lim_{t \to \infty} \exp(-\beta t) \sup_{a \in \mathcal{A}(\theta_{-t}\omega)} \|a\| = 0 \quad \text{for all } \beta > 0.$$

Remark 4.11. Let $\mathcal{A}$ be a random set and let $\rho : \Omega \to (0, \infty)$ be a tempered random variable. If for all $\omega \in \Omega$

$$\mathcal{A}(\omega) \subset B(0, \rho(\omega)),$$

then $\mathcal{A}$ is a tempered set.

4.1.3 Random dynamical systems

Definition 4.12 (Random dynamical system (RDS)). A *random dynamical system* (RDS) on a separable Banach space V over a metric dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ is a map

$$\varphi : \mathbb{R}^+ \times \Omega \times V \to V; \quad (t, \omega, v) \mapsto \varphi(t, \omega, v),$$

which is $(\mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F} \otimes \mathcal{B}(V), \mathcal{B}(V))$-measurable, such that $\varphi(0, \omega) = \mathrm{Id}_V$ for all $\omega \in \Omega$ and

$$\varphi(t + s, \omega, v) = \varphi(t, \theta_s \omega, \varphi(s, \omega, v)),$$

for all $s, t \in \mathbb{R}^+$, $v \in V$ and $\omega \in \Omega$. This last condition is called *cocycle property*, see Figure 4.1.

We say that φ is a continuous or differentiable RDS if $v \mapsto \varphi(t, \omega, v)$ is continuous or differentiable for all $t \in \mathbb{R}^+$ and every $\omega \in \Omega$.

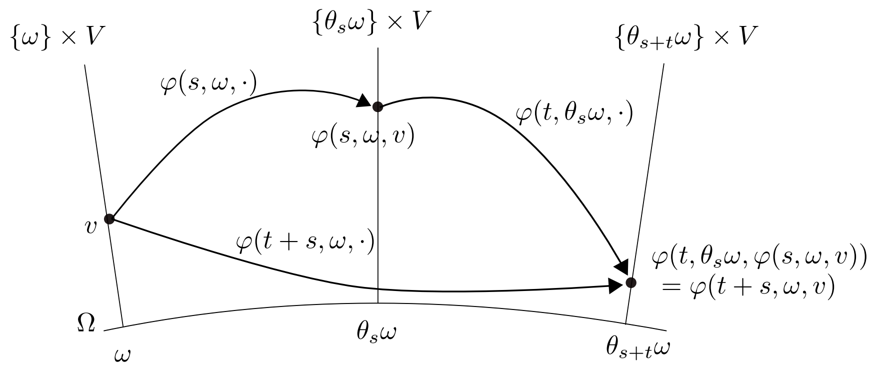

Figure 4.1: Illustration of the cocycle property, adapted from [Arn13, Fig. 1.2.].

Remark 4.13.

(i) The framework of random dynamical systems can be regarded as a generalization of the concept of non-autonomous deterministic dynamical systems, see [CH17] for further insight.

(ii) Definition 4.12 justifies the notion of the metric dynamical system as the driving force for the temporal evolution of the noise.

(iii) Finite-dimensional stochastic differential equations generate random dynamical system (see [Arn13, Chapter 1] and [Kun97, Section 4.5]). However, this does not hold in full generality for SPDEs since Kolmogorov's continuity theorem fails for random fields that are parametrized by an infinite-dimensional Hilbert space [MZZ06]. Nevertheless, in the special cases of linear multiplicative or additive noise it is possible to transform SPDEs into PDEs with random coefficients, which can be solved for every ω. More details on this approach can be found in 4.3 and 4.5.

(iv) Apart from the classical approach to generate a RDS from a SPDE via a transformation into a PDE with random coefficients as mentioned in (iii), there is also a possibility to use the concept of pathwise mild solutions as explored in [KNS20]. This allows to use the random dynamical systems approach for certain SPDEs that can not be handled via a Doss-Sussmann type transformation, e.g. because the differential operator depends on time and random coefficient ω.

A related notion is that of stochastic flows.

Definition 4.14 (Stochastic flow)**.** Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be a metric dynamical system and let V be a Banach space. A family of mappings

$$S(t, \tau, \omega) : V \to V \quad \text{for } t \geq \tau \in \mathbb{R},\ \omega \in \Omega,$$

is called a *stochastic flow* if the following conditions are fulfilled

(i) $S(t, s, \omega)S(s, \tau, \omega) = S(t, \tau, \omega)$ for all $\tau \leq s \leq t$ and $\omega \in \Omega$,

(ii) $S(t, \tau, \omega) = S(t - \tau, 0, \theta_\tau \omega)$ for all $\tau \leq t$ and $\omega \in \Omega$,

(iii) The map $(t, \tau, \omega, v) \mapsto S(t, \tau, \omega)v$ is $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{F} \otimes \mathcal{B}(V), \mathcal{B}(V))$-measurable.

If the map $v \mapsto S(t, \tau, v)$ is continuous for all $\tau \leq t$ and $\omega \in \Omega$, then S is called a *continuous stochastic flow.*

4.2 Random attractors

One of the main concepts to analyse the long-term behaviour of a random dynamical system is that of *random attractors.* A random attractor is defined as a compact invariant subset of the phase space with a certain pullback attraction property. This pullback approach is comparable to the concept of an attractor for a non-autonomous deterministic dynamical system, see Remark 4.13. We will present the precise mathematical formalism of random attractors in the following.

Random attractors have been introduced and intesively studied by Crauel and Flandoli [CF94], Debussche [Deb97, CDF97] and Schmalfuss [Sch92], amongst others.

Within this section we assume that φ is a random dynamical system on a separable Banach space $(V, \|\cdot\|)$ over a metric dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$.

Definition 4.15 (Attracting and absorbing set)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be random sets.

(i) $\mathcal{B}$ is said to *pullback attract* $\mathcal{A}$ for the RDS φ, if for every $\omega \in \Omega$

$$\operatorname{dist}(\varphi(t, \theta_{-t}\omega, \mathcal{A}(\theta_{-t}\omega)), \mathcal{B}(\omega)) \to 0 \quad \text{for } t \to \infty.$$

(ii) $\mathcal{B}$ is said to *absorb* $\mathcal{A}$ for the RDS φ, if for every $\omega \in \Omega$ there exists a (random) absorption time $t_\mathcal{A}(\omega)$ such that for all $t \geq t_\mathcal{A}(\omega)$

$$\varphi(t, \theta_{-t}\omega, \mathcal{A}(\theta_{-t}\omega)) \subset \mathcal{B}(\omega).$$

(iii) Let $\mathcal{D}$ be a collection of random sets, which is closed with respect to set inclusion. A set $\mathcal{B} \in \mathcal{D}$ is called $\mathcal{D}$-absorbing/$\mathcal{D}$-pullback attracting for the RDS φ, if $\mathcal{B}$ absorbs/pullback attracts every random set in $\mathcal{D}$.

Remark 4.16. Let $\mathcal{A}$ be a random set. If for every $v \in \mathcal{A}(\theta_{-t}\omega)$ and every $\omega \in \Omega$ it holds

$$\limsup_{t\to\infty} \|\varphi(t, \theta_{-t}\omega, v)\| \leq \rho(\omega), \tag{4.4}$$

where $\rho > 0$ is a tempered random variable, then for any $\delta > 0$ the random set $\mathcal{B}$ defined via

$$\mathcal{B}(\omega) := B(0, \rho(\omega) + \delta), \quad \text{for } \omega \in \Omega,$$

is a tempered absorbing set for $\mathcal{A}$. This is a convenient criterion to derive the existence of an absorbing set via a-priori estimates of the random dynamical system.

An attractor will be defined relative to a universe of sets that get attracted, the so-called *basin of attraction*. We choose the set of all tempered random subsets of V, denoted as $\mathcal{T}$, as the universe under consideration.

Definition 4.17 (Random pullback attractor)**.** A random set $\mathcal{A} \in \mathcal{T}$ is called a *$\mathcal{T}$-random pullback attractor* for the RDS φ if it possesses the following properties

(i) $\mathcal{A}(\omega)$ is compact for every $\omega \in \Omega$,

(ii) $\mathcal{A}$ is pullback invariant, i.e., $\varphi(t, \omega, \mathcal{A}(\omega)) = \mathcal{A}(\theta_t\omega)$ for all $t \geq 0$,

(iii) $\mathcal{A}$ is $\mathcal{T}$-pullback attracting.

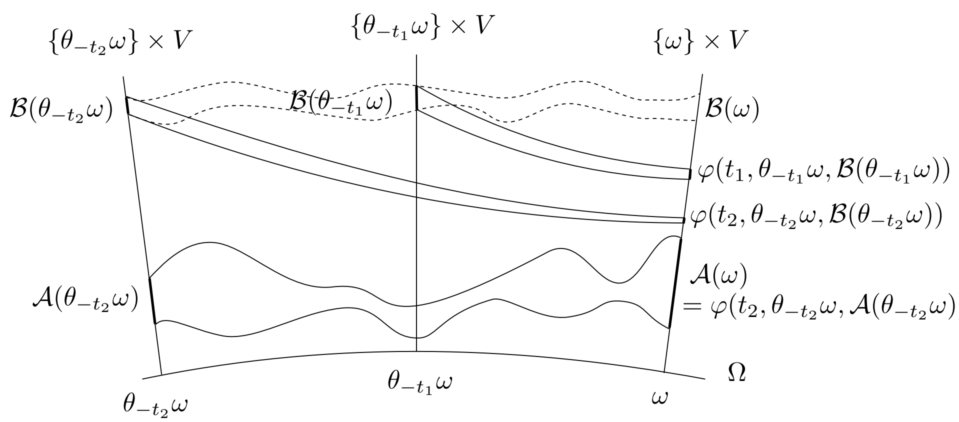

Figure 4.2: Illustration of the pullback invariance of the random attractor $\mathcal{A}$ and the pullback attraction of a set $\mathcal{B}$ by the attractor, adapted from [CH17, Fig. 4.1].

Both, the invariance and the attraction property, are defined in the pullback sense, i.e. states are moved from $-t$ to 0 while $t \to \infty$; see Figure 4.2 for an illustration. It is this pullback convergence that allows to analyse fixed fibres of the omega-limit sets, see the definition below.

Definition 4.18 (Omega-limit set)**.** For a random set $\mathcal{K}$ we define the *omega-limit set* as

$$\Omega_{\mathcal{K}}(\omega) := \bigcap_{T \geq 0} \overline{\bigcup_{t \geq T} \varphi(t, \theta_{-t}\omega, \mathcal{K}(\theta_{-t}\omega))}.$$

Note that $\Omega_{\mathcal{K}}(\omega)$ is closed by definition.

A very useful criterion for the existence of a random pullback attractor is given by the following theorem, a proof of which can be found in [FS96, Theorem 3.5]. The result is a generalisation of [AS96, CF94, Sch92].

Theorem 4.19. *Let φ be a continuous RDS over $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ and assume that there exists a compact random set $\mathcal{B} \in \mathcal{T}$ that absorbs every set $\mathcal{D} \in \mathcal{T}$, i.e. $\mathcal{B}$ is $\mathcal{T}$-absorbing. Then there exists a unique $\mathcal{T}$-random attractor $\mathcal{A}$, which is defined by*

$$\mathcal{A}(\omega) = \Omega_{\mathcal{B}}(\omega), \qquad \text{for each } \omega \in \Omega.$$

Remark 4.20. Note that if a RDS associated to a SPDE exhibits a random attractor, then there exists an invariant measure (see [CF94, Section 4]).

Remark 4.21. For many deterministic dynamical systems one can show that the dimension of the associated attractor is finite. In particular, this means that only a finite number of degrees of freedom is relevant for the long-term behaviour of the system. In [Deb98,Deb97] a method, based on global Lyapunov exponents, to deduce bounds on the dimensions of *random attractors* has been developed. However, so far only for few random systems it was possible to establish the finiteness of a random attractor [CLR01]. For lower bounds on the dimension of the attractor, one may also try to find invariant manifolds that lie within the attractor and whose dimension can be bounded.

4.3 Conjugacy

We introduce the notion of conjugated random dynamical systems. Again let $(V, \|\cdot\|)$ denote a separable Banach space.

Proposition 4.22 (see [CKS04, Lemma 2.2])**.** *Let φ_1 be a continuous RDS on V over a metric dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ and let $T : \Omega \times V \to V$ be a mapping with the following properties:*

(i) For fixed $\omega \in \Omega$ the mapping $v \mapsto T(\omega, v)$ is a homeomorphism on V.

(ii) For fixed $v \in V$ the mappings $\omega \mapsto T(\omega, v)$ and $\omega \mapsto T^{-1}(\omega, v)$ are measurable.

Then the mapping

$$(t, \omega, v) \mapsto \varphi_2(t, \omega, v) := T(\theta_t\omega, \varphi_1(t, \omega, T^{-1}(\omega, v)))$$

defines a RDS, which is called conjugate to the RDS φ_1.

There is a simple connection between random attractors of two conjugate RDS.

Theorem 4.23 (see [IS01, Theorem 2.1])**.** *Let φ_1 be a continuous RDS on V over the MDS $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ and let φ_2 be a RDS conjugated to φ_1 via the mapping T. Let $\mathcal{A}_1$ be a random attractor for the RDS φ_1. Furthermore, assume that $\{T(\mathcal{D})|\mathcal{D} \in \mathcal{T}\} \subset \mathcal{T}$, where $\mathcal{T}$ is the set of tempered subsets (i.e. T^{-1} preserves temperedness). Then $\mathcal{A}_2(\omega) := T(\omega, \mathcal{A}_1(\omega))$ is a random attractor for φ_2.*

Remark 4.24. A widely used strategy to show the existence of random attractors for stochastic partial differential equations is to find a suitable mapping that transforms the equation into a partial differential equation with random coefficients. If this PDE generates a random dynamical system φ_1 and if the mapping fulfils the conditions in Proposition 4.22, then the stochastic equation also generates a random dynamical system φ_2. In particular, it is sufficient to derive the existence of a random attractor for φ_1 as this implies the existence of a random attractor for the stochastic equation. These kinds of transformations are often based on suitable *Ornstein-Uhlenbeck processes*, which we are going to characterise further in the next section, see also Subsection 3.2.3. Details about the transformations can be found in Subsection 4.5.

4.4 Ornstein-Uhlenbeck processes

Recall that we have defined Ornstein-Uhlenbeck processes in a general setting in Subsection 3.2.3. Here we will look at two special cases, namely real-valued Ornstein-Uhlenbeck processes and Hilbert space valued Ornstein-Uhlenbeck processes associated to Q-Wiener processes, and show additional properties that are useful in the the theory of random dynamical systems.

4.4.1 Real-valued Ornstein-Uhlenbeck processes

We consider the ergodic MDS $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ associated to the two-sided real-valued Brownian motion $(B(t))_{t\in\mathbb{R}}$, see Example 4.2. The elements of Ω are identified with the paths of the Brownian motion, i.e. $W(t, \omega) = \omega(t)$ for $\omega \in \Omega$.

Let us consider the following equation

$$\mathrm{d}z = -z \; \mathrm{d}t + \mathrm{d}\omega. \tag{4.5}$$

The stationary solution to this equation is given by an Ornstein-Uhlenbeck process as detailed in the following.

Proposition 4.25 (cf. [CKS04, Lemma 4.1])**.** *There exists a θ-invariant subset $\overline{\Omega} \in \mathcal{F}$ of Ω of full $\mathbb{P}$-measure such that for every $\omega \in \overline{\Omega}$*

$$\lim_{t\to\pm\infty} \frac{|\omega(t)|}{|t|} = 0$$

and the random variable given by

$$z(\omega) := -\int_{-\infty}^{0} \exp(\tau)\omega(\tau)\mathrm{d}\tau,$$

is well-defined. Furthermore, for $\omega \in \overline{\Omega}$ *the mapping*

$$(t,\omega) \mapsto z(\theta_t\omega) = -\int_{-\infty}^{0} \exp(\tau)\theta_t\omega(\tau)\mathrm{d}\tau,$$

is a stationary solution of (4.5) *with continuous trajectories.*

For $\omega \in \overline{\Omega}$ *we have the following identities*

(i) $\lim_{t\to\pm\infty} \frac{|z(\theta_t\omega)|}{|t|} = 0,$

(ii) $\lim_{t\to\pm\infty} \frac{1}{t}\int_0^t z(\theta_\tau\omega)\mathrm{d}\tau = 0,$

(iii) $\lim_{t\to\pm\infty} \frac{1}{t}\int_0^t |z(\theta_\tau\omega)|\mathrm{d}\tau = \mathbb{E}|z| < \infty.$

Proof. For completeness we give the proof here. By the law of the iterated logarithm there exists a set of full measure $\tilde{\Omega} \subset \mathcal{F}$ such that for any $\omega \in \tilde{\Omega}$

$$\limsup_{t\to\pm\infty} \frac{|B(t,\omega)|}{\sqrt{2|t|\log\log|t|}} = 1.$$

Thus, for every $\omega \in \tilde{\Omega}$

$$0 \le \lim_{t\to\pm\infty} \frac{|\omega(t)|}{|t|} \le \limsup_{t\to\pm\infty} \frac{|\omega(t)|}{\sqrt{2|t|\log\log|t|}} \frac{\sqrt{2\log\log|t|}}{\sqrt{|t|}},$$

and the first statement follows. Furthermore, the sub-linear growth of $\omega \in \tilde{\Omega}$ guarantees that the integral

$$z(\omega) := -\int_{-\infty}^{0} \exp(\tau)\omega(\tau)\mathrm{d}\tau, \tag{4.6}$$

is well-defined for all $\omega \in \tilde{\Omega}$. A solution to

$$\mathrm{d}z = -z\ \mathrm{d}t + \mathrm{d}B(t)$$

is given by the stochastic convolution

$$(\omega, t) \mapsto \int_{-\infty}^{t} \exp(-(t-\tau))\mathrm{d}B(\tau,\omega). \tag{4.7}$$

Using integration by parts we obtain for $s \le t$

$$\begin{aligned}
&\int_s^t \exp(-(t-\tau))\mathrm{d}B(\tau,\omega)\\
&= \exp(0)B(t,\omega) - \exp(-(t-s))B(s,\omega) - \int_s^t \exp(-(t-\tau))B(\tau,\omega)\mathrm{d}\tau\\
&= \omega(t) - \exp(s-t)\omega(s) - \int_{s-t}^0 \exp(\tau)B(\tau+t,\omega)\mathrm{d}\tau\\
&= \omega(t)\int_{s-t}^0 \exp(\tau)\mathrm{d}\tau + \exp(s-t)(\omega(t)-\omega(s)) - \int_{s-t}^0 \exp(\tau)B(\tau+t,\omega)\mathrm{d}\tau\\
&= \exp(s-t)(\omega(t)-\omega(s)) - \int_{s-t}^0 \exp(\tau)\theta_t\omega(\tau)\mathrm{d}\tau.
\end{aligned}$$

Letting $s \to -\infty$ the right hand side tends to $z(\theta_t\omega)$ for those ω satisfying the above growth condition, while the left hand side tends to (4.7). Hence $z(\theta_t\omega)$ is a version of the solution (4.7). The stationarity of this solutions follows by the invariance of the Wiener measure $\mathbb{P}$ with respect to θ.

To prove continuity of the trajectories $t \mapsto F(t) := -\int_{-\infty}^0 \exp(\tau)\omega(t+\tau)\mathrm{d}\tau + \omega(t)$, we only need to show that the integral term is continuous as $\omega(t)$ is continuous by definition. Define $f(\tau,t) := \exp(\tau)\omega(\tau+t)$, then $t \mapsto f(\tau,t)$ is continuous for every $\tau \in \mathbb{R}$. Let $t_0 \in \mathbb{R}$ and $(t_n)_{n\in\mathbb{N}}$ be a sequence with $\lim_{n\to\infty} t_n = t_0$. Define $f_n(\tau) := f(\tau,t_n)$, then $\lim_{n\to\infty} f_n(\tau) = f(\tau,t_0)$ because of continuity. Let $g(\tau) := \exp(\tau)\sup_{\xi\in[t_0-1,t_0+1]}|\omega(\tau+\xi)|$, then $g(\tau)$ is integrable and for n sufficiently large $|f_n(\tau)| \le g(\tau)$ for $\tau \in (-\infty,0]$. By Lebesgue's dominated convergence theorem we therefore have

$$\lim_{n\to\infty} F(t_n) = F(t_0),$$

i.e. the trajectories are continuous.

Finally, we will prove the three identities given at the end of the proposition.

(i) By the law of the iterated logarithm for $\omega \in \tilde{\Omega}$ and $1/2 < \delta < 1$ there exists a constant $C_{\delta,\omega} > 0$ such that

$$|\omega(\tau+t)| \le C_{\delta,\omega} + |\tau+t|^\delta \le C_{\delta,\omega} + |\tau|^\delta + |t|^\delta, \qquad \tau < 0.$$

Thus,

$$\begin{aligned}
\lim_{t\to\pm\infty}\frac{|z(\theta_t\omega)|}{|t|} &= \lim_{t\to\pm\infty}\frac{1}{|t|}\left|-\int_{-\infty}^0 \exp(\tau)\omega(t+\tau)\mathrm{d}\tau + \omega(t)\right|\\
&\le \lim_{t\to\pm\infty}\frac{1}{|t|}\int_{-\infty}^0 \exp(\tau)|\omega(\tau+t)|\mathrm{d}\tau + \lim_{t\to\pm\infty}\frac{|\omega(t)|}{|t|}\\
&\le \lim_{t\to\pm\infty}\int_{-\infty}^0 \exp(\tau)\frac{C_{\delta,\omega}+|\tau|^\delta+|t|^\delta}{|t|}\mathrm{d}\tau + 0 = 0.
\end{aligned}$$

(ii) From (4.6) it follows $\mathbb{E}z = 0$. Furthermore, the expectation of the absolute value of a Gaussian random variable is finite, i.e. $\mathbb{E}|z| < \infty$. By Birkhoff's ergodic theorem (see Theorem 4.3) there exist a θ invariant set $\Omega' \in \mathcal{F}$ of full $\mathbb{P}$-measure such that for $\omega \in \Omega'$

$$\lim_{t\to\pm\infty} \frac{1}{t}\int_0^t z(\theta_\tau\omega)\mathrm{d}\tau = \mathbb{E}z = 0.$$

(iii) Again by the Birkhoff's ergodic theorem there exists a θ invariant set $\Omega^\circ \in \mathcal{F}$ of full $\mathbb{P}$ measure such that for every $\omega \in \Omega^\circ$ the stated identity holds.

Finally we set $\overline{\Omega} := \tilde{\Omega} \cap \Omega' \cap \Omega^\circ$. Then for every $\omega \in \overline{\Omega}$ the above statements hold, $\mathbb{P}(\overline{\Omega}) = 1$ and $\overline{\Omega}$ is θ invariant. This completes the proof. □

Remark 4.26. We now consider the restriction of the Wiener shift onto $\overline{\Omega}$, denoted as $\overline{\theta}$. Likewise, we restrict

$$\overline{\mathcal{F}} := \{\overline{\Omega} \cap F, F \in \mathcal{F}\}$$

and $\overline{\mathbb{P}}$ is the restriction of the Wiener measure to this σ-algebra. Then $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\theta})$ defines again a metric dynamical system. For further analysis we will always consider this MDS, however, for convenience, we will denote it again as $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$.

4.4.2 Ornstein-Uhlenbeck processes in Hilbert spaces

We now consider the metric dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ associated to a two-sided Q-Wiener process $(W(t))_{t\in\mathbb{R}}$ with trace class covariance operator Q on a separable Hilbert space $(H, \|\cdot\|)$, see Example 4.2. Again, we identify $W(t,\omega) = \omega(t)$ for $\omega \in \Omega$ and we consider the equation

$$\mathrm{d}z = -\mu z\ \mathrm{d}t + \mathrm{d}\omega, \tag{4.8}$$

with $\mu > 0$. The following generalization to Proposition 4.25 holds

Proposition 4.27 (cf. [KS99, Lemma 2.5])**.** *There exists a θ-invariant set $\overline{\Omega} \in \mathcal{F}$ of full $\mathbb{P}$-measure such that for $\omega \in \overline{\Omega}$ the mapping*

$$(t,\omega) \mapsto z(\theta_t\omega) = -\mu\int_{-\infty}^0 \exp(\mu\tau)\theta_t\omega(\tau)\ \mathrm{d}\tau,$$

is a stationary solution of (4.8) with continuous trajectories in H. Furthermore, for $\omega \in \overline{\Omega}$

(i) $\lim_{t\to\pm\infty} \frac{\|z(\theta_t\omega)\|}{|t|} = 0,$

(ii) For any $k > 0$

$$\lim_{t\to\pm\infty} \frac{1}{t}\int_0^t \|z(\theta_\tau\omega)\|^k\ \mathrm{d}\tau = \mathbb{E}\|z\|^k.$$

Proof. For completeness we give the proof here. The first statement follows by a similar calculation as in the beginning of the proof of Proposition 4.25. The continuity of the trajectories was proved in Proposition 3.32.

Furthermore, by applying Itô's formula (see Theorem 3.19) with the functional $F(u) = \frac{1}{2}\|u\|^2$ we obtain

$$\|z(\theta_t\omega)\|^2 = \|z(\omega)\|^2 + 2\int_0^t \langle z(\theta_s\omega), \mathrm{d}W(s)\rangle_H \,\mathrm{d}s - 2\mu\int_0^t \|z(\theta_s\omega)\|^2 \,\mathrm{d}s + t\ \mathrm{Tr}Q.$$

Taking expectations and noting that $\mathbb{E}\|z(\theta_t\omega)\|^2 = \mathbb{E}\|z(\omega)\|^2$ by the stationarity of z it follows

$$2\mu\int_0^t \mathbb{E}\|z(\omega)\|^2 \,\mathrm{d}s = t\ \mathrm{Tr}Q$$

and thus $\mathbb{E}\|z(\omega)\|^2 = \frac{\mathrm{Tr}Q}{2\mu} < \infty$.

(i) By Doob's inequality we have

$$\mathbb{E}\sup_{t\in[0,1]} \|z(\theta_t\omega)\|^2 \le C\mathbb{E}\|z(\theta_1\omega)\|^2 = C\mathbb{E}\|z\|^2 < \infty$$

and with a similar argument as in the proof of Proposition 4.6 the statement follows.

(ii) As mentioned above it holds $\mathbb{E}\|z\|^2 = \frac{\mathrm{Tr}Q}{2\mu}$, and since z is Gaussian similar estimates hold for any $k > 0$. The statement thus follows by Birkhoff's ergodic theorem (Theorem 4.3).

□

4.5 Doss-Sussmann transformations

As mentioned above for certain SPDEs it is possible to transform them into random PDEs, via a so-called *Doss-Sussmann-type transformation*, [Dos77,Sus78]. However, this kind of transformation is only possible for SPDEs with additive or linear multiplicative noise. In the following we will give examples for the transformation in both cases.

4.5.1 Additive noise

We consider the equation

$$\mathrm{d}u = [Au + f(u)]\mathrm{d}t + \mathrm{d}W,$$

where W denotes a Q-Wiener process on a Hilbert space H and A generates an analytic semigroup on H. Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be the associated metric dynamical system and let us identify $W(t,\omega) = \omega(t)$ for $\omega \in \Omega$. The unique stationary solution of the linear equation

$$\mathrm{d}z = Az\ \mathrm{d}t + \mathrm{d}\omega,$$

is given by the Ornstein-Uhlenbeck process

$$z(\theta_t\omega) = \int_{-\infty}^{t} e^{(t-s)A}\mathrm{d}\omega.$$

The following Doss-Sussmann-type transformation

$$v(t) = u(t) - z(\theta_t\omega),$$

yields a non-autonomous random PDE for each $\omega \in \Omega$, namely

$$\frac{\mathrm{d}v}{\mathrm{d}t} = Av + f(v + z(\theta_t\omega)).$$

Applications of Doss-Sussmann-type transformations in the case with additive noise can be found (amongst others) in [GLR11, You17, PY19, BH10, BLW09]. We also refer to Section 5.2.4.

4.5.2 Linear multiplicative noise

We consider the following Stratonovich SPDE with multiplicative noise

$$\mathrm{d}u = [Au + f(u)]\mathrm{d}t + u \circ \mathrm{d}B,$$

where B is a real-valued Brownian motion. Again, let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be the associated MDS and let $z(\theta_t\omega)$ denote the one-dimensional Ornstein-Uhlenbeck process solving (4.5). Then

$$v(t) = \exp(-z(\theta_t\omega))u(t),$$

satisfies the following random PDE for each $\omega \in \Omega$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = Av + z(\theta_t\omega)v + \exp(-z(\theta_t\omega))f(\exp(z(\theta_t\omega))v).$$

Applications in the linear multiplicative setting can, for instance, be found in [TY16, CGALdlC17, CL08, Pha20, CLR00, WZ11] and we also refer to Section 5.3.

Chapter 5

Random attractors for stochastic partly dissipative systems

We prove the existence of global random attractors for a class of so-called *stochastic partly dissipative systems*. These systems consist of two reaction-diffusion equations, where the diffusion constant vanishes in one of them. Both equations are linearly coupled and perturbed by (additive or multiplicative) noise. The result for additive noise (Section 5.2) was published in [KNP20] (joint work with Christian Kuehn and Alexandra Neamţu). Similarly, we prove the existence of a random attractor for the *stochastic Field-Noyes* system, a reaction-diffusion system with a *non-linear coupling* between different components. This chapter is based on joint works with Christian Kuehn and Alexandra Neamţu.

5.1 Introduction

Coupled deterministic *reaction-diffusion systems* appear in many models for the dynamical behaviour of biological, chemical and physical systems, see [Mur07] for an overview. They allow for many interesting spatio-temporal phenomena such as *pattern formation* [Tur90] and *oscillatory behaviour*, see also [Kin13].

Having a diffusive behaviour in only one component, leads to so-called *partly dissipative* systems of the form

$$\begin{aligned} \frac{\partial u_1}{\partial t} &= d\Delta u_1 + f(x, u_1, u_2), \\ \frac{\partial u_2}{\partial t} &= g(x, u_1, u_2), \end{aligned} \tag{5.1}$$

where f and g are suitable functions, the so-called *reaction terms*, and $d > 0$ is a diffusion constant. Famous examples of such systems are the *Hodgkin-Huxley* [HH52] and *FitzHugh-Nagumo* systems [Fit61, NAY62], which model the signal transmission across axons in neuroscience [ET10]. Other examples in the biological context concern the modelling of interactions between cellular or intracellular processes and diffusion growth factors, for instance in the field of *carcinogenesis*, see [MCK08, MCK06].

The mathematical analysis of reaction-diffusion systems often concerns the *existence* of solutions, *bifurcations*, properties of solutions (e.g. positivity) and their *long-term behaviour*. Due to their importance in the applied sciences a vast body of literature on theses equations has been accumulated over the years, see for example the monographs [Smo94, Vol14, GR11]. The analysis of partly dissipative systems of the form (5.1) can in parts be more complicated as the semigroup associated to the problem is no longer compact. In [Mar89] such systems have been analysed on a bounded domain and, under certain *polynomial growth* assumptions on the reaction terms, the existence of a global attractor that captures the long-term behaviour of solutions has been proved. Furthermore, bounds on the *Hausdorff* and *fractal dimension* of this attractor have been derived in the same work. In [RBW00] a similar partly dissipative systems was analysed, however, on $\mathbb{R}^n$. Again, the authors proved the existence of a global attractor using that the solutions are uniformly small at infinity for large times. In [Wan09a] a pullback attractor for a non-autonomous version of the FitzHugh-Nagumo system on unbounded domain was derived.

In many situations a *stochastic* version of the reaction-diffusion system provides a more realistic model of the underlying physical system [GOS12], see also [GSB11] in the context of modelling electrically active cells and recall the Introduction. It is thus of great interest to study random and stochastic reaction-diffusion models in a mathematical rigorous way. The type of noise and the way in which it is integrated into a system plays a central role and can lead to widely different effects.

In this chapter, we study a class of *stochastic partial differential equations* with a *partly dissipative* structure, that is, systems of the form

$$\begin{aligned} \mathrm{d}u_1 &= (d\Delta u_1 + f(x,u_1,u_2))\,\mathrm{d}t + B_1(x,u_1,u_2)\,\mathrm{d}W_1, \\ \mathrm{d}u_2 &= g(x,u_1,u_2)\,\mathrm{d}t + B_2(x,u_1,u_2)\,\mathrm{d}W_2, \end{aligned} \tag{5.2}$$

where $W_{1,2}$ are Wiener processes, the f, g are given functions, $B_{1,2}$ are operator-valued, Δ is the Laplace operator and $d > 0$ is the diffusion constant. The equation is posed on a bounded open domain $D \subset \mathbb{R}^n$, $u_{1,2} = u_{1,2}(t,x)$ are the unknowns for $(t,x) \in [0, T_{\max}) \times D$, and $T_{\max}$ is the maximal existence time. The assumptions on the reactions terms will be chosen similar to those imposed on the deterministic system analysed in [Mar89], see Section 5.2 for the precise setting. Firstly, we will consider an *additive random perturbation* by a Wiener process, that is, we choose

$$B_1(x,u_1,u_2) = B_1, \quad B_2(x,u_1,u_2) = B_2,$$

where B_1 and B_2 are fixed bounded linear operators. Secondly, in Section 5.3 we will consider the same system with a *linear multiplicative perturbation* by a real-valued Brownian motion in the Stratonovich sense.

In both cases, the goal will be to analyse the *long-term behaviour* of solutions using the *random dynamical systems approach* as introduced in Chapter 4. To this end, we will transform the stochastic equations into random equations that generate a random dynamical system, for which we can show the existence of a random attractor by deriving a compact absorbing set (cf. Theorem 4.19). More precisely, certain regularity assumptions on the noise terms together with the assumptions on the reaction terms allow us to compute *a-priori bounds* of the solution, which are used to construct a *bounded absorbing set*. Due to the absence of the regularizing effect of the Laplacian in the second component, a compact embedding result can not be applied directly in order to obtain a compact absorbing set. For this reason a more refined compactness argument based on a suitable *splitting technique* needs to be derived, similar to the deterministic setting, see [Mar89, Tem12].

Let us now briefly summarise, without claiming completeness, previous results on stochastic reaction-diffusion systems with a partly dissipative structure. The analysis focused so far mainly on the specific stochastic FitzHugh-Nagumo system and its variants. In [BM08] the authors proved the existence and uniqueness of mild solutions for this system on the bounded domain $(0, 1)$ where the additive random perturbation is given by a Q-Wiener process in both components. Furthermore, they showed the existence of an *invariant ergodic measure* associated to the transition semigroup. In [SS16] a similar neuronal model was analysed, here, a reaction-diffusion equation, additively perturbed by a cylindrical Wiener process, is coupled to a system of ODEs, which itself is driven by multiplicative noise. The authors prove existence and uniqueness of variational solutions under local Lipschitz and monotonicity assumptions on the reaction terms. They also discuss a numerical approximation scheme. In [HSZS18] deterministic PDEs are coupled to a SDE modelling acid-mediated tumor invasion. The authors focus on global well-posedness of the problem and simulations of solutions. Furthermore, we emphasize that other dynamical aspects for similar systems have been investigated, e.g. inertial manifolds and master-slave synchronization in [CS10].

Regarding the *asymptotic behaviour* of solutions in terms of random attractors, to the best of our knowledge, only the FitzHugh-Nagumo system has been analysed in detail. In [Wan09b] the existence of a random attractor for the FitzHugh-Nagumo system perturbed by additive real-valued Wiener processes on unbounded domain is proved. Furthermore, in [AW13b] respectively [AW13a] a non-autonomous version of the stochastic FitzHugh-Nagumo system on unbounded domain is analysed, and amongst others the existence of a random attractor is shown in case of additive respectively multiplicative real-valued Wiener noise. The regularity of the attractor is further studied in [LY16] for the additive setting and in [ZG17] for the multiplicative setting. Furthermore, in [ZW16] the authors consider a similar non-autonomous version of the FitzHugh-Nagumo system driven by multiplicative real-valued Wiener-processes, however, here defined on a bounded domain, and they show the existence of random attractors under slightly different assumptions. The setting with a non-autonomous FitzHugh-Nagumo system with real-valued multiplicative noise on unbounded thin domains was analysed in [SWLW19]. Finally, a result with colored noise for the non-autonomous system on a bounded domain can be found in [GW18].

In this chapter we will develop a more general theory of stochastic partly dissipative systems, that is, we will allow for a whole class of reaction terms, similar to that analysed by [Mar89] in the deterministic setting. Moreover, in the case with additive noise we will extend the theory to infinite-dimensional noise.

However, this class of reaction terms only includes systems where the coupling between different components is linear. Nevertheless, there are many reaction-diffusion equations appearing in the natural sciences with a non-linear coupling, for example the *Field-Noyes system* [NFK72, FKN72, FN74], which describes the Belousov-Zhabotinskii reaction in chemical kinetics. Marion analysed in a second part of her work in [Mar87] (fully dissipative) and [Mar89] (partly dissipative) deterministic reaction-diffusion systems that exhibit an invariant region and she proved the existence of a global attractor for such systems as well. As a first step towards extending such a result to the stochastic setting we will consider in Section 5.5 the fully dissipative *stochastic Field-Noyes system* and prove the existence of a random attractor as well.

Let us summarise the structure of this chapter: In Section 5.2 we will analyse partly dissipative systems with an additive random perturbation given by a Wiener process. We will start by stating our precise assumptions (Subsection 5.2.1) and by formulating the system as an abstract Cauchy problem (Subsection 5.2.2) for which solutions exists locally in time (Subsection 5.2.3). Subsequently, in Subsection 5.2.4 we transform the stochastic problem into a random equation via an Ornstein-Uhlenbeck process and formulate the associated random dynamical system. The existence of a bounded absorbing set for this system will be derived in Subsection 5.2.5 and the compactness argument mentioned above can be found in Subsection 5.2.6. The next section, Section 5.3, contains the same analysis for the setting with multiplicative Stratonovich noise. We will give two concrete examples arising from applications that fall into the class of analysed systems in Section 5.4. In the following Section 5.5 we analyse the fully dissipative stochastic Fields-Noyes model, a system with non-linear coupling, in a similar fashion as in Section 5.2 and 5.3. Finally, in Section 5.6 we conclude this chapter with an outlook on further interesting research questions that we plan to approach in the future.

5.2 Stochastic partly dissipative systems with additive noise

Note that this section is based on [KNP20], joint work with Christian Kuehn and Alexandra Neamţu; in particular, all the technical calculations are copied from there.

Let $D \subset \mathbb{R}^n$ be a bounded open set with regular boundary, set $H := L^2(D)$ and let U_1, U_2 be two separable Hilbert spaces. We consider the following coupled, partly dissipative system with additive noise

$$\begin{aligned} \mathrm{d}u_1 &= (d\Delta u_1 - h(x,u_1) - f(x,u_1,u_2))\,\mathrm{d}t + B_1\,\mathrm{d}W_1, \\ \mathrm{d}u_2 &= (-\sigma(x)u_2 - g(x,u_1))\,\mathrm{d}t + B_2\,\mathrm{d}W_2, \end{aligned} \tag{5.3}$$

where $u_{1,2} = u_{1,2}(t,x)$, $(t,x) \in [0,T] \times D$, $T > 0$, $W_{1,2}$ are cylindrical Wiener processes on U_1 respectively U_2, and Δ is the Laplace operator. Furthermore, $B_1 \in L(U_1, H)$, $B_2 \in L(U_2, H)$ and $d > 0$ is a parameter controlling the strength of the diffusion in the first component. The system is equipped with non-random initial conditions

$$u_1(0,x) = u_1^0(x) \in L^2(D), \quad u_2(0,x) = u_2^0(x) \in L^2(D),$$

and a Dirichlet boundary condition for the first component

$$u_1(t,x) = 0 \quad \text{on } [0,T] \times \partial D.$$

We will denote by A the realization of the Laplace operator with Dirichlet boundary conditions, more precisely we define the operator $A : \mathcal{D}(A) \to L^2(D)$ as $Au = d\Delta u$ with domain $\mathcal{D}(A) := H^2(D) \cap H_0^1(D) \subset L^2(D)$.

Remark 5.1. Note that A is a self-adjoint operator that possesses a complete orthonormal system of eigenfunctions $\{e_k\}_{k\in\mathbb{N}}$ of $L^2(D)$. Within this chapter we assume that there exists $\kappa > 0$ such that $|e_k(x)|^2 < \kappa$ for all $k \in \mathbb{N}$ and $x \in D$. This holds for example when $D = [0,\pi]^n$.

5.2.1 Assumptions

For the deterministic reaction terms appearing in (5.3) we impose the following assumptions.

Assumptions 5.2. (Reaction terms)

(i) $h \in C^2(\mathbb{R}^n \times \mathbb{R})$ and there exist $\delta_1, \delta_2, \delta_3 > 0$, $p > 2$ such that

$$\delta_1 |u_1|^p - \delta_3 \le h(x,u_1)u_1 \le \delta_2 |u_1|^p + \delta_3. \tag{5.4}$$

(ii) $f \in C^2(\mathbb{R}^n \times \mathbb{R} \times \mathbb{R})$ and there exist $\delta_4 > 0$ and $0 < p_1 < p - 1$ such that

$$|f(x,u_1,u_2)| \le \delta_4 (1 + |u_1|^{p_1} + |u_2|). \tag{5.5}$$

(iii) $\sigma \in C^2(\mathbb{R}^n)$ and there exist $\delta, \tilde{\delta} > 0$ such that

$$\delta \le \sigma(x) \le \tilde{\delta}. \tag{5.6}$$

(iv) $g \in C^2(\mathbb{R}^n \times \mathbb{R})$ and there exists $\delta_5 > 0$ such that

$$|g_u(x, u_1)| \leq \delta_5, \quad |g_{x_i}(x, u_1)| \leq \delta_5(1 + |u_1|), \quad i = 1, \ldots, n. \tag{5.7}$$

In particular, Assumptions 5.2 (i) and (iv) imply that there exist $\delta_7, \delta_8 > 0$ such that

$$|g(x, \xi)| \leq \delta_7(1 + |\xi|), \qquad \text{for all } \xi \in \mathbb{R},\ x \in D, \tag{5.8}$$

$$|h(x, \xi)| \leq \delta_8(1 + |\xi|^{p-1}), \qquad \text{for all } \xi \in \mathbb{R},\ x \in D. \tag{5.9}$$

Remark 5.3. Assumptions 5.2 are identical to those given in [Mar89], except that in the deterministic setting only a lower bound on σ was assumed.

We always consider an underlying filtered probability space denoted as $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ that will be specified later on. In order to guarantee certain regularity properties of the noise terms, we make the following additional assumptions.

Assumptions 5.4. (Noise)

(i) We assume that $B_2 : U_2 \to H$ is a Hilbert-Schmidt operator. In particular, this implies that $Q_2 := B_2 B_2^*$ is a trace class operator and $B_2 W_2$ is a Q_2-Wiener process.

(ii) We assume that $B_1 \in L(U_1, H)$ and that the operator Q_t defined by

$$Q_t u = \int_0^t \exp(sA)\, Q_1 \exp(sA^*)\, u \; \mathrm{d}s, \quad u \in H, t \geq 0,$$

where $Q_1 := B_1 B_1^*$, is of trace class. Hence, $B_1 W_1$ is a Q_1-Wiener process as well.

(iii) Let $U_1 = H$. There exists an orthonormal basis $\{e_k\}_{k=1}^{\infty}$ of H and sequences $\{\lambda_k\}_{k=1}^{\infty}$ and $\{\delta_k\}_{k=1}^{\infty}$ such that

$$A e_k = -\lambda_k e_k, \qquad Q_1 e_k = \delta_k e_k, \quad k \in \mathbb{N}.$$

Furthermore, we assume that there exists $\alpha \in \left(0, \frac{1}{2}\right)$ such that

$$\sum_{k=1}^{\infty} \delta_k \lambda_k^{2\alpha+1} < \infty.$$

Remark 5.5. Assumptions 5.4 guarantee that the stochastic convolution is a well-defined process with sufficient regularity properties, see Lemma 5.10. As an example, one could choose $B_1 = (-A)^{-\gamma/2}$ with $\gamma > \frac{n}{2} - 1$ to ensure that Assumptions 5.4 (ii)-(iii) hold for α with $2\alpha < \gamma - \frac{n}{2} + 1$, see [DP12, Chapter 4].

5.2.2 The Cauchy problem

We will formulate problem (5.3) as an abstract Cauchy problem. Let us define the following product space

$$\mathbb{H} := L^2(D) \times L^2(D);$$

with the norm $\|(u_1, u_2)^\top\|_{\mathbb{H}}^2 = \|u_1\|_2^2 + \|u_2\|_2^2$ this becomes a separable Hilbert space. Let $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ denote the corresponding scalar product. Furthermore, we set

$$\mathbb{V} := H_0^1(D) \times L^2(D),$$

with the norm $\|(u_1, u_2)^\top\|_{\mathbb{V}}^2 = \|u_1\|_{H^1(D)}^2 + \|u_2\|_2^2$. We define the following linear operator

$$\mathbf{A} := \begin{pmatrix} A & 0 \\ 0 & -\sigma(x) \end{pmatrix},$$

where $\mathbf{A} : \mathcal{D}(\mathbf{A}) \subset \mathbb{H} \to \mathbb{H}$ with $\mathcal{D}(\mathbf{A}) = \mathcal{D}(A) \times L^2(D)$. Since all the reaction terms are twice continuously differentiable they obey in particular the Carathéodory conditions [Zei89]. Thus, the corresponding Nemytskii operator is defined by

$$\begin{aligned} \mathbf{F}((u_1, u_2)^\top)(x) &:= \begin{pmatrix} F_1((u_1, u_2)^\top)(x) \\ F_2((u_1, u_2)^\top)(x) \end{pmatrix}, \\ &:= \begin{pmatrix} -h(x, u_1(x)) - f(x, u_1(x), u_2(x)) \\ -g(x, u_1(x)) \end{pmatrix}, \end{aligned}$$

where $\mathbf{F} : \mathcal{D}(\mathbf{F}) \subset \mathbb{H} \to \mathbb{H}$ and $\mathcal{D}(\mathbf{F}) := \mathbb{H}$. By setting

$$\mathbf{W} := \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}, \qquad \mathbf{B} := \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad \text{and} \quad u := \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$

we can rewrite the system (5.3) as an abstract Cauchy problem on the space $\mathbb{H}$

$$\mathrm{d}u = (\mathbf{A}u + \mathbf{F}(u))\ \mathrm{d}t + \mathbf{B}\ \mathrm{d}\mathbf{W}, \tag{5.10}$$

with initial condition

$$u(0) = u^0 := \begin{pmatrix} u_1^0 \\ u_2^0 \end{pmatrix}. \tag{5.11}$$

5.2.3 Mild solutions and the stochastic convolution

We are interested in *mild solutions* to (5.10)-(5.11). First of all, let us note the following

Lemma 5.6. $\mathbf{A}$ *generates an analytic semigroup* $\{\exp(t\mathbf{A})\}_{t \geq 0}$ *on* $\mathbb{H}$.

Proof. Consider

$$\mathbf{A} = \underbrace{\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}}_{=:A_1} + \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & -\sigma(x) \end{pmatrix}}_{=:A_2}.$$

It is well known that A_1 generates an analytic semigroup on $\mathbb{H}$ [RR06, Theorem 12.40]. Furthermore, A_2 is a bounded multiplication operator on $\mathbb{H}$. Hence the statement follows by Lemma C.22. □

Remark 5.7. Also note that A with $\mathcal{D}(A) = W^{2,p}(D) \cap W_0^{1,p}(D)$ generates an analytic semigroup $\{\exp(tA)\}_{t\geq 0}$ on $L^p(D)$ for every $1 < p < \infty$ and $-A$ is a positive, sectorial operator [SY02, Theorem 38.2]. In particular, we have for $u \in L^p(D)$ that for every $\alpha \geq 0$ there exists a constant $C_\alpha > 0$ such that

$$\|(-A)^\alpha \exp(tA)\, u\|_p \leq C_\alpha t^{-\alpha} \exp(-\delta t)\, \|u\|_p, \qquad \text{for all } t > 0,$$

where $\delta > 0$, see Lemma C.21. Furthermore, the domain $\mathcal{D}((-A)^\alpha)$ can be identified with the Sobolev space $W^{2\alpha,p}(D)$ and thus we have in our setting for $t > 0$

$$\|\exp(tA)\, u\|_{W^{\alpha,p}(D)} \leq C_\alpha t^{-\alpha/2} \exp(-\delta t)\, \|u\|_p. \tag{5.12}$$

The *stochastic convolution* corresponding to (5.10) is given by (see [Nag89, Proposition 3.1])

$$\begin{aligned} W_{\mathbf{A}}(t) &= \int_0^t \begin{pmatrix} \exp((t-s)A) & 0 \\ 0 & \exp(-(t-s)\sigma(x)) \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \,\mathrm{d}\mathbf{W}(s) \\ &= \begin{pmatrix} \int_0^t \exp((t-s)A)\, B_1 \,\mathrm{d}W_1(s) \\ \int_0^t \exp(-(t-s)\sigma(x))\, B_2 \,\mathrm{d}W_2(s) \end{pmatrix}. \end{aligned}$$

This is a well-defined $\mathbb{H}$-valued Gaussian process. Furthermore, Assumptions 5.4 (i) and (ii) ensure that $W_{\mathbf{A}}(t)$ is mean-square continuous and $\mathcal{F}_t$-measurable, see Theorem 3.31.

Remark 5.8. As $W_{\mathbf{A}}$ is a Gaussian process, we can bound all its higher-order moments, i.e. for $p \geq 1$ we have

$$\sup_{t\in[0,T]} \mathbb{E}\|W_{\mathbf{A}}(t)\|_{\mathbb{H}}^p < \infty. \tag{5.13}$$

This follows from the Kahane-Khintchine inequality, see [vN08, Theorem 3.12].

Proposition 5.9. *Let Assumptions 5.2 and 5.4 (i)-(ii) hold. Then a mild solution*

$$u(t) = \exp(t\mathbf{A})u^0 + \int_0^t \exp((t-s)\mathbf{A})\mathbf{F}(u(s)) \,\mathrm{d}s + W_{\mathbf{A}}(t),$$

of (5.10)-(5.11) *exists locally-in-time in*

$$L^2(\Omega; C([0,T];\mathbb{H})) \cap L^2(\Omega; L^2([0,T];\mathbb{V})),$$

for some $T > 0$.

Proof. Since the reactions terms are locally Lipschitz continuous, the existence of local in time solutions follows from classical SPDE theory, see for example [DW14, Theorem 4.17] and the comments thereafter. □

5.2.4 Associated RDS

We consider $\mathbb{H} := L^2(D) \times L^2(D)$ and let $\mathcal{T}$ denote the set of all tempered subsets of $\mathbb{H}$. In the sequel, we consider the fixed canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ corresponding to a two-sided $\mathbb{H}$-valued Wiener process, see Example 4.2. Here $\mathbb{P}$ is the distribution of the trace class Wiener process $\tilde{W}(t) := (\tilde{W}_1(t), \tilde{W}_2(t)) = (B_1 W_1(t), B_2 W_2(t))$ (extended to $t \in \mathbb{R}$), where we recall that B_1 and B_2 fulfil Assumptions 5.4. We identify the elements of Ω with the paths of these Wiener processes, more precisely

$$\tilde{W}(t,\omega) := (\tilde{W}_1(t,\omega_1), \tilde{W}_2(t,\omega_2)) = (\omega_1(t), \omega_2(t)) =: \omega(t), \text{ for } \omega \in \Omega.$$

Together with the Wiener shift the quadruple $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ defines a metric dynamical system.

Let us now consider the following equations

$$\mathrm{d}z_1 = A z_1 \,\mathrm{d}t + \,\mathrm{d}\omega_1, \tag{5.14}$$

$$\mathrm{d}z_2 = -\sigma(x) z_2 \,\mathrm{d}t + \mathrm{d}\omega_2. \tag{5.15}$$

The stationary solutions of (5.14)-(5.15) are given by the following *Ornstein-Uhlenbeck processes* (see Subsection 4.4.2)

$$(t,\omega) \mapsto z_1(\theta_t \omega) = \int_{-\infty}^{t} e^{(t-s)A} \,\mathrm{d}\omega_1(s) = \int_{-\infty}^{0} e^{-sA} \,\mathrm{d}\theta_t \omega_1(s), \tag{5.16}$$

$$(t,\omega) \mapsto z_2(\theta_t \omega) = \int_{-\infty}^{t} e^{-(t-s)\sigma(x)} \,\mathrm{d}\omega_2(s) = \int_{-\infty}^{0} e^{s\sigma(x)} \,\mathrm{d}\theta_t \omega_2(s). \tag{5.17}$$

Here, we observe that for $t = 0$

$$z_1(\omega) = \int_{-\infty}^{0} e^{-sA} \,\mathrm{d}\omega_1(s), \quad z_2(\omega) = \int_{-\infty}^{0} e^{s\sigma(x)} \,\mathrm{d}\omega_2(s).$$

In the following we will need a spatial regularity result of the Ornstein-Uhlenbeck processes (5.16)-(5.17). For this, Assumption 5.4 (iii) is crucial.

Lemma 5.10. *Suppose that Assumptions 5.2 and 5.4 hold. Then for every $p \geq 1$*

$$\|z_1(\omega)\|_p^p \text{ and } \|z_2(\omega)\|_2^2$$

are tempered random variables.

Proof. Using the assumption $0 < \delta \leq \sigma(x) \leq \tilde{\delta}$, the statement for $\|z_2(\omega)\|_2^2$ follows by the same argument as in the proof of Proposition 4.27 (i). Furthermore, under Assumptions 5.4 all conditions required in Proposition 3.34 are fulfilled (in particular recall Remark 5.7) and from there we can infer

$$\mathbb{E}\left(\sup_{t\in[0,1]} \|z_1(\theta_t\omega)\|_p^p\right) < \infty,$$

i.e., temperedness of $\|z_1(\omega)\|_p^p$ follows by Proposition 4.6. □

Remark 5.11.

(i) Note that Assumption 5.4 (iii) together with the boundedness of e_k for $k \in \mathbb{N}$ were essential for the proof of Proposition 3.34. One can extend such statements for general open bounded domains in $D \subset \mathbb{R}^n$, according to Remark 5.27 and Theorem 5.28 in [DPZ92].

(ii) One can show in a similar way as in the proof of Proposition 3.34 that $z_1 \in W^{1,p}(D)$ and in particular also $\|\nabla z_1(\omega)\|_p^p$ is a tempered random variable for all $p \geq 1$.

Let us perform the following *Doss-Sussmann transformation*

$$v(t) = u(t) - z(\theta_t\omega),$$

where $v(t) = (v_1(t), v_2(t))^\top$, $z(\omega) = (z_1(\omega_1), z_2(\omega_2))^\top$ and $u(t) = (u_1(t), u_2(t))^\top$ is a solution to (5.10)-(5.11). Then $v(t)$ satisfies the following random equation for each $\omega \in \Omega$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \mathbf{A}v + \mathbf{F}(v + z(\theta_t\omega)), \tag{5.18}$$

$$v(0) = u^0 - z(\omega) =: v^0. \tag{5.19}$$

Component-wise this reads as follows

$$\frac{\mathrm{d}v_1(t)}{\mathrm{d}t} = d\Delta v_1(t) - h(x, v_1(t) + z_1(\theta_t\omega)) - f(x, v_1(t) + z_1(\theta_t\omega), v_2(t) + z_2(\theta_t\omega)), \tag{5.20}$$

$$\frac{\mathrm{d}v_2(t)}{\mathrm{d}t} = -\sigma(x)v_2(t) - g(x, v_1(t) + z_1(\theta_t\omega)). \tag{5.21}$$

In the equations above no stochastic differentials appear, hence they can be considered path-wise, i.e., for every ω instead just for $\mathbb{P}$-almost every ω. For every ω (5.18) is a deterministic equation, where $z(\theta_t\omega)$ can be regarded as a non-autonomous, time-continuous perturbation. In particular, we have

Lemma 5.12. *The map* $\psi : \mathbb{R}^+ \times \Omega \times \mathbb{H} \to \mathbb{H}$ *with*

$$\psi(t,\omega,(v_1^0,v_2^0)) := \begin{pmatrix} v_1(t,\omega,v_1^0) \\ v_2(t,\omega,v_2^0) \end{pmatrix}, \tag{5.22}$$

defines a continuous RDS over the MDS $(\Omega,\mathcal{F},\mathbb{P},\theta)$.

Proof. [CV96] guarantees that for all $v^0 = (v_1^0, v_2^0)^\top \in \mathbb{H}$ there exists a weak solution $v(\cdot,\omega,v^0) \in C([0,\infty),\mathbb{H})$ with $v_1(0,\omega,v_1^0) = v_1^0$, $v_2(0,\omega,v_2^0) = v_2^0$. Moreover, the mapping $\mathbb{H} \ni v_0 \mapsto v(t,\omega,v_0) \in \mathbb{H}$ is continuous and also the measurability with respect to ω is guaranteed. In particular, ψ is jointly measurable with respect to (t,ω,v_0). Furthermore, note the equivalence of weak and mild solutions [Bal77, Już14]. The cocycle property can be verified easily using the mild formulation. We have

$$v(t,\omega,v^0) = e^{t\mathbf{A}}v^0 + \int_0^t e^{(t-r)\mathbf{A}}\mathbf{F}(v(r,\omega,v^0) + z(\theta_r\omega))\ \mathrm{d}r,$$

and thus

$$\begin{aligned}
&\psi(t,\theta_s\omega,\psi(s,\omega,v^0)) \\
&= e^{t\mathbf{A}}\psi(s,\omega,v^0) + \int_0^t e^{(t-r)\mathbf{A}}\mathbf{F}(v(r+s,\omega,v^0) + z(\theta_{s+r}\omega))\ \mathrm{d}r \\
&= e^{t\mathbf{A}}\left(e^{s\mathbf{A}}v^0 + \int_0^s e^{(s-r)\mathbf{A}}\mathbf{F}(v(r,\omega,v^0) + z(\theta_r\omega))\mathrm{d}r\right) \\
&\quad + \int_0^t e^{(t-r)\mathbf{A}}\mathbf{F}(v(r+s,\omega,v^0) + z(\theta_{s+r}\omega))\ \mathrm{d}r \\
&= e^{(t+s)\mathbf{A}}v^0 + \int_0^s e^{(t+s-r)\mathbf{A}}\mathbf{F}(v(r,\omega,v^0) + z(\theta_r\omega))\mathrm{d}r \\
&\quad + \int_s^{t+s} e^{(t+s-r)\mathbf{A}}\mathbf{F}(v(r,\omega,v^0) + z(\theta_r\omega))\ \mathrm{d}r \\
&= \psi(t+s,\omega,v^0).
\end{aligned}$$

□

Lemma 5.13. *The map* $\varphi : \mathbb{R}^+ \times \Omega \times \mathbb{H} \to \mathbb{H}$ *with*

$$\varphi(t,\omega,(u_1^0,u_2^0)) := \begin{pmatrix} u_1(t,\omega,u_1^0) \\ u_2(t,\omega,u_2^0) \end{pmatrix}$$

defines a continuous RDS over the MDS $(\Omega,\mathcal{F},\mathbb{P},\theta)$, *conjugated to the RDS* ψ.

Proof. Define the mapping $T : \Omega \times \mathbb{H} \to \mathbb{H}$ by

$$T(\omega,v) = v + z(\omega).$$

Clearly this map fulfils conditions (i) and (ii) from Proposition 4.22, where $T^{-1}(\omega, v) = v - z(\omega)$. Thus, together with Lemma 5.12,

$$\begin{aligned}\varphi(t,\omega,(u_1^0,u_2^0)) :&= T(\theta_t\omega, \psi(t,\omega,T^{-1}(\omega,u^0)))\\ &= \psi(t,\omega,T^{-1}(\omega,u^0)) + z(\theta_t\omega)\\ &= \psi(t,\omega,u^0 - z(\omega)) + z(\theta_t\omega)\\ &= \begin{pmatrix} u_1(t,\omega,u_1^0)\\ u_2(t,\omega,u_2^0)\end{pmatrix},\end{aligned}$$

defines a continuous RDS over the MDS, conjugated to the RDS ψ. □

The RDS φ is associated to the stochastic partly dissipative system (5.10) and conjugated to the RDS ψ. In the following we will prove the existence of a random attractor for the RDS ψ. By Theorem 4.23 this implies the existence of a random attractor for φ, i.e. for our original stochastic system.

5.2.5 Bounded absorbing set

In the following we will prove the existence of a bounded absorbing set for the RDS (5.22). In the calculations we will make use of some standard analytical inequalities, which are stated in Appendix B.1.

Lemma 5.14. *Suppose Assumptions 5.2 and 5.4 hold. Then there exists a set $\mathcal{B} \in \mathcal{T}$ such that $\mathcal{B}$ is a bounded $\mathcal{T}$-absorbing set for the RDS ψ.*

Proof. To show the existence of a bounded absorbing set, we want to make use of Remark 4.16, i.e. we need an a-priori estimate in $\mathbb{H}$. Let $v = (v_1, v_2)^\top$ be the solution of (5.18), then

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\|v_1\|_2^2 + \|v_2\|_2^2\right) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|v\|_{\mathbb{H}}^2 = \left\langle \frac{\mathrm{d}}{\mathrm{d}t}v, v\right\rangle_{\mathbb{H}}\\ &= \langle \mathbf{A}v + \mathbf{F}(v + z(\theta_t\omega)), v\rangle_{\mathbb{H}}\\ &= \langle dAv_1, v_1\rangle + \langle F_1(v + z(\theta_t\omega)), v_1\rangle - \langle \sigma(x)v_2, v_2\rangle + \langle F_2(v + z(\theta_t\omega)), v_2\rangle\\ &= -d\|\nabla v_1\|_2^2 \underbrace{-\langle h(x, v_1 + z_1(\theta_t\omega)), v_1\rangle}_{=:I_1} \underbrace{-\langle f(x, v_1 + z_1(\theta_t\omega), v_2 + z_2(\theta_t\omega)), v_1\rangle}_{=:I_2}\\ &\qquad - \delta\|v_2\|_2^2 \underbrace{-\langle g(x, v_1 + z_1(\theta_t\omega)), v_2\rangle}_{=:I_3},\end{aligned}$$

where we have used (5.6). We now estimate I_1-I_3 separately. Deterministic constants denoted as $C, C_1, C_2, ...$ may change from line to line. Using (5.4) and (5.9) we calculate

$$
\begin{aligned}
I_1 &= -\int_D h(x, v_1 + z_1(\theta_t\omega))v_1 \ \mathrm{d}x \\
&= -\int_D h(x, v_1 + z_1(\theta_t\omega))(v_1 + z_1(\theta_t\omega)) \ \mathrm{d}x \\
&\quad + \int_D h(x, v_1 + z_1(\theta_t\omega))z_1(\theta_t\omega) \ \mathrm{d}x \\
&\leq -\int_D \delta_1 |u_1|^p \ \mathrm{d}x + \int_D \delta_3 \ \mathrm{d}x + \int_D |h(x, v_1 + z_1(\theta_t\omega))||z_1(\theta_t\omega)| \ \mathrm{d}x \\
&\leq -\delta_1 \|u_1\|_p^p + C + \delta_8 \int_D (1 + |u_1|^{p-1})|z_1(\theta_t\omega)| \ \mathrm{d}x \\
&= -\delta_1 \|u_1\|_p^p + C + \delta_8 \|z_1(\theta_t\omega)\|_1 + \delta_8 \int_D |u_1|^{p-1}|z_1(\theta_t\omega)| \ \mathrm{d}x \\
&\leq -\delta_1 \|u_1\|_p^p + C + C_1 \|z_1(\theta_t\omega)\|_2^2 + \frac{\delta_1}{2}\|u_1\|_p^p + C_2 \|z_1(\theta_t\omega)\|_p^p \\
&= -\frac{\delta_1}{2}\|u_1\|_p^p + C + C_1 \left(\|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_p^p\right).
\end{aligned}
$$

Furthermore, with (5.5) we estimate

$$
\begin{aligned}
I_2 &= -\int_D f(x, v_1 + z_1(\theta_t\omega), v_2 + z_2(\theta_t\omega))v_1 \ \mathrm{d}x \\
&\leq \int_D |f(x, v_1 + z_1(\theta_t\omega), v_2 + z_2(\theta_t\omega))||u_1 - z_1(\theta_t\omega)| \ \mathrm{d}x \\
&\leq \int_D \delta_4 (1 + |u_1|^{p_1} + |u_2|)|u_1| \ \mathrm{d}x \\
&\quad + \int_D \delta_4 (1 + |u_1|^{p_1} + |u_2|)|z_1(\theta_t\omega)| \ \mathrm{d}x \\
&= \int_D \delta_4 (|u_1| + |u_1|^{p_1+1}) \ \mathrm{d}x + \int_D \delta_4 |u_1||u_2| \ \mathrm{d}x + \delta_4 \|z_1(\theta_t\omega)\|_1 \\
&\quad + \int_D \delta_4 |u_1|^{p_1}|z_1(\theta_t\omega)| \ \mathrm{d}x + \int_D \delta_4 |u_2||z_1(\theta_t\omega)| \ \mathrm{d}x \\
&\leq \int_D \delta_4 (|u_1| + |u_1|^{p_1+1}) \ \mathrm{d}x + \int_D \delta_4 |u_1||u_2| \ \mathrm{d}x + \delta_4 \|z_1(\theta_t\omega)\|_2^2 + C \\
&\quad + \int_D \frac{\delta_4}{2}|u_1|^{p_1+1} \ \mathrm{d}x + C_1 \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1} + \int_D \delta_4 |u_2||z_1(\theta_t\omega)| \ \mathrm{d}x \\
&\leq \int_D \delta_4 \frac{3}{2}(|u_1| + |u_1|^{p_1+1}) \ \mathrm{d}x + \int_D \delta_4 |u_1||u_2| \ \mathrm{d}x + C \\
&\quad + C_1 \left(\|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1}\right) + \int_D \delta_4 |u_2||z_1(\theta_t\omega)| \ \mathrm{d}x.
\end{aligned}
$$

With (5.8) we compute

$$\begin{aligned}
I_3 &= -\int_D g(x, v_1 + z_1(\theta_t\omega))v_2 \,\mathrm{d}x \\
&\le \int_D |g(x,u_1)||u_2 - z_2(\theta_t\omega)| \,\mathrm{d}x \\
&\le \int_D \delta_7(1+|u_1|)|u_2| \,\mathrm{d}x + \int_D \delta_7(1+|u_1|)|z_2(\theta_t\omega)| \,\mathrm{d}x \\
&= \int_D \delta_7(1+|u_1|)|u_2| \,\mathrm{d}x + \delta_7\|z_2(\theta_t\omega)\|_1 + \int_D \delta_7|u_1||z_2(\theta_t\omega)| \,\mathrm{d}x \\
&\le \int_D \delta_7(1+|u_1|)|u_2| \,\mathrm{d}x + \delta_7\|z_2(\theta_t\omega)\|_2^2 + C + \int_D \delta_7|u_1||z_2(\theta_t\omega)| \,\mathrm{d}x.
\end{aligned}$$

Combining the estimates for I_2 and I_3 yields

$$\begin{aligned}
&I_2 + I_3 \\
&\le \int_D \delta_7(1+|u_1|)|u_2| \,\mathrm{d}x + \int_D \delta_7|u_1||z_2(\theta_t\omega)| \,\mathrm{d}x + \int_D \delta_4\frac{3}{2}(|u_1| + |u_1|^{p_1+1}) \,\mathrm{d}x \\
&\quad + \int_D \delta_4|u_1||u_2| \,\mathrm{d}x + \int_D \delta_4|u_2||z_1(\theta_t\omega)| \,\mathrm{d}x \\
&\quad + C + C_1\left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1}\right) \\
&\le (\delta_4+\delta_7)\int_D (1+|u_1|)|u_2| \,\mathrm{d}x + \int_D \delta_7|u_1||z_2(\theta_t\omega)| \,\mathrm{d}x + \int_D \delta_4\frac{3}{2}(|u_1| + |u_1|^{p_1+1}) \,\mathrm{d}x \\
&\quad + \int_D \delta_4|u_2||z_1(\theta_t\omega)| \,\mathrm{d}x + C + C_1\left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1}\right) \\
&\le \frac{\delta}{16}\|u_2\|_2^2 + C_2\int_D (1+|u_1|)^2 \,\mathrm{d}x + \int_D \delta_7|u_1||z_2(\theta_t\omega)| \,\mathrm{d}x \\
&\quad + \int_D \delta_4\frac{3}{2}(|u_1| + |u_1|^{p_1+1}) \,\mathrm{d}x + \int_D \delta_4|u_2||z_1(\theta_t\omega)| \,\mathrm{d}x \\
&\quad + C + C_1\left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1}\right) \\
&= \frac{\delta}{16}\|u_2\|_2^2 + \delta_4\frac{3}{2}\int_D \left(|u_1| + |u_1|^{p_1+1} + C_2(1+|u_1|)^2\right) \,\mathrm{d}x \\
&\quad + \int_D \delta_7|u_1||z_2(\theta_t\omega)| \,\mathrm{d}x + \int_D \delta_4|u_2||z_1(\theta_t\omega)| \,\mathrm{d}x \\
&\quad + C + C_1\left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1}\right) \\
&\le \frac{\delta}{16}\|u_2\|_2^2 + C_2\int_D (1+|u_1|^q) \,\mathrm{d}x + \frac{\delta_1}{8}\|u_1\|_2^2 + \frac{\delta}{16}\|u_2\|_2^2 \\
&\quad + C + C_1\left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1}\right),
\end{aligned}$$

where we have used that for $q = \max\{p_1+1, 2\} < p$ there exists a constant C_2 such that

$$C_1 \left(|\xi| + |\xi|^{p_1+1} + C(1+|\xi|)^2\right) \leq C_2(|\xi|^q + 1), \quad \text{for all } \xi \in \mathbb{R}. \tag{5.23}$$

Thus,

$$\begin{aligned}
&I_2 + I_3 \\
&\leq \frac{\delta}{8}\|u_2\|_2^2 + \frac{\delta_1}{8}\|u_1\|_2^2 + \frac{\delta_1}{4}\|u_1\|_p^p \\
&\qquad + C + C_1 \left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1}\right) \\
&\leq \frac{\delta}{4}\|v_2\|_2^2 + \frac{\delta_1 3}{8}\|u_1\|_p^p + C \\
&\qquad + C_1 \left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1}\right).
\end{aligned}$$

Hence, in total we obtain

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\|v_1\|_2^2 + \|v_2\|_2^2) \\
&\leq -d\|\nabla v_1\|_2^2 - \frac{\delta_1}{2}\|u_1\|_p^p - \delta\|v_2\|_2^2 + \frac{\delta}{4}\|v_2\|_2^2 + \frac{\delta_1 3}{8}\|u_1\|_p^p \\
&\qquad + C + C_1 \left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1} + \|z_1(\theta_t\omega)\|_p^p\right) \\
&= -d\|\nabla v_1\|_2^2 - \frac{\delta_1}{8}\|u_1\|_p^p - \frac{3\delta}{4}\|v_2\|_2^2 \\
&\qquad + C + C_1 \left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_{p_1+1}^{p_1+1} + \|z_1(\theta_t\omega)\|_p^p\right) \\
&\leq -\frac{d}{2}\|\nabla v_1\|_2^2 - \frac{d}{2c}\|v_1\|_2^2 - \frac{3\delta}{4}\|v_2\|_2^2 + C + C_1 \left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_p^p\right)
\end{aligned} \tag{5.24}$$

and thus

$$\frac{\mathrm{d}}{\mathrm{d}t}(\|v_1\|_2^2 + \|v_2\|_2^2) \leq -C_2 \left(\|v_1\|_2^2 + \|v_2\|_2^2\right) + C + C_1 \left(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_p^p\right). \tag{5.25}$$

Applying Gronwall's inequality yields

$$\begin{aligned}
&\|v_1\|_2^2 + \|v_2\|_2^2 \\
&\leq \left(\|v_1^0\|_2^2 + \|v_2^0\|_2^2\right)\exp(-C_2 t) + C_3\left(1 - \exp(-C_2 t)\right) \\
&\qquad + C_1 \int_0^t \exp(-C_2(t-s)) \left(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p\right) \, \mathrm{d}s \\
&\leq \left(\|v_1^0\|_2^2 + \|v_2^0\|_2^2\right)\exp(-C_2 t) + C_3 \\
&\qquad + C_1 \int_0^t \exp(-C_2(t-s)) \left(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p\right) \, \mathrm{d}s.
\end{aligned} \tag{5.26}$$

We replace ω by $\theta_{-t}\omega$ (note the $\mathbb{P}$-preserving property of the MDS) and carry out a change of variables

$$\begin{aligned}
&\|v_1(t,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2+\|v_2(t,\theta_{-t}\omega,v_2^0(\theta_{-t}\omega))\|_2^2\\
&\le\left(\|v_1^0(\theta_{-t}\omega)\|_2^2+\|v_2^0(\theta_{-t}\omega)\|_2^2\right)\exp\left(-C_2t\right)+C_3\\
&\qquad+C_1\int_0^t\exp\left(-C_2(t-s)\right)\left(\|z_2(\theta_{s-t}\omega)\|_2^2+\|z_1(\theta_{s-t}\omega)\|_p^p\right)\,\mathrm{d}s\\
&\le\left(\|v_1^0(\theta_{-t}\omega)\|_2^2+\|v_2^0(\theta_{-t}\omega)\|_2^2\right)\exp\left(-C_2t\right)+C_3\\
&\qquad+C_1\int_{-t}^0\exp\left(C_2s\right)\left(\|z_2(\theta_s\omega)\|_2^2+\|z_1(\theta_s\omega)\|_p^p\right)\,\mathrm{d}s.
\end{aligned}$$

Thus for arbitrary $\mathcal{D}\in\mathcal{T}$ and $(v_1^0,v_2^0)(\theta_{-t}\omega)\in\mathcal{D}(\theta_{-t}\omega)$

$$\begin{aligned}
&\|\psi(t,\theta_{-t}\omega,(v_1^0,v_2^0)(\theta_{-t}\omega))\|_{\mathbb{H}}^2\\
&\le\left(\|v_1^0(\theta_{-t}\omega)\|_2^2+\|v_2^0(\theta_{-t}\omega)\|_2^2\right)\exp\left(-C_2t\right)+C_3\\
&\qquad+C_1\int_{-t}^0\exp\left(C_2s\right)\left(\|z_2(\theta_s\omega)\|_2^2+\|z_1(\theta_s\omega)\|_p^p\right)\,\mathrm{d}s.
\end{aligned}$$

Since $\mathcal{D}\in\mathcal{T}$ we have

$$\limsup_{t\to\infty}\left(\|v_1^0(\theta_{-t}\omega)\|_2^2+\|v_2^0(\theta_{-t}\omega)\|_2^2\right)\exp\left(-C_2t\right)=0.$$

Hence,

$$\begin{aligned}
&\limsup_{t\to\infty}\|\psi(t,\theta_{-t}\omega,(v_1^0,v_2^0)(\theta_{-t}\omega))\|_{\mathbb{H}}^2\\
&\le C_3+C_1\int_{-\infty}^0\exp\left(C_2s\right)\left(\|z_2(\theta_s\omega)\|_2^2+\|z_1(\theta_s\omega)\|_p^p\right)\,\mathrm{d}s\\
&=:\rho(\omega). \qquad (5.27)
\end{aligned}$$

Due to the temperedness of $\|z_1(\omega)\|_p^p$ for $p\ge1$ and $\|z_2(\omega)\|_2^2$, the improper integral above exists and $\rho(\omega)>0$ is a tempered constant. As described in Remark 4.16, we can define for some $\varepsilon>0$

$$\mathcal{B}(\omega)=B(0,\rho(\omega)+\varepsilon).$$

Then $\mathcal{B}=\{\mathcal{B}(\omega)\}_\omega\in\mathcal{T}$ is a $\mathcal{T}$-absorbing set for the RDS φ with finite absorption time $t_{\mathcal{T}}(\omega)=\sup_{\mathcal{D}\in\mathcal{T}}t_{\mathcal{D}}(\omega)$. □

Remark 5.15. The random radius $\rho(\omega)$ depends on the restrictions imposed on the non-linearity and the noise. These were heavily used in Lemma 5.14 in order to derive the expression (5.27) for $\rho(\omega)$.

In order to make use of Theorem 4.19 we have to show the existence of a *compact* $\mathcal{T}$-absorbing set. So far we have only shown the existence of a bounded absorbing set, which is, being a ball in an infinite-dimensional Hilbert space, not compact.

5.2.6 Compact absorbing set

The classical strategy to find a compact absorbing set in $L^2(D)$ for a reaction-diffusion equation is the following: Firstly, one needs to find an absorbing set in $L^2(D)$. Secondly, this set is used to find an absorbing set $\mathcal{B}$ in $H^1(D)$ and due to compact embedding $H^1(D) \subset\subset L^2(D)$ (see Theorem C.2), $\overline{\mathcal{B}}$ defines a compact absorbing set in $L^2(D)$.

In the given setting the construction of an absorbing set in $H^1(D)$ is more complicated as the regularizing effect of the Laplacian is missing in the second component of (5.18). That is, solutions with initial conditions in $L^2(D)$ will in general only belong to $L^2(D)$ and not to $H^1(D)$.

To overcome this difficulty, we split the solution of the second component into two parts: one, which is regular enough, in the sense that it belongs to $H^1(D)$ and another one, which asymptotically tends to zero in $L^2(D)$. This splitting method has been used by other authors in the context of partly dissipative systems as well, see for instance [Mar89, Wan09a]. Let us now explain the strategy for our setting in more detail. We consider the equations

$$\frac{\mathrm{d}v_2^1(t)}{\mathrm{d}t} = -\sigma(x)v_2^1(t) - g(x, v_1(t) + z_1(\theta_t\omega)), \quad v_2^1(0) = 0, \tag{5.28}$$

and

$$\frac{\mathrm{d}v_2^2}{\mathrm{d}t} = -\sigma(x)v_2^2, \quad v_2^2(0) = v_2^0, \tag{5.29}$$

and one can easily verify that $v_2 = v_2^1 + v_2^2$ solves (5.21). Note at this point that we associate the initial condition $v_2^0 \in L^2(D)$ to the second part. Now, let $\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2) \in \mathcal{T}$ be arbitrary and $v^0 = (v_1^0, v_2^0) \in \mathcal{D}$. Then

$$\begin{aligned}\psi(t, \theta_{-t}\omega, v^0(\theta_{-t}\omega)) &= \begin{pmatrix} v_1(t, \theta_{-t}\omega, v_1^0(\theta_{-t}\omega)) \\ v_2(t, \theta_{-t}\omega, v_2^0(\theta_{-t}\omega)) \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} v_1(t, \theta_{-t}\omega, v_1^0(\theta_{-t}\omega)) \\ v_2^1(t, \theta_{-t}\omega, 0) \end{pmatrix}}_{=:\psi_1(t,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))} + \underbrace{\begin{pmatrix} 0 \\ v_2^2(t, \theta_{-t}\omega, v_2^0(\theta_{-t}\omega)) \end{pmatrix}}_{=:\psi_2(t,\theta_{-t}\omega,v_2^0(\theta_{-t}\omega))}.\end{aligned}$$

If we can show that for a certain $t^* \geq t_{\mathcal{D}}(\omega)$ there exist tempered random variables $\rho_1(\omega)$, $\rho_2(\omega)$ such that

$$\|v_1(t^*, \theta_{-t^*}\omega, v_1^0(\theta_{-t^*}\omega))\|_{H^1(D)} < \rho_1(\omega), \tag{5.30}$$

$$\|v_2^1(t^*, \theta_{-t^*}\omega, 0)\|_{H^1(D)} < \rho_2(\omega), \tag{5.31}$$

then by compact embedding $\overline{\psi_1(t^*, \theta_{-t^*}\omega, \mathcal{D}_1(\theta_{-t^*}\omega))}$ is a compact set in $\mathbb{H}$. If, furthermore,

$$\lim_{t\to\infty} \|v_2^2(t, \theta_{-t}\omega, v_2^0(\theta_{-t}\omega))\|_2 = 0, \tag{5.32}$$

then $\overline{\psi_2(t,\theta_{-t}\omega,\mathcal{D}_2(\theta_{-t}\omega))}$ can be regarded as a (random) bounded perturbation and $\overline{\psi(t,\theta_{-t}\omega,\mathcal{D}(\theta_{-t}\omega))}$ is compact in $\mathbb{H}$ as well, see [Tem12, Theorem 2.1]. Then,

$$\overline{\psi(t^*,\theta_{-t^*}\omega,\mathcal{B}(\theta_{-t^*}\omega))} \tag{5.33}$$

is a compact absorbing set for the RDS ψ. We will now prove the necessary estimates (5.30)-(5.32).

Lemma 5.16. *Let Assumptions 5.2 and 5.4 hold. Let $\mathcal{D}_2 \subset L^2(D)$ be tempered and $u_2^0 \in \mathcal{D}_2$. Then*

$$\lim_{t\to\infty} \|v_2^2(t,\theta_{-t}\omega, v_2^0(\theta_{-t}\omega))\|_2^2 = 0.$$

Proof. The solution to (5.29) is given by

$$v_2^2(t) = v_2^0 \exp\left(-\sigma(x)t\right)$$

and thus

$$\begin{aligned}\lim_{t\to\infty} \|v_2^2(t,\theta_{-t}\omega, v_2^0(\theta_{-t}\omega))\|_2^2 &= \lim_{t\to\infty} \left\|v_2^0(\theta_{-t}\omega)\exp\left(-\sigma(x)t\right)\right\|_2^2 \\ &\leq \lim_{t\to\infty} \|v_2^0(\theta_{-t}\omega)\|_2^2 \exp\left(-\delta t\right) \\ &\leq \lim_{t\to\infty} \left(\|u_2^0(\theta_{-t}\omega)\|_2^2 + \|z_2(\theta_{-t}\omega)\|_2^2\right)\exp\left(-\delta t\right) = 0,\end{aligned}$$

as $u_2^0 \in \mathcal{D}_2$ and $\|z_2(\omega)\|_2^2$ is a tempered random variable. □

We now prove boundedness of v_1 and v_2^1 in $H^1(D)$. Therefore we need some auxiliary estimates. First, let us derive uniform estimates for $u_1 \in L^p(D)$ and for $v_1 \in H^1(D)$.

Lemma 5.17. *Let Assumptions 5.2 and 5.4 hold. Let $\mathcal{D}_1 \subset L^2(D)$ be tempered and $u_1^0 \in \mathcal{D}_1$. Assume $t \geq 0$, $r > 0$, then*

$$\begin{aligned}\int_t^{t+r} \|u_1(s,\omega,u_1^0(\omega))\|_p^p \,\mathrm{d}s \leq Cr + C_1 \int_t^{t+r} \left(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p\right)\mathrm{d}s \\ + \|v_1(t,\omega,v_1^0(\omega))\|_2^2 + \|v_2(t,\omega,v_2^0(\omega))\|_2^2,\end{aligned} \tag{5.34}$$

$$\begin{aligned}\int_t^{t+r} \|\nabla v_1(s,\omega,v_1^0(\omega))\|_2^2 \,\mathrm{d}s \leq Cr + C_1 \int_t^{t+r} \left(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p\right)\mathrm{d}s \\ + \|v_1(t,\omega,v_1^0(\omega))\|_2^2 + \|v_2(t,\omega,v_2^0(\omega))\|_2^2,\end{aligned} \tag{5.35}$$

where C, C_1 are deterministic constants.

Proof. From (5.24) we can derive

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(\|v_1\|_2^2+\|v_2\|_2^2)\\
&\le -d\|\nabla v_1\|_2^2-\frac{\delta_1}{4}\|u_1\|_p^p+C+C_1\left(\|z_2(\theta_t\omega)\|_2^2+\|z_1(\theta_t\omega)\|_p^p\right),
\end{aligned}$$

and thus by integration

$$\begin{aligned}
&d\int_t^{t+r}\|\nabla v_1(s,\omega,v_1^0(\omega))\|_2^2\ \mathrm{d}s+\frac{\delta_1}{4}\int_t^{t+r}\|u_1(s,\omega,u_1^0(\omega))\|_p^p\ \mathrm{d}s\\
&\le Cr+C_1\int_t^{t+r}\left(\|z_2(\theta_s\omega)\|_2^2+\|z_1(\theta_s\omega)\|_p^p\right)\mathrm{d}s\\
&\qquad+\|v_1(t,\omega,v_1^0(\omega))\|_2^2+\|v_2(t,\omega,v_2^0(\omega))\|_2^2.
\end{aligned}$$

The two statements of the lemma follow directly from this estimate. □

Lemma 5.18. *Let Assumptions 5.2 and 5.4 hold. Let $\mathcal{D}_1\subset L^2(D)$ be tempered and $u_1^0\in\mathcal{D}_1$. Assume $t\ge r$, then*

$$\begin{aligned}
&\int_t^{t+r}\|u_1(s,\omega,u_1^0(\omega))\|_{2p-2}^{2p-2}\ \mathrm{d}s\\
&\le C_6r+\int_{t-r}^{t+r}C_2\|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p}+C_3\|z_2(\theta_s\omega)\|_2^2+C_4\|v_2(s,\omega,v_2^0(\omega))\|_2^2\ \mathrm{d}s\\
&\quad+C_5\|v_1(t-r,\omega,v_1^0(\omega))\|_2^2+C_5\|v_2(t-r,\omega,v_2^0(\omega))\|_2^2, \qquad (5.36)
\end{aligned}$$

where C_2,C_3,C_4,C_5,C_6 are deterministic constants.

Proof. Multiplying equation (5.20) by $|v_1|^{p-2}v_1$ and integrating over D yields

$$\begin{aligned}
&\frac{1}{p}\frac{\mathrm{d}}{\mathrm{d}t}\int_D|v_1|^p\ \mathrm{d}x\\
&=d\int_D\Delta v_1(t)|v_1|^{p-2}v_1\ \mathrm{d}x-\int_D h(x,v_1(t)+z_1(\theta_t\omega))|v_1|^{p-2}v_1\ \mathrm{d}x\\
&\qquad-\int_D f(x,v_1(t)+z_1(\theta_t\omega),v_2(t)+z_2(\theta_t\omega))|v_1|^{p-2}v_1\ \mathrm{d}x\\
&=-d(p-1)\int_D|\nabla v_1|^2|v_1|^{p-2}\ \mathrm{d}x-\int_D h(x,v_1(t)+z_1(\theta_t\omega))|v_1|^{p-2}v_1\ \mathrm{d}x\\
&\qquad-\int_D f(x,v_1(t)+z_1(\theta_t\omega),v_2(t)+z_2(\theta_t\omega))|v_1|^{p-2}v_1\ \mathrm{d}x\\
&\le-\int_D\left(\frac{\delta_1}{2^p}|v_1|^p-C-C_1(|z_1(\theta_t\omega)|^2+|z_1(\theta_t\omega)|^p)\right)|v_1|^{p-2}\ \mathrm{d}x\\
&\qquad+\int_D|f(x,v_1(t)+z_1(\theta_t\omega),v_2(t)+z_2(\theta_t\omega))||v_1|^{p-2}v_1\ \mathrm{d}x,
\end{aligned}$$

where we have used the inequality

$$h(x, v_1 + z_1)v_1 \geq \frac{\delta_1}{2^p}|v_1|^p - C - C_1(|z_1|^2 + |z_1|^p),$$

which can be proved by using conditions (5.4) and (5.9)

$$\begin{aligned}
h(x, v_1 + z_1)v_1 &= h(x, v_1 + z_1)(v_1 + z_1) - h(x, v_1 + z_1)z_1 \\
&\geq \delta_1|v_1 + z_1|^p - \delta_3 - |h(x, v_1 + z_1)||z_1| \\
&\geq \delta_1|v_1 + z_1|^p - \delta_3 - (\delta_8 + \delta_8|v_1 + z_1|^{p-1})|z_1| \\
&\geq \delta_1|v_1 + z_1|^p - C - C_1|z_1|^2 - \delta_1/2|v_1 + z_1|^p - C_2|z_1|^p \\
&= \frac{\delta_1}{2}|v_1 + z_1|^p - C - C_1(|z_1|^2 + |z_1|^p) \\
&\geq \frac{\delta_1}{2}||v_1| - |z_1||^p - C - C_1(|z_1|^2 + |z_1|^p) \\
&\geq \frac{\delta_1}{2^p}|v_1|^p - C - C_1(|z_1|^2 + |z_1|^p).
\end{aligned}$$

We compute further, using condition (5.5) and the relations $p-1, p-2, p_1+p-1 < 2p-2$

$$\begin{aligned}
&\frac{1}{p}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |v_1|^p \,\mathrm{d}x \\
&\leq -\int_D \frac{\delta_1}{2^p}|v_1|^{2p-2}\,\mathrm{d}x + C\int_D |v_1|^{p-2}\,\mathrm{d}x \\
&\quad + C_1\int_D (|z_1(\theta_t\omega)|^2 + |z_1(\theta_t\omega)|^p)|v_1|^{p-2}\,\mathrm{d}x \\
&\quad + \int_D \delta_4(1 + |v_1 + z_1(\theta_t\omega)|^{p_1} + |v_2 + z_2(\theta_t\omega)|)|v_1|^{p-2}v_1\,\mathrm{d}x \\
&\leq -\int_D \frac{\delta_1}{2^p}|v_1|^{2p-2}\,\mathrm{d}x + C\int_D |v_1|^{p-2}\,\mathrm{d}x + C_1\int_D |v_1|^{p-1}\,\mathrm{d}x \\
&\quad + C_2\int_D (|z_1(\theta_t\omega)|^{2p-2} + |z_1(\theta_t\omega)|^{p^2-p})\,\mathrm{d}x \\
&\quad + \int_D \delta_4\left(|v_1|^{p-1} + C_3\left(|v_1|^{p_1+p-1} + |z_1(\theta_t\omega)|^{p_1}|v_1|^{p-1} + |v_2||v_1|^{p-1} + |z_2(\theta_t\omega)||v_1|^{p-1}\right)\right)\,\mathrm{d}x \\
&\leq -\int_D \frac{\delta_1}{2^p}|v_1|^{2p-2}\,\mathrm{d}x + \frac{\delta_1}{2^p 4}\int_D |v_1|^{2p-2}\,\mathrm{d}x + C_6 \\
&\quad + C_2\int_D (|z_1(\theta_t\omega)|^{2p-2} + |z_1(\theta_t\omega)|^{p^2-p})\,\mathrm{d}x \\
&\quad + \int_D C_3(|z_1(\theta_t\omega)|^{p_1}|v_1|^{p-1} + |v_2||v_1|^{p-1} + |z_2(\theta_t\omega)||v_1|^{p-1})\,\mathrm{d}x.
\end{aligned}$$

Hence we have

$$
\begin{aligned}
&\frac{1}{p}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |v_1|^p \ \mathrm{d}x + \int_D \frac{3}{4}\frac{\delta_1}{2^p}|v_1|^{2p-2} \ \mathrm{d}x \\
&\le C_6 + C_2\int_D (|z_1(\theta_t\omega)|^{2p-2} + |z_1(\theta_t\omega)|^{p^2-p}) \ \mathrm{d}x \\
&\quad + \int_D C_3(|z_1(\theta_t\omega)|^{p_1} + |v_2| + |z_2(\theta_t\omega)|)|v_1|^{p-1} \ \mathrm{d}x \\
&\le C_6 + C_2\int_D (|z_1(\theta_t\omega)|^{2p-2} + |z_1(\theta_t\omega)|^{p^2-p}) \ \mathrm{d}x + \int_D \frac{1}{4}\frac{\delta_1}{2^p}|v_1|^{2p-2} \ \mathrm{d}x \\
&\quad + \int_D C_3(|z_1(\theta_t\omega)|^{p_1} + |v_2| + |z_2(\theta_t\omega)|)^2 \ \mathrm{d}x,
\end{aligned}
$$

and thus

$$
\begin{aligned}
&\frac{1}{p}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |v_1|^p \ \mathrm{d}x + \int_D \frac{1}{2}\frac{\delta_1}{2^p}|v_1|^{2p-2} \ \mathrm{d}x \\
&\le C_6 + C_2\int_D (|z_1(\theta_t\omega)|^{2p-2} + |z_1(\theta_t\omega)|^{p^2-p}) \ \mathrm{d}x \\
&\quad + \int_D C_3(|z_1(\theta_t\omega)|^{2p_1} + |v_2(t)|^2 + |z_2(\theta_t\omega)|^2) \ \mathrm{d}x. \qquad (5.37)
\end{aligned}
$$

We arrive at the following inequality

$$
\frac{1}{p}\frac{\mathrm{d}}{\mathrm{d}t}\|v_1\|_p^p + \frac{\delta_1}{2^{p+1}}\|v_1\|_{2p-2}^{2p-2} \le C_6 + C_2\|z_1(\theta_t\omega)\|_{p^2-p}^{p^2-p} + C_3\|z_2(\theta_t\omega)\|_2^2 + C_3\|v_2\|_2^2, \quad (5.38)
$$

and thus

$$
\frac{\mathrm{d}}{\mathrm{d}t}\|v_1\|_p^p \le C_6 + C_2\|z_1(\theta_t\omega)\|_{p^2-p}^{p^2-p} + C_3\|z_2(\theta_t\omega)\|_2^2 + C_3\|v_2\|_2^2 - \frac{\delta_1}{2^{p+1}}\|v_1\|_p^p. \quad (5.39)
$$

With (5.34) we have

$$
\begin{aligned}
\int_t^{t+r} \|v_1(s,\omega,v_1^0(\omega))\|_p^p \ \mathrm{d}s &= \int_t^{t+r} \|u_1(s,\omega,v_1^0(\omega)) - z_1(\theta_s\omega)\|_p^p \ \mathrm{d}s \\
&\le Cr + C_1\int_t^{t+r} \left(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p\right) \mathrm{d}s \\
&\quad + C_2\|v_1(t,\omega,v_1^0(\omega))\|_2^2 + C_2\|v_2(t,\omega,v_2^0(\omega))\|_2^2.
\end{aligned}
$$

Thus by applying the uniform Gronwall Lemma to (5.39) we have

$$
\begin{aligned}
&\|v_1(t+r,\omega,v_1^0(\omega))\|_p^p \\
&\le rC_6 + \int_t^{t+r} C_2\|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p} + C_3\|z_2(\theta_s\omega)\|_2^2 + C_4\|v_2(s,\omega,v_2^0(\omega))\|_2^2 \ \mathrm{d}s \\
&\quad + C_5\|v_1(t,\omega,v_1^0(\omega))\|_2^2 + C_5\|v_2(t,\omega,v_2^0(\omega))\|_2^2. \qquad (5.40)
\end{aligned}
$$

Now integrating (5.38) between t and $t+r$ yields

$$\begin{aligned}
&\int_t^{t+r} \|v_1(s,\omega,v_1(\omega))\|_{2p-2}^{2p-2}\,\mathrm{d}s \\
&\leq C_6 r + \int_t^{t+r} C_2\|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p} + C_3\|z_2(\theta_s\omega)\|_2^2 + C_3\|v_2(s,\omega,v_2^0(\omega))\|_2^2\,\mathrm{d}s \\
&\qquad + C\|v_1(t,\omega,v_1^0(\omega))\|_p^p,
\end{aligned}$$

and thus for $t \geq r$ using (5.40)

$$\begin{aligned}
&\int_t^{t+r} \|v_1(s,\omega,v_1(\omega))\|_{2p-2}^{2p-2}\,\mathrm{d}s \\
&\leq C_6 r + \int_{t-r}^{t+r} C_2\|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p} + C_3\|z_2(\theta_s\omega)\|_2^2 + C_4\|v_2(s,\omega,v_2^0(\omega))\|_2^2\,\mathrm{d}s \\
&\qquad + C_5\|v_1(t-r,\omega,v_1^0(\omega))\|_2^2 + C_5\|v_2(t-r,\omega,v_2^0(\omega))\|_2^2.
\end{aligned}$$

In total this leads to

$$\begin{aligned}
&\int_t^{t+r} \|u_1(s,\omega,v_1(\omega))\|_{2p-2}^{2p-2}\,\mathrm{d}s \\
&\leq C_6 r + \int_{t-r}^{t+r} C_2\|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p} + C_3\|z_2(\theta_s\omega)\|_2^2 + C_4\|v_2(s,\omega,v_2^0(\omega))\|_2^2\,\mathrm{d}s \\
&\qquad + C_5\|v_1(t-r,\omega,v_1^0(\omega))\|_2^2 + C_5\|v_2(t-r,\omega,v_2^0(\omega))\|_2^2 \\
&\qquad + \int_t^{t+r} \|z_1(\theta_s\omega)\|_{2p-2}^{2p-2}\,\mathrm{d}s \\
&\leq C_6 r + \int_{t-r}^{t+r} C_2\|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p} + C_3\|z_2(\theta_s\omega)\|_2^2 + C_4\|v_2(s,\omega,v_2^0(\omega))\|_2^2\,\mathrm{d}s \\
&\qquad + C_5\|v_1(t-r,\omega,v_1^0(\omega))\|_2^2 + C_5\|v_2(t-r,\omega,v_2^0(\omega))\|_2^2,
\end{aligned}$$

and this finishes the proof. □

Remark 5.19. One can also use appropriate shifts within the integrals on the left hand sides in (5.34), (5.35), (5.36) to obtain simpler forms of the ω-dependent constants on the right hand side, see for instance [Wan09b, Lemma 4.3, 4.4]. More precisely, in case of (5.34) one can for instance obtain an estimate of the form

$$\int_t^{t+r} \|u_1(s,\theta_{-t-r}\omega,u_1^0(\theta_{-t-r}\omega))\|_p^p \leq c(1+\widetilde{\rho}(\omega)),$$

where $\tilde{\rho}(\omega)$ is a random constant. Nevertheless such estimates hold for every ω, independent of the shift that one inserts inside the integral on the left hand side. Without the appropriate shifts on the left hand sides, as in the lemmas above, the constants on the right hand sides depend on the shift.

Next, we are going to show the boundedness of v_1 in $H^1(D)$.

Lemma 5.20. *Let Assumptions 5.2 and 5.4 hold. Let $\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2) \in \mathcal{T}$ and $u^0 \in \mathcal{D}$. Assume $t \geq t_{\mathcal{D}}(\omega) + 2r$ for some $r > 0$ then*

$$\|\nabla v_1(t, \theta_{-t}\omega, v_1^0(\theta_{-t}\omega))\|_2^2 \leq \rho_1(\omega), \tag{5.41}$$

where $\rho_1(\omega)$ is a tempered random variable.

Proof. We recall that v_1 satisfies equation (5.20) and thus we can compute

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla v_1\|_2^2 = \left\langle \frac{\mathrm{d}}{\mathrm{d}t}v_1, -\Delta v_1 \right\rangle \\
&= \langle d\Delta v_1 - h(x, v_1 + z_1(\theta_t\omega)) - f(x, v_1 + z_1(\theta_t\omega), v_2 + z_2(\theta_t\omega)), -\Delta v_1\rangle \\
&= -d\|\Delta v_1\|_2^2 + \langle h(x, v_1 + z_1(\theta_t\omega)), \Delta v_1\rangle + \langle f(x, v_1 + z_1(\theta_t\omega), v_2 + z_2(\theta_t\omega)), \Delta v_1\rangle \\
&\leq -d\|\Delta v_1\|_2^2 + \int_D \delta_8(1 + |u_1|^{p-1})|\Delta v_1| \ \mathrm{d}x + \int_D \delta_4(1 + |u_1|^{p_1} + |u_2|)|\Delta v_1| \ \mathrm{d}x \\
&\leq -d\|\Delta v_1\|_2^2 + C\int_D (2 + |u_1|^{p-1} + |u_1|^{p_1} + |u_2|)|\Delta v_1| \ \mathrm{d}x \\
&\leq -\frac{d}{2}\|\Delta v_1\|_2^2 + C\int_D (1 + |u_1|^{p-1} + |u_1|^{p_1} + |u_2|)^2 \ \mathrm{d}x \\
&\leq -\frac{d}{2}\|\Delta v_1\|_2^2 + C\int_D (1 + |u_1|^{2p-2} + |u_2|^2) \ \mathrm{d}x \\
&= -\frac{d}{2}\|\Delta v_1\|_2^2 + C_1 + C\|u_1\|_{2p-2}^{2p-2} + C\|u_2\|_2^2 \\
&\leq -\frac{dc}{2}\|\nabla v_1\|_2^2 + C_1 + C\|u_1\|_{2p-2}^{2p-2} + C\|u_2\|_2^2.
\end{aligned}$$

We want to apply the uniform Gronwall Lemma. Therefore, note

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\|\nabla v_1(t, \omega, v_1^0(\omega))\|_2^2}_{:=y(t)} \leq& \underbrace{-dc}_{:=g(t)}\|\nabla v_1(t, \omega, v_1^0(\omega))\|_2^2 \\
&+ \underbrace{C_1 + C\|u_1(t, \omega, u_1^0(\omega))\|_{2p-2}^{2p-2} + C\|u_2(t, \omega, u_2^0(\omega))\|_2^2}_{:=h(t)}.
\end{aligned}$$

We calculate

$$\int_t^{t+r} g(s) \ \mathrm{d}s \leq 0,$$

and

$$\begin{aligned}
\int_t^{t+r} \|\nabla v_1(s, \omega, v_1^0(\omega))\|_2^2 \ \mathrm{d}s \leq Cr + C_1\int_t^{t+r} &\left(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p\right)\mathrm{d}s \\
&+ C_2\left(\|v_1(t, \omega, v_1^0(\omega))\|_2^2 + \|v_2(t, \omega, v_2^0(\omega))\|_2^2\right),
\end{aligned}$$

where we have applied Lemma 5.17. By Lemma 5.18 for $t \geq r$

$$\begin{aligned}
&\int_t^{t+r} \|u_1(s,\omega,u_1^0(\omega))\|_{2p-2}^{2p-2} \,\mathrm{d}s \\
&\leq C_6 r + \int_{t-r}^{t+r} C_2\|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p} + C_3\|z_2(\theta_s\omega)\|_2^2 + C_4\|u_2(s,\omega,v_2^0(\omega))\|_2^2 \,\mathrm{d}s \\
&\qquad + C_5\|v_1(t-r,\omega,v_1^0(\omega))\|_2^2 + C_5\|v_2(t-r,\omega,v_2^0(\omega))\|_2^2.
\end{aligned}$$

Now, the uniform Gronwall Lemma yields for $t \geq r$

$$\begin{aligned}
&\|\nabla v_1(t+r,\omega,v_1^0(\omega))\|_2^2 \\
&\leq C + C_1 \int_t^{t+r} \left(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p\right) \mathrm{d}s \\
&\quad + C_2\left(\|v_1(t,\omega,v_1^0(\omega))\|_2^2 + \|v_2(t,\omega,v_2^0(\omega))\|_2^2\right) \\
&\quad + C_3 \int_{t-r}^{t+r} \|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p} + \|z_2(\theta_s\omega)\|_2^2 + \|u_2(s,\omega,v_2^0(\omega))\|_2^2 \,\mathrm{d}s \\
&\quad + C_4\left(\|v_1(t-r,\omega,v_1^0(\omega))\|_2^2 + \|v_2(t-r,\omega,v_2^0(\omega))\|_2^2\right) \\
&\quad + C_5 \int_t^{t+r} \|u_2(s,\omega,u_2^0(\omega))\|_2^2 \,\mathrm{d}s \\
&\leq C + C_1 \int_{t-r}^{t+r} \|u_2(s,\omega,u_2^0(\omega))\|_2^2 \,\mathrm{d}s + C_2 \int_{t-r}^{t+r} \|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p} + \|z_2(\theta_s\omega)\|_2^2 \,\mathrm{d}s \\
&\quad + C_3\left(\|v_1(t,\omega,v_1^0(\omega))\|_2^2 + \|v_2(t,\omega,v_2^0(\omega))\|_2^2\right. \\
&\qquad \left. + \|v_1(t-r,\omega,v_1^0(\omega))\|_2^2 + \|v_2(t-r,\omega,v_2^0(\omega))\|_2^2\right).
\end{aligned}$$

That is, for $t \geq 0$ we have

$$\begin{aligned}
&\|\nabla v_1(t+2r,\omega,v_1^0(\omega))\|_2^2 \\
&\leq C + C_1 \int_t^{t+2r} \|v_2(s,\omega,u_2^0(\omega))\|_2^2 \,\mathrm{d}s \\
&\quad + C_2 \int_t^{t+2r} \|z_1(\theta_s\omega)\|_{p^2-p}^{p^2-p} + \|z_2(\theta_s\omega)\|_2^2 \,\mathrm{d}s \\
&\quad + C_3\left(\|v_1(t+r,\omega,v_1^0(\omega))\|_2^2 + \|v_2(t+r,\omega,v_2^0(\omega))\|_2^2 \ldots\right. \\
&\qquad \left. + \|v_1(t,\omega,v_1^0(\omega))\|_2^2 + \|v_2(t,\omega,v_2^0(\omega))\|_2^2\right).
\end{aligned}$$

Let us recall that our goal is to find a $t^* \geq t_{\mathcal{D}}(\omega)$ such that (5.30) holds. Now assume that $t \geq t_{\mathcal{D}}(\omega)$. We replace ω by $\theta_{-t-2r}\omega$ (again note the $\mathbb{P}$-preserving property of the MDS), then

$$
\begin{aligned}
&\|\nabla v_1(t+2r,\theta_{-t-2r}\omega,v_1^0(\theta_{-t-2r}\omega))\|_2^2\\
&\leq C + C_1\int_t^{t+2r}\|v_2(s,\theta_{-t-2r}\omega,u_2^0(\theta_{-t-2r}\omega))\|_2^2\,\mathrm{d}s\\
&\quad + C_2\int_t^{t+2r}\|z_1(\theta_{s-t-2r}\omega)\|_{p^2-p}^{p^2-p}+\|z_2(\theta_{s-t-2r}\omega)\|_2^2\,\mathrm{d}s\\
&\quad + C_3\left(\|v_1(t+r,\theta_{-t-2r}\omega,v_1^0(\theta_{-t-2r}\omega))\|_2^2+\|v_2(t+r,\theta_{-t-2r}\omega,v_2^0(\theta_{-t-2r}\omega))\|_2^2\right.\\
&\qquad\left.+\|v_1(t,\theta_{-t-2r}\omega,v_1^0(\theta_{-t-2r}\omega))\|_2^2+\|v_2(t,\theta_{-t-2r}\omega,v_2^0(\theta_{-t-2r}\omega))\|_2^2\right).
\end{aligned}
$$

As $t \geq t_{\mathcal{D}}(\omega)$ we know by the absorption property that there exists a $\tilde{\rho}(\omega)$ such that

$$\|v_1(t,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2 \leq \tilde{\rho}(\omega),$$

and thus replacing ω by $\theta_{-2r}\omega$

$$\|v_1(t,\theta_{-t-2r}\omega,v_1^0(\theta_{-t-2r}\omega))\|_2^2 \leq \tilde{\rho}(\theta_{-2r}\omega).$$

Similarly, we know that

$$\|v_1(t+r,\theta_{-t-r}\omega,v_1^0(\theta_{-t-r}\omega))\|_2^2 \leq \tilde{\rho}(\theta_{-r}\omega),$$

and thus by replacing ω by $\theta_{-r}\omega$

$$\|v_1(t+r,\theta_{-t-2r}\omega,v_1^0(\theta_{-t-2r}\omega))\|_2^2 \leq \tilde{\rho}(\theta_{-2r}\omega).$$

The same arguments hold for v_2. Furthermore, as $t \geq t_{\mathcal{D}}(\omega)$ and we know from Lemma 5.14 that there exists a tempered random variable $\hat{\rho}(\omega)$ such that for $s \in (t,t+2r)$

$$\|v_2(s,\theta_{-s}\omega,u_2^0(\theta_{-s}\omega))\|_2^2 \leq \hat{\rho}(\omega)$$

and thus

$$
\begin{aligned}
&\int_t^{t+2r}\|v_2(s,\theta_{-t-2r}\omega,u_2^0(\theta_{-t-2r}\omega))\|_2^2\mathrm{d}s\\
&\leq \int_t^{t+2r}\hat{\rho}(\theta_{s-t-2r}\omega)\,\mathrm{d}s = \int_0^{2r}\hat{\rho}(\theta_{\tau-2r}\omega)\,\mathrm{d}\tau = \int_{-2r}^{0}\hat{\rho}(\theta_y\omega)\mathrm{d}y.
\end{aligned}
$$

With similar substitutions in the integral over $\|z_1(\theta_{s-t-2r}\omega)\|_{p^2-p}^{p^2-p}$ and $\|z_2(\theta_{s-t-2r}\omega)\|_2^2$ we arrive at

$$
\begin{aligned}
&\|\nabla v_1(t+2r,\theta_{-t-2r}\omega,v_1^0(\theta_{-t-2r}\omega))\|_2^2\\
&\leq C + C_1\int_{-2r}^{0}\hat{\rho}(\theta_y\omega)\mathrm{d}y + C_2\int_{-2r}^{0}\|z_1(\theta_y\omega)\|_{p^2-p}^{p^2-p}+\|z_2(\theta_y\omega)\|_2^2\,\mathrm{d}y + C_3\tilde{\rho}(\theta_{-2r}\omega),
\end{aligned}
$$

where the right hand side is independent of t. Due to the temperedness of all terms involved, they can be combined into one tempered random variable $\rho_1(\omega)$ such that for $t \geq t_{\mathcal{D}}(\omega) + 2r =: t^*$ we have

$$\|\nabla v_1(t, \theta_{-t}\omega, v_1^0(\theta_{-t}\omega))\|_2^2 \leq \rho_1(\omega),$$

this concludes the proof. □

We are now able to prove the boundedness of the first term of v_2 in $H^1(D)$.

Lemma 5.21. *Let Assumptions 5.2 and 5.4 hold. Let $\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2) \in \mathcal{T}$ and $u^0 \in \mathcal{D}$. Assume $t \geq t_{\mathcal{D}}(\omega) + 2r$ for some $r > 0$. Then we have*

$$\|\nabla v_2^1(t, \theta_{-t}\omega, 0)\|_2^2 \leq \rho_2(\omega), \tag{5.42}$$

where $\rho_2(\omega)$ is a tempered random variable.

Proof. Remember that v_2^1 satisfies the equation (5.28) and thus

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla v_2^1\|_2^2 &= \langle \frac{\mathrm{d}}{\mathrm{d}t} v_2^1, -\Delta v_2^1\rangle \\ &= \langle -\sigma(x) v_2^1 - g(x, v_1 + z_1), -\Delta v_2^1\rangle \\ &= \underbrace{\langle \sigma(x) v_2^1, \Delta v_2^1\rangle}_{=:L_1} + \underbrace{\langle g(x, v_1 + z_1), \Delta v_2^1\rangle}_{=:L_2}.\end{aligned}$$

We estimate L_1 and L_2 separately

$$\begin{aligned}L_1 = \int_D \sigma(x) v_2^1 \Delta v_2^1 \mathrm{d}x &= -\int_D \nabla(\sigma(x) v_2^1) \cdot \nabla v_2^1 \mathrm{d}x \\ &\leq -\delta\|\nabla v_2^1\|_2^2 - \int_D \nabla\sigma(x) v_2^1 \cdot \nabla v_2^1 \mathrm{d}x,\end{aligned}$$

and

$$\begin{aligned}L_2 &= \int_D g(x, v_1 + z_1)\Delta v_2^1 \,\mathrm{d}x = -\int_D \nabla g(x, v_1 + z_1) \cdot \nabla v_2^1 \,\mathrm{d}x \\ &= -\int_D (\nabla g(x, v_1 + z_1) + \partial_\xi g(x, v_1 + z_1)\nabla(v_1 + z_1)) \cdot \nabla v_2^1 \,\mathrm{d}x,\end{aligned}$$

where in the last equation the gradient is to be understood as

$$\nabla g(x, v_1 + z_1) = (\partial_{x_1} g(x, v_1 + z_1), ..., \partial_{x_n} g(x, v_1 + z_1))^\top.$$

Hence,

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla v_2^1\|_2^2 + 2\delta\|\nabla v_2^1\|_2^2 \\ &\leq 2\int_D \left|\nabla\sigma(x) v_2^1 + \nabla g(x, v_1 + z_1) + \partial_\xi g(x, v_1 + z_1)\nabla(v_1 + z_1)\right| |\nabla v_2^1| \,\mathrm{d}x \\ &\leq \frac{1}{\delta}\int_D \left|\nabla\sigma(x) v_2^1 + \nabla g(x, v_1 + z_1) + \partial_\xi g(x, v_1 + z_1)\nabla(v_1 + z_1)\right|^2 \,\mathrm{d}x + \delta\|\nabla v_2^1\|_2^2,\end{aligned}$$

and further with (5.7)

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla v_2^1\|_2^2 + \delta\|\nabla v_2^1\|_2^2 \\
&\leq \frac{1}{\delta}\int_D \sum_{i=1}^n \left(|\partial_{x_i}\sigma(x)v_2^1| + |\partial_{x_i}g(x, v_1+z_1)| + |\partial_\xi g(x, v_1+z_1)\partial_{x_i}(v_1+z_1)|\right)^2 \, \mathrm{d}x \\
&\leq \frac{1}{\delta}\int_D \sum_{i=1}^n \left(C|v_2^1| + \delta_5(1+|v_1+z_1|) + \delta_5|\partial_{x_i}(v_1+z_1)|\right)^2 \, \mathrm{d}x \\
&\leq \frac{2}{\delta}(C+\delta_5)^2 n\int_D \left(|v_2^1| + 1 + |v_1+z_1|\right)^2 \, \mathrm{d}x + \frac{2\delta_5^2}{\delta}\int_D \sum_{i=1}^n |\partial_{x_i}(v_1+z_1)|^2 \, \mathrm{d}x \\
&= \frac{2}{\delta}(C+\delta_5)^2 n\int_D \left(|v_2^1| + 1 + |v_1+z_1|\right)^2 \, \mathrm{d}x + \frac{2\delta_5^2}{\delta}\|\nabla(v_1+z_1)\|_2^2 \\
&\leq C_1 + C_2(\|v_2^1\|_2^2 + \|v_1\|_2^2 + \|z_1\|_2^2) + C_3(\|\nabla v_1\|_2^2 + \|\nabla z_1\|_2^2),
\end{aligned}
$$

where $C := \max_{1\leq i\leq n}\max_{x\in\overline{D}}|\partial_{x_i}\sigma(x)|$. Next, we apply Gronwall's inequality while taking the initial condition into account and we obtain for $t \geq 0$

$$
\begin{aligned}
\|\nabla v_2^1\|_2^2 \leq \int_0^t &\left[C_1 + C_2(\|v_2^1\|_2^2 + \|v_1\|_2^2 + \|z_1\|_2^2) + C_3(\|\nabla v_1\|_2^2 + \|\nabla z_1\|_2^2)\right] \\
&\times \exp\left((s-t)\delta\right) \, \mathrm{d}s. \qquad (5.43)
\end{aligned}
$$

We have from (5.24) the following equation

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(\|v_1\|_2^2 + \|v_2\|_2^2) + M(\|v_1\|_2^2 + \|v_2\|_2^2) + d\|\nabla v_1\|_2^2 \\
&\leq \hat{C} + \tilde{C}(\|z_2(\theta_t\omega)\|_2^2 + \|z_1(\theta_t\omega)\|_p^p), \qquad (5.44)
\end{aligned}
$$

where $M = \min\{d/c, \delta\}$ and certain constants $\hat{C}, \tilde{C}$. We multiply (5.44) by $\exp(Mt)$ and integrate between 0 and t

$$
\begin{aligned}
&\int_0^t \exp(Ms)\frac{\mathrm{d}}{\mathrm{d}s}(\|v_1\|_2^2 + \|v_2\|_2^2)\mathrm{d}s + M\int_0^t \exp(Ms)(\|v_1\|_2^2 + \|v_2\|_2^2)\mathrm{d}s \\
&\qquad + d\int_0^t \exp(Ms)\|\nabla v_1\|_2^2\mathrm{d}s \\
&\leq \int_0^t \hat{C}\exp(Ms)\mathrm{d}s + \tilde{C}\int_0^t \exp(Ms)(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p)\mathrm{d}s.
\end{aligned}
$$

This yields

$$
\begin{aligned}
&\int_0^t \exp(M(s-t))\|\nabla v_1(s,\omega,v_1^0(\omega))\|_2^2\mathrm{d}s \\
&\leq \frac{1}{d}\exp(-Mt)(\|v_1^0(\omega)\|_2^2 + \|v_2^0(\omega)\|_2^2) + \hat{C} \\
&\qquad + \tilde{C}\int_0^t \exp(M(s-t))(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p)\mathrm{d}s, \qquad (5.45)
\end{aligned}
$$

as well as

$$
\begin{aligned}
&\|v_1(t,\omega,v_1^0(\omega))\|_2^2+\|v_2(t,\omega,v_2^0(\omega))\|_2^2\\
&\le \left(\|v_1^0(\omega)\|_2^2+\|v_2^0(\omega)\|_2^2\right)\exp\left(-Mt\right)+\hat{C}\\
&\quad +\tilde{C}\int_0^t \exp\left(M(s-t)\right)\left(\|z_2(\theta_s\omega)\|_2^2+\|z_1(\theta_s\omega)\|_p^p\right)\,\mathrm{d}s.
\end{aligned}
$$

In particular, from the last estimate we obtain

$$
\begin{aligned}
&\int_0^{t_{\mathcal{D}}(\omega)}(\|v_1(s,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega)\|_2^2+\|v_2(s,\theta_{-t}\omega,v_2^0(\theta_{-t}\omega))\|_2^2)\exp(M(s-t))\mathrm{d}s\\
&\le \int_0^{t_{\mathcal{D}}(\omega)}\left(\|v_1^0(\theta_{-t}\omega)\|_2^2+\|v_2^0(\theta_{-t}\omega)\|_2^2\right)\exp\left(-Mt\right)\mathrm{d}s+\hat{C}\int_0^{t_{\mathcal{D}}(\omega)}\exp(M(s-t))\mathrm{d}s\\
&\quad +\tilde{C}\int_0^{t_{\mathcal{D}}(\omega)}\int_0^s \exp\left(M(\tau-t)\right)\left(\|z_2(\theta_{\tau-t}\omega)\|_2^2+\|z_1(\theta_{\tau-t}\omega)\|_p^p\right)\,\mathrm{d}\tau\mathrm{d}s\\
&\le \left(\|v_1^0(\theta_{-t}\omega)\|_2^2+\|v_2^0(\theta_{-t}\omega)\|_2^2\right)\exp\left(-Mt\right)t_{\mathcal{D}}(\omega)+\hat{C}\\
&\quad +\tilde{C}t_{\mathcal{D}}(\omega)\int_0^{t_{\mathcal{D}}(\omega)}\exp\left(M(\tau-t)\right)\left(\|z_2(\theta_{\tau-t}\omega)\|_2^2+\|z_1(\theta_{\tau-t}\omega)\|_p^p\right)\,\mathrm{d}\tau.
\end{aligned}
\tag{5.46}
$$

where we have replaced ω by $\theta_{-t}\omega$ after integrating and we have used that $t\ge t_{\mathcal{D}}(\omega)$. Now, replacing ω by $\theta_{-t}\omega$ in (5.43), noting that $\delta\ge M$ and assuming that $t\ge t_{\mathcal{D}}(\omega)$, we compute using (5.45)

$$
\begin{aligned}
&\|\nabla v_2^1(t,\theta_{-t}\omega,0)\|_2^2\\
&\le \frac{C_1}{\delta}+C_2\int_0^t\big[\|v_2^1(s,\theta_{-t}\omega,0)\|_2^2+\|v_1(s,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2+\|z_1(\theta_{s-t}\omega)\|_2^2\\
&\qquad +\|\nabla v_1(s,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2+\|\nabla z_1(\theta_{s-t}\omega)\|_2^2\big]\exp\left((s-t)M\right)\,\mathrm{d}s\\
&\le C_1+C_2\int_0^{t_{\mathcal{D}}(\omega)}\big[\|v_2^1(s,\theta_{-t}\omega,0)\|_2^2+\|v_1(s,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2\big]\exp\left((s-t)M\right)\mathrm{d}s\\
&\quad +C_2\int_{t_{\mathcal{D}}(\omega)}^t\big[\|v_2^1(s,\theta_{-t}\omega,0)\|_2^2+\|v_1(s,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2\big]\exp\left((s-t)M\right)\mathrm{d}s\\
&\quad +C_3\exp(-Mt)(\|v_1^0(\theta_{-t}\omega)\|_2^2+\|v_2^0(\theta_{-t}\omega)\|_2^2)+C_4\int_0^t\exp(M(s-t))\\
&\quad \times(\|z_2(\theta_{s-t}\omega)\|_2^2+\|z_1(\theta_{s-t}\omega)\|_p^p+\|z_1(\theta_{s-t}\omega)\|_2^2+\|\nabla z_1(\theta_{s-t}\omega)\|_2^2)\mathrm{d}s,
\end{aligned}
$$

and using (5.46) as well as the absorption property, this can be estimated further

$$\begin{aligned}
&\|\nabla v_2^1(t,\theta_{-t}\omega,0)\|_2^2\\
&\leq C_1 + C_2\left(\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2\right)\exp\left(-Mt\right)t_{\mathcal{D}}(\omega)\\
&\quad + C_5 t_{\mathcal{D}}(\omega)\int_0^{t_{\mathcal{D}}(\omega)}\exp\left(M(\tau-t)\right)\left(\|z_2(\theta_{\tau-t}\omega)\|_2^2 + \|z_1(\theta_{\tau-t}\omega)\|_p^p\right)\,\mathrm{d}\tau\\
&\quad + C_2\int_{t_{\mathcal{D}}(\omega)}^t \rho(\omega)\exp\left((s-t)M\right)\mathrm{d}s\\
&\quad + C_3\exp(-Mt)(\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2)\\
&\quad + C_4\int_{-\infty}^0 \exp(Ms)(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p + \|z_1(\theta_s\omega)\|_2^2 + \|\nabla z_1(\theta_s\omega)\|_2^2)\mathrm{d}s\\
&\leq C_1 + C_2(t_{\mathcal{D}}(\omega))\left(\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2\right)\exp\left(-Mt\right) + C_3\rho(\omega)\\
&\quad + C_4(t_{\mathcal{D}}(\omega))\int_{-\infty}^0 \exp(Ms)\\
&\quad \times(\|z_2(\theta_s\omega)\|_2^2 + \|z_1(\theta_s\omega)\|_p^p + \|z_1(\theta_s\omega)\|_2^2 + \|\nabla z_1(\theta_s\omega)\|_2^2)\mathrm{d}s.
\end{aligned}$$

Finally, since $\|z_2(\theta_s\omega)\|_2^2$, $\|z_1(\theta_s\omega)\|_p^p$, $\|z_1(\theta_s\omega)\|_2^2$, $\|\nabla z_1(\theta_s\omega)\|_2^2$ (see Lemma 5.10 and Remark 5.11) and $\|v_1^0(\theta_{-t}\omega)\|_2^2$, $\|v_2^0(\theta_{-t}\omega)\|_2^2$ (by assumption) are tempered random variables, we can combine the right hand side into one tempered random variable $\rho_2(\omega)$ and this concludes the proof. □

Theorem 5.22. *Let Assumptions 5.2 and 5.4 hold. The random dynamical system defined in Lemma 5.13 has a unique $\mathcal{T}$-random attractor $\mathcal{A}$.*

Proof. By the previous lemmas there exist a compact absorbing set given by (5.33) in $\mathcal{T}$ for the RDS ψ. Thus Theorem 4.19 guarantees the existence of a unique $\mathcal{T}$-random attractor. By conjugacy the existence of a unique $\mathcal{T}$-random attractor for φ follows. □

5.3 Stochastic partly dissipative systems with multiplicative noise

In this section we analyse the same partly dissipative system as before, however, this time we perturb it by *linear multiplicative noise.* Note that the general strategy to prove the existence of an attractor is similar to the case with additive noise.

Again, let $D \subset \mathbb{R}^n$ be a bounded open set with regular boundary. We consider the following partly dissipative system perturbed by the same multiplicative Stratonovich noise in both components

$$\begin{aligned}
\mathrm{d}u_1 &= (d\Delta u_1 - h(x,u_1) - f(x,u_1,u_2))\,\mathrm{d}t + u_1\circ\,\mathrm{d}B,\\
\mathrm{d}u_2 &= (-\sigma(x)u_2 - g(x,u_1))\,\mathrm{d}t + u_2\circ\,\mathrm{d}B,
\end{aligned}\tag{5.47}$$

where $u_{1,2} = u_{1,2}(t,x)$, $(t,x) \in [0,T] \times D$, $T > 0$, $(B(t))_{t\in\mathbb{R}}$ is a two-sided, real-valued Wiener process. The symbol $\circ$ indicates that the equation is understood in the *Stratonovich* sense. The system is equipped with initial conditions

$$u_1(0,x) = u_1^0(x) \in L^2(D), \quad u_2(0,x) = u_2^0(x) \in L^2(D),$$

and a homogeneous Dirichlet boundary condition for the first component. We make the same assumptions on the reaction terms as in the setting with additive noise, that is, throughout this section let Assumptions 5.2 hold. Again, we let A denote the realization of the Laplace operator with Dirichlet boundary conditions.

5.3.1 Associated RDS

We consider the MDS $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ associated to the two-sided Brownian motion, as defined in Example 4.2. Let $(t,\omega) \mapsto z(\theta_t\omega)$ denote the stationary *Ornstein-Uhlenbeck process* defined via

$$\mathrm{d}z = -z \;\mathrm{d}t + \mathrm{d}\omega. \tag{5.48}$$

Let $u(t) = (u_1(t), u_2(t))^\top$ be a solution of (5.47) and consider the following *Doss-Sussmann transformations*

$$v_1(t) := \exp(-z(\theta_t\omega))u_1(t), \tag{5.49}$$
$$v_2(t) := \exp(-z(\theta_t\omega))u_2(t). \tag{5.50}$$

Using the chain rule, we verify that v_1 and v_2 satisfy the following equations

$$\begin{aligned}\mathrm{d}v_1 &= \exp(-z(\theta_t\omega))\mathrm{d}u_1 - \exp(-z(\theta_t\omega))u_1 \circ \mathrm{d}z_1(\theta_t\omega)\\ &= d\Delta v_1 \mathrm{d}t + v_1 z(\theta_t\omega)\mathrm{d}t - \exp(-z(\theta_t\omega))(h(x, \exp(z(\theta_t\omega))v_1)\\ &\quad + f(x, \exp(z(\theta_t\omega))v_1, \exp(z(\theta_t\omega))v_2)) \;\mathrm{d}t,\end{aligned} \tag{5.51}$$

and

$$\begin{aligned}\mathrm{d}v_2 &= \exp(-z(\theta_t\omega))\mathrm{d}u_2 - \exp(-z(\theta_t\omega))u_2 \circ \mathrm{d}z(\theta_t\omega)\\ &= -\sigma(x)v_2\mathrm{d}t + v_2 z(\theta_t\omega) \;\mathrm{d}t - \exp(-z(\theta_t\omega))g(x, \exp(z(\theta_t\omega))v_1)\mathrm{d}t.\end{aligned} \tag{5.52}$$

Furthermore, the transformed initial conditions read

$$\begin{aligned}v_1^0(x) &:= v_1(0,x) = \exp(-z(\omega))u_1^0(x),\\ v_2^0(x) &:= v_2(0,x) = \exp(-z(\omega))u_2^0(x).\end{aligned}$$

Remark 5.23. The setting where the equations are perturbed multiplicatively by finitely many real-valued Wiener processes can be treated analogously by adapting the transformation (5.49)-(5.50) accordingly.

Lemma 5.24. *The map* $\psi : \mathbb{R}^+ \times \Omega \times \mathbb{H} \to \mathbb{H}$,

$$\psi(t,\omega,(v_1^0,v_2^0)) := (v_1(t,\omega,v_1^0), v_2(t,\omega,v_2^0))^\top,$$

defines a continuous RDS over the MDS $(\Omega,\mathcal{F},\mathbb{P},\theta)$.

Proof. The Lemma can be proved similar to Lemma 5.12. □

Lemma 5.25. *The map* $\varphi : \mathbb{R}^+ \times \Omega \times \mathbb{H} \to \mathbb{H}$,

$$\varphi(t,\omega,(u_1^0,u_2^0)) := (u_1(t,\omega,u_1^0), u_2(t,\omega,u_2^0))^\top,$$

defines a continuous RDS over the MDS $(\Omega,\mathcal{F},\mathbb{P},\theta)$.

Proof. The argument is similar to the proof of Lemma 5.13; here, with the transformation $T : \Omega \times \mathbb{H} \to \mathbb{H}$

$$T(\omega, v) := \exp(z(\omega))v.$$

□

As in the case for additive noise, we will first derive the existence of an absorbing set for the RDS ψ and subsequently we will use a splitting argument to construct a compact absorbing set.

5.3.2 Bounded absorbing set

Lemma 5.26. *Let Assumptions 5.2 hold. Then there exists a bounded* $\mathcal{T}$*-absorbing set* $\mathcal{B}$ *for the RDS* ψ.

Proof. Let (v_1, v_2) be a solution of (5.51)-(5.52), then we compute

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\|v_1\|_2^2 + \|v_2\|_2^2)\\
&= -d\|\nabla v_1\|_2^2 + z(\theta_t\omega)(\|v_1\|_2^2 + \|v_2\|_2^2) - \int_D \sigma(x)|v_2|^2\mathrm{d}x\\
&\quad \underbrace{- \exp(-z(\theta_t\omega))\int_D h(x, \exp(z(\theta_t\omega))v_1)v_1\mathrm{d}x}_{=:I_1}\\
&\quad \underbrace{- \exp(-z(\theta_t\omega))\int_D f(x, \exp(z(\theta_t\omega))v_1, \exp(z(\theta_t\omega))v_2))v_1\mathrm{d}x}_{=:I_2}\\
&\quad \underbrace{- \exp(-z(\theta_t\omega))\int_D g(x, \exp(z(\theta_t\omega))v_1)v_2\mathrm{d}x}_{=:I_3}\,.
\end{aligned}$$

By using the boundedness of $\sigma(x)$, we derive

$$-\int_D \sigma(x)v_2^2(t)\mathrm{d}x \le -\delta\|v_2\|_2^2.$$

Furthermore, making use of Assumptions 5.2 we compute

$$\begin{aligned}
I_1 =& -\exp(-2z(\theta_t\omega))\int_D h(x,\exp(z(\theta_t\omega))v_1)\exp(z(\theta_t\omega))v_1\mathrm{d}x\\
\le& \exp(-2z(\theta_t\omega))\int_D \delta_3 - \delta_1|v_1\exp(z(\theta_t\omega))|^p\mathrm{d}x\\
=& C_1\exp(-2z(\theta_t\omega)) - \delta_1\exp(-2z(\theta_t\omega))\|u_1\|_p^p,
\end{aligned}$$

$$\begin{aligned}
I_2 &\le \exp(-z(\theta_t\omega))\delta_4\int_D (1+|\exp(z(\theta_t\omega))v_1|^{p_1}+|\exp(z(\theta_t\omega))v_2|)v_1\mathrm{d}x\\
&= \exp(-2z(\theta_t\omega))\delta_4\int_D (1+|u_1|^{p_1}+|u_2|)|u_1|\mathrm{d}x,
\end{aligned}$$

and

$$\begin{aligned}
I_3 &= -\exp(-2z(\theta_t\omega))\int_D g(x,u_1)u_2\mathrm{d}x\\
&\le \delta_7\exp(-2z(\theta_t\omega))\int_D (1+|u_1|)|u_2|\mathrm{d}x.
\end{aligned}$$

Combining the last two estimates yields

$$\begin{aligned}
I_2+I_3 &\le \exp(-2z(\theta_t\omega))(\delta_4+\delta_7)\int_D |u_1|+|u_1|^{p_1+1}+(1+|u_1|)|u_2|\mathrm{d}x\\
&\le \exp(-2z(\theta_t\omega))(\delta_4+\delta_7)\int_D |u_1|+|u_1|^{p_1+1}+C(1+|u_1|)^2+\frac{\delta}{2(\delta_4+\delta_7)}|u_2|^2\mathrm{d}x\\
&\le \exp(-2z(\theta_t\omega))\frac{\delta}{2}\|u_2\|_2^2+\exp(-2z(\theta_t\omega))C\int_D 1+|u_1|^q\mathrm{d}x\\
&\le \frac{\delta}{2}\|v_2\|_2^2+\exp(-2z(\theta_t\omega))C+\exp(-2z(\theta_t\omega))\frac{\delta_1}{2}\|u_1\|_p^p,
\end{aligned}$$

where we have used equation (5.23) as in the case with additive noise. In total, we arrive at

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(\|v_1\|_2^2+\|v_2\|_2^2) \le& -2d\|\nabla v_1\|_2^2+2z(\theta_t\omega)(\|v_1\|_2^2+\|v_2\|_2^2)-2\delta\|v_2\|_2^2 \qquad (5.53)\\
&+C\exp(-2z(\theta_t\omega))-2\delta_1\exp(-2z(\theta_t\omega))\|u_1\|_p^p\\
&+\delta\|v_2\|_2^2+\exp(-2z(\theta_t\omega))\delta_1\|u_1\|_p^p\\
\le& -\frac{2d}{c}\|v_1\|_2^2+2z(\theta_t\omega)(\|v_1\|_2^2+\|v_2\|_2^2)-\delta\|v_2\|_2^2\\
&+C\exp(-2z(\theta_t\omega))-\delta_1\exp(-2z(\theta_t\omega))\|u_1\|_p^p\\
\le& (2z(\theta_t\omega)-C_1)(\|v_1\|_2^2+\|v_2\|_2^2)+C\exp(-2z(\theta_t\omega)),
\end{aligned}$$

where we have used Poincaré's inequality. Applying Gronwall's inequality we obtain

$$\begin{aligned}\|v_1(\omega)\|_2^2 + \|v_2(\omega)\|_2^2 \le{}& (\|v_1^0(\omega)\|_2^2 + \|v_2^0(\omega)\|_2^2)\exp\left(\int_0^t (2z(\theta_\tau\omega) - C_1)\mathrm{d}\tau\right)\\ &+ \int_0^t C\exp(-2z(\theta_s\omega))\exp\left(\int_s^t (2z(\theta_\tau\omega) - C_1)\mathrm{d}\tau\right)\mathrm{d}s\\ ={}& (\|v_1^0(\omega)\|_2^2 + \|v_2^0(\omega)\|_2^2)\exp\left(2\int_0^t z(\theta_\tau\omega)\mathrm{d}\tau - C_1 t\right)\\ &+ \int_0^t C\exp\left(-2z(\theta_s\omega) + 2\int_s^t z(\theta_\tau\omega)\mathrm{d}\tau - C_1(t-s)\right)\mathrm{d}s.\end{aligned}$$

We replace ω by $\theta_{-t}\omega$ and perform a change of variables

$$\begin{aligned}&\|v_1(t,\theta_{-t}\omega, v_1^0(\theta_{-t}\omega))\|_2^2 + \|v_2(t,\theta_{-t}\omega, v_2^0(\theta_{-t}\omega))\|_2^2\\ &\le (\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2)\exp\left(2\int_0^t z(\theta_{\tau-t}\omega)\mathrm{d}\tau - C_1 t\right)\\ &\quad + \int_0^t C\exp\left(-2z(\theta_{s-t}\omega) + 2\int_s^t z(\theta_{\tau-t}\omega)\mathrm{d}\tau - C_1(t-s)\right)\mathrm{d}s\\ &= (\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2)\exp\left(2\int_{-t}^0 z(\theta_\tau\omega)\mathrm{d}\tau - C_1 t\right)\\ &\quad + C\int_0^t \exp\left(-2z(\theta_{s-t}\omega) + 2\int_{s-t}^0 z(\theta_\tau\omega)\mathrm{d}\tau - C_1(t-s)\right)\mathrm{d}s\\ &= (\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2)\exp\left(2\int_{-t}^0 z(\theta_\tau\omega)\mathrm{d}\tau - C_1 t\right)\\ &\quad + C\int_{-t}^0 \exp\left(-2z(\theta_s\omega) + 2\int_s^0 z(\theta_\tau\omega)\mathrm{d}\tau + C_1 s\right)\mathrm{d}s.\end{aligned}$$

Now, let $\mathcal{D} \in \mathcal{T}$ be an arbitrary tempered set and $(v_1^0, v_2^0)(\theta_{-t}\omega) \in \mathcal{D}(\theta_{-t}\omega)$. By (4.3) we have

$$\lim_{t\to\infty}(\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2)\exp(-C_1/2t) = 0.$$

Furthermore, by Proposition 4.25 (ii)

$$\lim_{t\to\infty}\frac{1}{t}\int_{-t}^0 2z(\theta_\tau\omega)\mathrm{d}\tau = 2\lim_{t\to-\infty}\frac{1}{t}\int_0^t z(\theta_\tau\omega)\mathrm{d}\tau = 0,$$

thus for every $\omega \in \Omega$ and every $0 < \varepsilon < C_1/2$ there exists t_0 such that for all $t \ge t_0$

$$\left|\frac{1}{t}\int_{-t}^0 2z(\theta_\tau\omega)\mathrm{d}\tau\right| < C_1/2 - \varepsilon,$$

and thus

$$\exp\left(-C_1/2t+\int_{-t}^{0}2z(\theta_\tau\omega)\mathrm{d}\tau\right)=\exp\left(t\left(-C_1/2+\frac{1}{t}\int_{-t}^{0}2z(\theta_\tau\omega)\mathrm{d}\tau\right)\right)$$
$$<\exp(-\varepsilon t).$$

Hence,

$$\lim_{t\to\infty}\exp(-C_1/2t)\exp\left(\int_{-t}^{0}2z(\theta_\tau\omega)\mathrm{d}\tau\right)=0,$$

and we have in total

$$\lim_{t\to\infty}(\|v_1^0(\theta_{-t}\omega)\|_2^2+\|v_2^0(\theta_{-t}\omega)\|_2^2)\exp(-C_1t)\exp\left(\int_{-t}^{0}2z(\theta_\tau\omega)\mathrm{d}\tau\right)=0. \qquad (5.54)$$

Therefore, we obtain

$$\begin{aligned}&\lim_{t\to\infty}\|v_1(t,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2+\|v_2(t,\theta_{-t}\omega,v_2^0(\theta_{-t}\omega))\|_2^2\\&\le C\int_{-\infty}^{0}\exp\left(-2z(\theta_s\omega)+\int_s^0 2z(\theta_\tau\omega)\mathrm{d}\tau+C_1s\right)\mathrm{d}s=:\rho(\omega)^2.\end{aligned}$$

By Proposition 4.25 (i) and (iii) for any $\varepsilon>0$ and $t<0$ small enough

$$|z(\theta_{s+t}\omega)|<\varepsilon|s+t|,\qquad\int_t^0|z(\theta_\tau\omega)|-\mathbb{E}|z(\omega)|\mathrm{d}\tau<\varepsilon|t|.$$

Thus for any $c>0$ we have for $t<0$ small enough

$$\begin{aligned}&\exp(ct)\rho(\theta_t\omega)^2\\&\le\exp(ct)C\int_{-\infty}^{0}\exp\left(-2\varepsilon(t+s)+2\int_{s+t}^{0}|z(\theta_\tau\omega)|-\mathbb{E}|z(\omega)|\mathrm{d}\tau\right.\\&\qquad\left.+2\int_t^0|z(\theta_\tau\omega)|-\mathbb{E}|z(\omega)|\mathrm{d}\tau+s\mathbb{E}|z(\omega)|+C_1s\right)\mathrm{d}s\\&\le\exp(ct-6\varepsilon t)C\int_{-\infty}^{0}\exp\left(-4\varepsilon s+s\mathbb{E}|z(\omega)|+C_1s\right)\mathrm{d}s,\end{aligned}$$

that is, choosing $\varepsilon<\min\{c/6,(\mathbb{E}|z(\omega)|+C_1)/4\}$, the right hand side converges to zero for $t\to-\infty$. A similar argument can be made for $t\to\infty$. It follows, that $\rho(\omega)$ is a tempered random variable.

Hence, there exists $\eta>0$ such that

$$\mathcal{B}(\omega)=B(0,\rho(\omega)+\eta)$$

is a bounded, tempered $\mathcal{T}$-absorbing set for the RDS ψ. We denote the absorption time as $t_{\mathcal{D}}(\omega)$.

□

5.3.3 Compact absorbing set

We perform a splitting argument as in the case with additive noise. Consider the equations

$$\frac{\mathrm{d}v_2^1(t)}{\mathrm{d}t} = -\sigma(x)v_2^1(t) + v_2^1(t)z(\theta_t\omega) - \exp(-z(\theta_t\omega))g(x, \exp(z(\theta_t\omega))v_1), \tag{5.55}$$

with $v_2^1(0) = 0$ and

$$\frac{\mathrm{d}v_2^2(t)}{\mathrm{d}t} = -\sigma(x)v_2^2(t) + v_2^2(t)z(\theta_t\omega) = (z(\theta_t\omega) - \sigma(x))v_2^2(t), \tag{5.56}$$

with $v_2^2(0) = v_2^0$.

Lemma 5.27. *Let Assumptions 5.2 hold. Let $\mathcal{D} \subset L^2(D)$ be tempered. Then for $v_2^0 \in \mathcal{D}$ we have*

$$\lim_{t\to\infty} \|v_2^2(t, \theta_{-t}\omega, v_2^0(\theta_{-t}\omega))\|_2^2 = 0.$$

Proof. The solution to (5.56) is given by

$$v_2^2(t) = v_2^0 \exp\left(\int_0^t (z(\theta_s\omega) - \sigma(x))\mathrm{d}s\right)$$

and thus

$$\begin{aligned}
&\lim_{t\to\infty} \|v_2^2(t, \theta_{-t}\omega, v_2^0(\theta_{-t}\omega))\|_2^2 \\
&= \lim_{t\to\infty} \left\| v_2^0(\theta_{-t}\omega) \exp\left(\int_0^t z(\theta_{s-t}\omega)\mathrm{d}s - t\sigma(x)\right)\right\|_2^2 \\
&\leq \lim_{t\to\infty} \exp(-2\delta t) \exp\left(2\int_{-t}^0 z(\theta_s\omega)\mathrm{d}s\right) \|v_2^0(\theta_{-t}\omega)\|_2^2 \\
&= 0,
\end{aligned}$$

where for the last equality the same argument that was used to prove equation (5.54) can be applied. □

Lemma 5.28. *Let Assumptions 5.2 hold. Let $\mathcal{D} \in \mathcal{T}$ and $(v_1^0, v_2^0) \in \mathcal{D}$. Then, for $t \geq t_{\mathcal{D}}(\omega)$ and $r > 0$*

$$\int_t^{t+r} \|\nabla v_1(s, \theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_2^2 \mathrm{d}s \leq R(\omega) \tag{5.57}$$

and

$$\int_t^{t+r} \|v_1(s, \theta_{-t-r}, v_1^0(\theta_{-t-r}\omega))\|_p^p \mathrm{d}s \leq \hat{R}(\omega) \tag{5.58}$$

where $R(\omega), \hat{R}(\omega)$ are tempered random variables.

Proof. From (5.53) we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(\|v_1\|_2^2 + \|v_2\|_2^2) \\
&\le -2d\|\nabla v_1\|_2^2 + (2z(\theta_t\omega) - C_1)(\|v_1\|_2^2 + \|v_2\|_2^2) + C\exp(-2z(\theta_t\omega)).
\end{aligned}$$

Applying Gronwall's inequality over the interval $[t, t+r]$ yields

$$\begin{aligned}
&\|v_1(t+r)\|_2^2 + \|v_2(t+r)\|_2^2 \\
&\le (\|v_1(t)\|_2^2 + \|v_2(t)\|_2^2)\exp\left(\int_t^{t+r}(2z(\theta_s\omega) - C_1)\mathrm{d}s\right) \\
&\quad + \int_t^{t+r}(C\exp(-2z(\theta_s\omega) - 2d\|\nabla v_1(s)\|_2^2)\exp\left(\int_s^{t+r}(2z(\theta_\tau\omega) - C_1)\mathrm{d}\tau\right)\mathrm{d}s,
\end{aligned}$$

and thus

$$\begin{aligned}
&\int_t^{t+r} 2d\|\nabla v_1(s)\|_2^2 \exp\left(\int_s^{t+r}(2z(\theta_\tau\omega) - C_1)\mathrm{d}\tau\right)\mathrm{d}s \\
&\le (\|v_1(t)\|_2^2 + \|v_2(t)\|_2^2)\exp\left(\int_t^{t+r} 2z(\theta_s\omega)\mathrm{d}s - C_1 r\right) \\
&\quad + \int_t^{t+r} C\exp(-2z(\theta_s\omega))\exp\left(\int_s^{t+r}(2z(\theta_\tau\omega) - C_1)\mathrm{d}\tau\right)\mathrm{d}s.
\end{aligned}$$

Now, we replace ω by $\theta_{-t-r}\omega$ and obtain

$$\begin{aligned}
&\int_t^{t+r} 2d\|\nabla v_1(s, \theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_2^2 \exp\left(-2\max_{-r\le\tau\le 0}|z(\theta_\tau\omega)| - C_1 r\right)\mathrm{d}s \\
&\le \int_t^{t+r} 2d\|\nabla v_1(s, \theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_2^2 \exp\left(\int_s^{t+r}(2z(\theta_{\tau-t-r}\omega) - C_1)\mathrm{d}\tau\right)\mathrm{d}s \\
&\le (\|v_1(t, \theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_2^2 + \|v_2(t, \theta_{-t-r}\omega, v_2^0(\theta_{-t-r}\omega))\|_2^2) \\
&\qquad \times \exp\left(\int_t^{t+r} 2z(\theta_{s-t-r}\omega)\mathrm{d}s - C_1 r\right) \\
&\qquad + \int_t^{t+r} C\exp(-2z(\theta_{s-t-r}\omega))\exp\left(\int_s^{t+r}(2z(\theta_{\tau-t-r}\omega) - C_1)\mathrm{d}\tau\right)\mathrm{d}s,
\end{aligned}$$

and therefore, making use of $t \geq t_{\mathcal{D}}(\omega)$,

$$
\begin{aligned}
&\int_t^{t+r} \|\nabla v_1(s,\theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_2^2 \,\mathrm{d}s \\
&\leq \frac{1}{2d}(\|v_1(t,\theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_2^2 + \|v_2(t,\theta_{-t-r}\omega, v_2^0(\theta_{-t-r}\omega))\|_2^2) \\
&\quad \times \exp\left(\int_t^{t+r} 2z(\theta_{s-t-r}\omega)\mathrm{d}s - C_1 r\right) \exp\left(2 \max_{-r\leq\tau\leq 0} |z(\theta_\tau\omega)| + C_1 r\right) \\
&\quad + \frac{1}{2d}\int_t^{t+r} C\exp(-2z(\theta_{s-t-r}\omega))\exp\left(\int_s^{t+r}(2z(\theta_{\tau-t-r}\omega) - C_1)\mathrm{d}\tau\right)\mathrm{d}s \\
&\quad \times \exp\left(2\max_{-r\leq\tau\leq 0}|z(\theta_\tau\omega)| + C_1 r\right) \\
&\leq \frac{1}{2d}\rho(\theta_{-r}\omega)^2 \exp\left(4\max_{-r\leq\tau\leq 0}|z(\theta_\tau\omega)|\right) + \exp\left(2\max_{-r\leq\tau\leq 0}|z(\theta_\tau\omega)| + C_1 r\right) \\
&\quad \times \frac{1}{2d}\int_{-r}^0 C\exp(-2z(\theta_s\omega))\exp\left(\int_s^0 (2z(\theta_\tau\omega))\mathrm{d}\tau + C_1 s\right)\mathrm{d}s \\
&\leq \frac{1}{2d}\rho(\theta_{-r}\omega)^2 \exp\left(4\max_{-r\leq\tau\leq 0}|z(\theta_\tau\omega)|\right) + \exp\left(2\max_{-r\leq\tau\leq 0}|z(\theta_\tau\omega)| + C_1 r\right) \\
&\quad \times \frac{1}{2d}\int_{-r}^0 C\exp\left(4\max_{-r\leq s\leq 0}|z(\theta_s\omega)|\right)\exp(C_1 s)\,\mathrm{d}s \\
&\leq \frac{1}{2d}\rho(\theta_{-r}\omega)^2\exp\left(4\max_{0\leq\tau\leq r}|z(\theta_{-\tau}\omega)|\right) \\
&\quad + C\exp\left(6\max_{0\leq s\leq r}|z(\theta_{-s}\omega)| + C_1 r\right)(1-\exp(-C_1 r)) \\
&=: R(\omega).
\end{aligned}
$$

Now, $\rho(\omega)$ is tempered and so is $\rho(\theta_{-r}\omega)$. Furthermore, by Proposition 4.25 (i) we have for $K > 0$

$$
\lim_{t\to\pm\infty}\frac{\log(\exp(K\max_{0\leq\tau\leq r}|z(\theta_{t-\tau}\omega)|))}{|t|} = K\lim_{t\to\pm\infty}\frac{\max_{0\leq\tau\leq r}|z(\theta_{t-\tau}\omega)|}{|t|} = 0.
$$

Therefore, recalling Lemma 4.7, $R(\omega)$ is a tempered random variable and (5.57) follows. Likewise, (5.53) yields

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(\|v_1\|_2^2 + \|v_2\|_2^2) \\
&\leq -\delta_1\exp((p-2)z(\theta_t\omega))\|v_1\|_p^p + (2z(\theta_t\omega) - C_1)(\|v_1\|_2^2 + \|v_2\|_2^2) + C\exp(-2z(\theta_t\omega)),
\end{aligned}
$$

and with a similar argument as used above we can also show (5.58). □

Lemma 5.29. *Let Assumptions 5.2 hold. Let $\mathcal{D} \in \mathcal{T}$ and $(v_1^0, v_2^0) \in \mathcal{D}$. Furthermore, let $r > 0$ be arbitrary. Then for $t \geq t_{\mathcal{D}}(\omega) + r$ we have*

$$\int_t^{t+r} \|v_1(s, \theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_{2p-2}^{2p-2} \,\mathrm{d}s \leq Q(\omega),$$

where $Q(\omega)$ is a tempered random variable.

Proof. We multiply (5.51) by $|v_1|^{p-2}v_1$ and integrate over D. Making use of Assumptions 5.2, we compute

$$\begin{aligned}
&\frac{1}{p}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |v_1|^p \mathrm{d}x \\
&= d\int_D \triangle v_1 |v_1|^{p-2}v_1 \mathrm{d}x + z(\theta_t\omega)\int_D |v_1|^p \mathrm{d}x \\
&\quad - \exp(-z(\theta_t\omega))\int_D h(x, \exp(z(\theta_t\omega))v_1)|v_1|^{p-2}v_1 \mathrm{d}x \\
&\quad - \exp(-z(\theta_t\omega))\int_D f(x, \exp(z(\theta_t\omega))v_1, \exp(z(\theta_t\omega))v_2)|v_1|^{p-2}v_1 \mathrm{d}x \\
&\leq -d(p-1)\int_D |\nabla v_1|^2|v_1|^{p-2}\mathrm{d}x + z(\theta_t\omega)\int_D |v_1|^p \mathrm{d}x \\
&\quad - \exp(-2z(\theta_t\omega))\int_D |v_1|^{p-2}(\delta_1|\exp(z(\theta_t\omega))v_1|^p - \delta_3)\mathrm{d}x \\
&\quad + \exp(-z(\theta_t\omega))\int_D \delta_4(1 + |\exp(z(\theta_t\omega))v_1|^{p_1} + |\exp(z(\theta_t\omega))v_2|)|v_1|^{p-1}\mathrm{d}x \\
&\leq z(\theta_t\omega)\int_D |v_1|^p \mathrm{d}x + \delta_3\exp(-2z(\theta_t\omega))\int_D |v_1|^{p-2}\mathrm{d}x \\
&\quad - \delta_1\exp((p-2)z(\theta_t\omega))\int_D |v_1|^{2p-2}\mathrm{d}x + \delta_4\exp(-z(\theta_t\omega))\int_D |v_1|^{p-1}\mathrm{d}x \\
&\quad + \delta_4\exp((p_1-1)z(\theta_t\omega))\int_D |v_1|^{p_1+p-1}\mathrm{d}x + \delta_4\int_D |v_2||v_1|^{p-1}\mathrm{d}x \\
&\leq z(\theta_t\omega)\int_D |v_1|^p \mathrm{d}x - \frac{\delta_1}{4}\exp((p-2)z(\theta_t\omega))\int_D |v_1|^{2p-2}\mathrm{d}x + \delta_4\int_D |v_2||v_1|^{p-1}\mathrm{d}x \\
&\quad + C_1\exp(C_2|z(\theta_t\omega)|) \\
&\leq z(\theta_t\omega)\int_D |v_1|^p \mathrm{d}x - \frac{\delta_1}{4}\exp((p-2)z(\theta_t\omega))\int_D |v_1|^{2p-2}\mathrm{d}x + \frac{\delta_1}{8}\exp((p-2)z(\theta_t\omega)) \\
&\quad \times \int_D |v_1|^{2p-2}\mathrm{d}x + C_3\exp(C_4|z(\theta_t\omega)|)\|v_2\|_2^2 + C_1\exp(C_2|z(\theta_t\omega)|).
\end{aligned}$$

Rearranging the terms further yields

$$\begin{aligned}
&\frac{1}{p}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |v_1|^p \mathrm{d}x \\
&\le \frac{\delta_1(p-2)}{8} z(\theta_t\omega)\int_D |v_1|^{2p-2}\mathrm{d}x + C_4 z(\theta_t\omega) - \frac{\delta_1}{8}\exp((p-2)z(\theta_t\omega)) \\
&\quad \times \int_D |v_1|^{2p-2}\mathrm{d}x + C_1\exp(C_2|z(\theta_t\omega)|)(1+\|v_2\|_2^2) \\
&\le \frac{\delta_1(p-2)}{8} z(\theta_t\omega)\int_D |v_1|^{2p-2}\mathrm{d}x - \frac{\delta_1}{8}((p-2)z(\theta_t\omega)+1)\int_D |v_1|^{2p-2}\mathrm{d}x \\
&\quad + C_4|z(\theta_t\omega)| + C_1\exp(C_2|z(\theta_t\omega)|)(1+\|v_2\|_2^2) \\
&\le -\frac{\delta_1}{8}\int_D |v_1|^{2p-2}\mathrm{d}x + C_1\exp(C_2|z(\theta_t\omega)|)(1+\|v_2\|_2^2). \qquad (5.59)
\end{aligned}$$

Therefore

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|v_1\|_p^p &\le -\int_D |v_1|^p\mathrm{d}x + C_1\exp(C_2|z(\theta_t\omega)|)(1+\|v_2\|_2^2) + C_3 \\
&= -\|v_1\|_p^p + C_1\exp(C_2|z(\theta_t\omega)|) + C_1\exp(C_2|z(\theta_t\omega)|)\|v_2\|_2^2 + C_3.
\end{aligned}$$

Now, applying the uniform Gronwall lemma we obtain

$$\begin{aligned}
&\|v_1(t+r,\omega,v_1^0(\omega))\|_p^p \\
&\le \frac{1}{r}\int_t^{t+r}\|v_1(s,\omega,v_1^0(\omega))\|_p^p\mathrm{d}s + C_1\int_t^{t+r}\exp(C_2|z(\theta_s\omega)|)\mathrm{d}s \\
&\quad + C_1\int_t^{t+r}\exp(C_2 z(\theta_s\omega))\|v_2(s,\omega,v_2^0(\omega))\|_2^2\mathrm{d}s + C_3 r,
\end{aligned}$$

and replacing ω by $\theta_{-t-r}\omega$ yields for $t \ge t_{\mathcal{D}}(\omega)$

$$\begin{aligned}
&\|v_1(t+r,\theta_{-t-r}\omega,v_1^0(\theta_{-t-r}\omega))\|_p^p \\
&\le \frac{1}{r}\int_t^{t+r}\|v_1(s,\theta_{-t-r}\omega,v_1^0(\theta_{-t-r}\omega))\|_p^p\mathrm{d}s + C_1\int_t^{t+r}\exp(C_2|z(\theta_{s-t-r}\omega)|)\mathrm{d}s \\
&\quad + \int_t^{t+r} C_1\exp(C_2 z(\theta_{s-t-r}\omega))\|v_2(s,\theta_{-t-r}\omega,v_2^0(\theta_{-t-r}\omega))\|_2^2\mathrm{d}s + C_3 r \\
&\le \frac{1}{r}\hat{R}(\omega) + C_1 r\exp(C_2\max_{0\le s\le r}|z(\theta_{-s}\omega)|) \\
&\quad + C_1 r\exp(C_2\max_{0\le s\le r}|z(\theta_{-s}\omega))|\max_{0\le s\le r}\rho(\theta_{-s}\omega)^2 + C_3 r \\
&=: S(\omega), \qquad (5.60)
\end{aligned}$$

where we have used (5.58) and the absorption property from Lemma 5.26. By similar arguments as in the proof of Lemma 5.28 $S(\omega)$ is a tempered random variable.

Now, (5.59) also yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\|v_1\|_p^p \leq -\frac{\delta_1}{8}\|v_1\|_{2p-2}^{2p-2} + C_1\exp(C_2|z(\theta_t\omega)|)(1+\|v_2\|_2^2).$$

Integrating between t and $t+r$

$$\begin{aligned}&\frac{\delta_1}{8}\int_t^{t+r}\|v_1(s,\omega,v_1^0(\omega))\|_{2p-2}^{2p-2}\mathrm{d}s\\&\leq \|v_1(t,\omega,v_1^0(\omega))\|_p^p + \int_t^{t+r} C_1\exp(C_2|z(\theta_s\omega)|)(1+\|v_2(s,\omega,v_2^0(\omega))\|_2^2)\mathrm{d}s,\end{aligned}$$

replacing ω by $\theta_{-t-r}\omega$, yields further

$$\begin{aligned}&\int_t^{t+r}\|v_1(s,\theta_{-t-r}\omega,v_1^0(\theta_{-t-r}\omega))\|_{2p-2}^{2p-2}\mathrm{d}s\\&\leq C\|v_1(t,\theta_{-t-r}\omega,v_1^0(\theta_{-t-r}\omega))\|_p^p\\&\qquad+\int_t^{t+r} C_1\exp(C_2|z(\theta_{s-t-r}\omega)|)(1+\|v_2(s,\theta_{-t-r}\omega,v_2^0(\theta_{-t-r}\omega))\|_2^2)\mathrm{d}s\\&\leq C\|v_1(t,\theta_{-t-r}\omega,v_1^0(\theta_{-t-r}\omega))\|_p^p\\&\qquad+C_1\exp(C_2\max_{0\leq s\leq r}|z(\theta_{-s}\omega)|)\left(r+\int_t^{t+r}\|v_2(s,\theta_{-t-r}\omega,v_2^0(\theta_{-t-r}\omega))\|_2^2\mathrm{d}s\right).\end{aligned}$$

For $t \geq t_{\mathcal{D}}(\omega)+r$ we can conclude with (5.60) and with the absorption property from Lemma 5.26

$$\begin{aligned}&\int_t^{t+r}\|v_1(s,\theta_{-t-r}\omega,v_1^0(\theta_{-t-r}\omega))\|_{2p-2}^{2p-2}\mathrm{d}s\\&\leq CS(\theta_{-r}\omega)+C_1\exp(C_2\max_{0\leq s\leq r}|z(\theta_{-s}\omega)|)(r+r\max_{0\leq s\leq r}\rho(\theta_{-s}\omega)^2)\\&=: Q(\omega),\end{aligned}$$

where $Q(\omega)$ is tempered. This concludes the proof. □

Lemma 5.30. *Let Assumptions 5.2 hold. Let $\mathcal{D}\in\mathcal{T}$ and $(v_1^0,v_2^0)\in\mathcal{D}$. Assume that $t\geq t_{\mathcal{D}}(\omega)+2r$ for some $r>0$, then*

$$\|\nabla v_1(t,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2 \leq \rho_1(\omega), \tag{5.61}$$

where $\rho_1(\omega)$ is tempered.

Proof. We take formally the inner product of $\frac{\mathrm{d}}{\mathrm{d}t}v_1$ with $-\triangle v_1$ and calculate

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla v_1\|_2^2 = \left\langle \frac{\mathrm{d}}{\mathrm{d}t}v_1, -\triangle v_1 \right\rangle \\
&= -d\|\triangle v_1\|_2^2 - z(\theta_t\omega)\int_D v_1 \triangle v_1 \mathrm{d}x \\
&\quad + \exp(-z(\theta_t\omega))\int_D \delta_8(1+|\exp(z(\theta_t\omega))v_1|^{p-1})|\triangle v_1|\mathrm{d}x \\
&\quad + \exp(-z(\theta_t\omega))\int_D \delta_4(1+|\exp(z(\theta_t\omega))v_1|^{p_1} + |\exp(z(\theta_t\omega))v_2|)|\triangle v_1|\mathrm{d}x \\
&= -d\|\triangle v_1\|_2^2 + z(\theta_t\omega)\int_D \nabla v_1 \nabla v_1 \mathrm{d}x + C\exp(-z(\theta_t\omega)) \\
&\quad \times \int_D (1+|\exp(z(\theta_t\omega))v_1|^{p_1} + |\exp(z(\theta_t\omega))v_2| + |\exp(z(\theta_t\omega))v_1|^{p-1})|\triangle v_1|\mathrm{d}x \\
&\leq -d\|\triangle v_1\|_2^2 + z(\theta_t\omega)\|\nabla v_1\|_2^2 + \frac{d_1}{2}\|\triangle v_1\|_2^2 + C\exp(-2z(\theta_t\omega)) \\
&\quad \times \int_D (1+|\exp(z(\theta_t\omega))v_1|^{p_1} + |\exp(z(\theta_t\omega))v_2| + |\exp(z(\theta_t\omega))v_1|^{p-1})^2\mathrm{d}x \\
&\leq -\frac{dc}{2}\|\nabla v_1\|_2^2 + z(\theta_t\omega)\|\nabla v_1\|_2^2 \\
&\quad + C\exp(-2z(\theta_t\omega))\int_D (1+|\exp(z(\theta_t\omega))v_2|^2 + |\exp(z(\theta_t\omega))v_1|^{2p-2})\mathrm{d}x,
\end{aligned}$$

and thus

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla v_1\|_2^2 \\
&\leq (2z(\theta_t\omega) - dc)\|\nabla v_1\|_2^2 \\
&\qquad + \underbrace{C_1\exp(-2z(\theta_t\omega)) + C_2\|v_2\|_2^2 + C\exp((2p-4)z(\theta_t\omega))\|v_1\|_{2p-2}^{2p-2}}_{=:h(t)}.
\end{aligned}$$

Now, we apply once more the uniform Gronwall lemma and we obtain for $t \geq r$

$$\begin{aligned}
&\|\nabla v_1(t+r,\omega,v_1^0(\omega))\|_2^2 \\
&\leq \left(\frac{1}{r}\int_t^{t+r} \|\nabla v_1(s,\omega,v_1^0(\omega))\|_2^2\mathrm{d}s + \int_t^{t+r} C_1\exp(-2z(\theta_s\omega))\right. \\
&\qquad \left. + C_2\|v_2(s,\omega,v_2^0(\omega))\|_2^2 + C\exp((2p-4)z(\theta_s\omega))\|v_1(s,\omega,v_1^0(\omega))\|_{2p-2}^{2p-2}\mathrm{d}s\right) \\
&\qquad \times \exp\left(\int_t^{t+r}(2z(\theta_s\omega) - dc)\mathrm{d}s\right).
\end{aligned}$$

Replacing ω by $\theta_{-t-r}\omega$, we obtain for $t \geq r + t_{\mathcal{D}}(\omega)$

$$
\begin{aligned}
&\|\nabla v_1(t+r, \theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_2^2 \\
&\leq \left(\frac{1}{r}\int_t^{t+r} \|\nabla v_1(s, \theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_2^2 \mathrm{d}s + \int_t^{t+r} h(s, \theta_{-t-r}\omega)\mathrm{d}s\right) \\
&\qquad \times \exp\left(\int_t^{t+r} (2z(\theta_{s-t-r}\omega) - dc)\mathrm{d}s\right) \\
&\leq \left(\frac{1}{r}R(\omega) + \int_t^{t+r} C_1 \exp(-2z(\theta_{s-t-r}\omega)) + C_2\|v_2(s, \theta_{-t-r}\omega, v_2^0(\theta_{-t-r}\omega))\|_2^2\right. \\
&\qquad \left. + C\exp((2p-4)z(\theta_{s-t-r}\omega))\|v_1(s, \theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_{2p-2}^{2p-2}\mathrm{d}s\right) \\
&\qquad \times \exp\left(\int_{-r}^0 2z(\theta_s\omega)\mathrm{d}s\right) \\
&\leq \left(\frac{1}{r}R(\omega) + \int_{-r}^0 C_1\exp(-2z(\theta_s\omega))\mathrm{d}s + \int_t^{t+r} C_2\|v_2(s, \theta_{-t-r}\omega, v_2^0(\theta_{-t-r}\omega))\|_2^2\mathrm{d}s\right. \\
&\qquad \left. +C\int_t^{t+r} \exp((2p-4)z(\theta_{s-t-r}\omega))\|v_1(s, \theta_{-t-r}\omega, v_1^0(\theta_{-t-r}\omega))\|_{2p-2}^{2p-2}\mathrm{d}s\right) \\
&\qquad \times \exp\left(\int_{-r}^0 2z(\theta_s\omega)\mathrm{d}s\right) \\
&\leq \left(\frac{1}{r}R(\omega) + C_1 r\exp(2\max_{0\leq s\leq r}|z(\theta_{-s}\omega)|) + \int_t^{t+r} C_2\|v_2(s, \theta_{-t-r}\omega, v_2^0(\theta_{-t-r}\omega))\|_2^2\mathrm{d}s\right. \\
&\qquad \left. +C\exp((2p-4)\max_{0\leq\tau\leq r}|z(\theta_{-\tau}\omega)|)\int_t^{t+r} \|v_1(s, \theta_{-t-r}\omega, \theta_{-t-r}\omega)\|_{2p-2}^{2p-2}\mathrm{d}s\right) \\
&\qquad \times \exp\left(2r\max_{0\leq s\leq r}|z(\theta_{-s}\omega)|\right) \\
&\leq \left(\frac{1}{r}R(\omega) + C_1 r\exp(2\max_{0\leq s\leq r}|z(\theta_{-s}\omega)|) + \int_t^{t+r} C_2\rho(\theta_{s-t-r}\omega)^2\mathrm{d}s\right. \\
&\qquad \left. +C\exp\left((2p-4)\max_{0\leq\tau\leq r}|z(\theta_{-\tau}\omega)|\right)Q(\omega)\right)\exp\left(2r\max_{0\leq s\leq r}|z(\theta_{-s}\omega)|\right) \\
&\leq \left(\frac{1}{r}R(\omega) + C_1 r\exp(2\varepsilon\max_{0\leq s\leq r}|z(\theta_{-s}\omega)|) + C_2 r\max_{0\leq s\leq r}\rho(\theta_{-s}\omega)^2\right. \\
&\qquad \left. +C\exp\left((2p-4)\max_{0\leq\tau\leq r}|z(\theta_{-\tau}\omega)|\right)Q(\omega)\right)\exp\left(2r\max_{0\leq s\leq r}|z(\theta_{-s}\omega)|\right) \\
&= \rho_1(\omega)
\end{aligned}
$$

The temperedness of all terms involved imply the temperedness of $\rho_1(\omega)$. Hence, for $t \geq t_{\mathcal{D}} + 2r$ the statement follows. □

Lemma 5.31. *Let Assumptions 5.2 hold. Let $\mathcal{D} \in \mathcal{T}$ and $(v_1^0, v_2^0) \in \mathcal{D}$. Assume that $t \geq t_{\mathcal{D}} + 2r$ for some $r > 0$. Then*

$$\|\nabla v_2^1(t, \theta_{-t}\omega, 0)\|_2^2 \leq \rho_2(\omega), \tag{5.62}$$

where $\rho_2(\omega)$ is a tempered random variable.

Proof. The proof is very similar to the one for the analogous lemma in the case of additive noise. We have

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla v_2^1\|_2^2 = \left\langle \frac{\mathrm{d}}{\mathrm{d}t}v_2^1, -\Delta v_2^1 \right\rangle \\
&= \left\langle (z(\theta_t\omega) - \sigma(x))v_2^1 - \exp(-z(\theta_t\omega))g(x, \exp(z(\theta_t\omega))v_1), -\Delta v_2^1 \right\rangle \\
&= \underbrace{\left\langle (\sigma(x) - z(\theta_t\omega))v_2^1, \Delta v_2^1 \right\rangle}_{=:L_1} + \underbrace{\left\langle \exp(-z(\theta_t\omega))g(x, \exp(z(\theta_t\omega))v_1), \Delta v_2^1 \right\rangle}_{=:L_2}.
\end{aligned}$$

We estimate L_1 and L_2 separately,

$$\begin{aligned}
L_1 &= \int_D (\sigma(x) - z(\theta_t\omega))v_2^1 \Delta v_2^1 \mathrm{d}x \\
&= -\int_D \nabla((\sigma(x) - z(\theta_t\omega))v_2^1)\nabla v_2^1 \mathrm{d}x \\
&\leq (z(\theta_t\omega) - \delta)\|\nabla v_2^1\|_2^2 - \int_D \nabla\sigma(x)v_2^1\nabla v_2^1 \mathrm{d}x,
\end{aligned}$$

and

$$\begin{aligned}
L_2 &= \int_D \exp(-z(\theta_t\omega))g(x, \exp(z(\theta_t\omega))v_1)\Delta v_2^1 \ \mathrm{d}x \\
&= -\exp(-z(\theta_t\omega))\int_D \nabla g(x, \exp(z(\theta_t\omega))v_1) \cdot \nabla v_2^1 \ \mathrm{d}x \\
&= -\exp(-z(\theta_t\omega)) \\
&\quad \times \int_D (\nabla g(x, \exp(z(\theta_t\omega))v_1) + \partial_\xi g(x, \exp(z(\theta_t\omega))v_1)\exp(z(\theta_t\omega))\nabla v_1) \cdot \nabla v_2^1 \ \mathrm{d}x.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla v_2^1\|_2^2 + 2(\delta - z(\theta_t\omega))\|\nabla v_2^1\|_2^2 \\
&\leq \frac{1}{\delta}\int_D \left(|\nabla\sigma(x)v_2^1| + |\exp(-z(\theta_t\omega))\nabla g(x, \exp(z(\theta_t\omega))v_1) + \partial_\xi g(x, \exp(z(\theta_t\omega))v_1)\nabla v_1|\right)^2 \ \mathrm{d}x \\
&\quad + \delta\|\nabla v_2^1\|_2^2,
\end{aligned}$$

and further with (5.7)

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla v_2^1\|_2^2 + (\delta - 2z(\theta_t\omega))\|\nabla v_2^1\|_2^2 \\
&\leq \frac{1}{\delta}\int_D \sum_{i=1}^n \big(|\partial_{x_i}\sigma(x)v_2^1| + \exp(-z(\theta_t\omega))|\partial_{x_i}g(x,\exp(z(\theta_t\omega))v_1)| \\
&\qquad +|\partial_\xi g(x,\exp(z(\theta_t\omega))v_1)\partial_{x_i}v_1|\big)^2 \,\mathrm{d}x \\
&\leq \frac{1}{\delta}\int_D \sum_{i=1}^n \big(C|v_2^1| + \delta_5\exp(-z(\theta_t\omega))(1+\exp(z(\theta_t\omega))|v_1|) + \delta_5|\partial_{x_i}v_1|\big)^2 \,\mathrm{d}x \\
&\leq \frac{2}{\delta}(C+\delta_5\exp(-z(\theta_t\omega)))^2 n\int_D \big(|v_2^1| + 1 + \exp(z(\theta_t\omega))|v_1|\big)^2 \,\mathrm{d}x \\
&\qquad + \frac{2\delta_5^2}{\delta}\int_D \sum_{i=1}^n |\partial_{x_i}v_1|^2 \,\mathrm{d}x \\
&\leq (C + C_1\exp(-2z(\theta_t\omega)))(\|v_2^1\|_2^2 + 1 + \exp(2z(\theta_t\omega))\|v_1\|_2^2) + C_3\|\nabla v_1\|_2^2 \\
&\leq C_1 + C_2\exp(-2z(\theta_t\omega)) + C_3\exp(2|z(\theta_t\omega)|)(\|v_2^1\|_2^2 + \|v_1\|_2^2) + C_4\|\nabla v_1\|_2^2,
\end{aligned}
$$

where $C := \max_{1\leq i\leq n}\max_{x\in\overline{D}}|\partial_{x_i}\sigma(x)|$.

Next, we apply Gronwall's inequality while taking the initial condition into account and we obtain for $t \geq 0$

$$
\begin{aligned}
&\|\nabla v_2^1\|_2^2 \\
&\leq \int_0^t \big[C_1 + C_2\exp(-2z(\theta_s\omega)) + C_3\exp(2|z(\theta_s\omega)|)(\|v_2^1\|_2^2 + \|v_1\|_2^2) + C_4\|\nabla v_1\|_2^2\big] \\
&\qquad \times \exp\left(\int_s^t (2z(\theta_\tau\omega) - \delta)\mathrm{d}\tau\right) \,\mathrm{d}s. \qquad (5.63)
\end{aligned}
$$

We have by (5.53), where $M = \min\{\delta, d/c\}$,

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(\|v_1\|_2^2 + \|v_2\|_2^2) \\
&\leq -d\|\nabla v_1\|_2^2 + (2z(\theta_t\omega) - M)(\|v_1\|_2^2 + \|v_2\|_2^2) + C\exp(-2z(\theta_t\omega)).
\end{aligned}
$$

Using Gronwall, this yields in particular

$$
\begin{aligned}
&\int_0^t \|\nabla v_1\|_2^2 \exp\left(\int_s^t (2z(\theta_\tau\omega) - M)\mathrm{d}\tau\right)\mathrm{d}s \\
&\leq \frac{1}{d}(\|v_1^0\|_2^2 + \|v_2^0\|_2^2)\exp\left(\int_0^t (2z(\theta_s\omega) - M)\mathrm{d}s\right) \\
&\qquad + C\int_0^t \exp(-2z(\theta_s\omega))\exp\left(\int_s^t (2z(\theta_\tau\omega) - M)\mathrm{d}\tau\right)\mathrm{d}s, \qquad (5.64)
\end{aligned}
$$

as well as

$$\begin{aligned}
&\|v_1\|_2^2 + \|v_2\|_2^2 \\
&\le \frac{1}{d}(\|v_1^0\|_2^2 + \|v_2^0\|_2^2) \exp\left(\int_0^t (2z(\theta_s\omega) - M)\mathrm{d}s\right) \\
&\quad + C\int_0^t \exp(-2z(\theta_s\omega)) \exp\left(\int_s^t (2z(\theta_\tau\omega) - M)\mathrm{d}\tau\right)\mathrm{d}s.
\end{aligned}$$

Using the last estimate, we obtain

$$\begin{aligned}
&\int_0^t \exp(2|z(\theta_s\omega)|)(\|v_2^1\|_2^2 + \|v_1\|_2^2) \exp\left(\int_s^t (2z(\theta_\tau\omega) - M)\mathrm{d}\tau\right)\mathrm{d}s \\
&\le \frac{1}{d}\int_0^t \exp(2|z(\theta_s\omega)|)(\|v_1^0\|_2^2 + \|v_2^0\|_2^2) \exp\left(\int_0^s (2z(\theta_\tau\omega) - M)\mathrm{d}\tau\right) \\
&\quad \times \exp\left(\int_s^t (2z(\theta_\tau\omega) - M)\mathrm{d}\tau\right)\mathrm{d}s \\
&\quad + C\int_0^t \exp(2|z(\theta_s\omega)|)\int_0^s \exp(-2z(\theta_\tau\omega)) \exp\left(\int_\tau^s (2z(\theta_r\omega) - M)\mathrm{d}r\right)\mathrm{d}\tau \\
&\quad \times \exp\left(\int_s^t (2z(\theta_\tau\omega) - M)\mathrm{d}\tau\right)\mathrm{d}s \\
&\le \frac{1}{d}(\|v_1^0\|_2^2 + \|v_2^0\|_2^2)\exp(-Mt)\exp\left(\int_0^t 2z(\theta_\tau\omega)\mathrm{d}\tau\right)\int_0^t \exp(2|z(\theta_s\omega)|)\mathrm{d}s \\
&\quad + C\left(\int_0^t \exp(2|z(\theta_s\omega)|)\exp(M(s-t))\exp\left(\int_s^t 2z(\theta_\tau\omega)\mathrm{d}\tau\right)\mathrm{d}s\right)^2 \qquad (5.65)
\end{aligned}$$

Now, using (5.64) and (5.65) and replacing ω by $\theta_{-t}\omega$ we have from (5.63)

$$\begin{aligned}
&\|\nabla v_2^1(t,\theta_{-t}\omega,0)\|_2^2 \\
&\le \int_0^t [C_1 + C_2\exp(-2z(\theta_{s-t}\omega))]\exp\left(\int_s^t (2z(\theta_{\tau-t}\omega) - \delta)\mathrm{d}\tau\right)\,\mathrm{d}s \\
&\quad + C_3\int_0^t \exp(2|z(\theta_{s-t}\omega)|)(\|v_2^1(s,\theta_{-t}\omega,0)\|_2^2 + \|v_1(s,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2) \\
&\quad \times \exp\left(\int_s^t (2z(\theta_{\tau-t}\omega) - \delta)\mathrm{d}\tau\right)\,\mathrm{d}s \\
&\quad + C_4\int_0^t \|\nabla v_1(s,\theta_{-t}\omega,v_1^0(\theta_{-t}\omega))\|_2^2 \exp\left(\int_s^t (2z(\theta_{\tau-t}\omega) - \delta)\mathrm{d}\tau\right)\,\mathrm{d}s,
\end{aligned}$$

and thus

$$
\begin{aligned}
&\|\nabla v_2^1(t,\theta_{-t}\omega,0)\|_2^2 \\
&\leq \int_0^t [C_1 + C_2\exp(-2z(\theta_{s-t}\omega))]\exp\left(\int_{s-t}^0 (2z(\theta_\tau\omega)\mathrm{d}\tau\right)\exp(\delta(s-t))\ \mathrm{d}s \\
&\qquad + C_3(\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2)\exp(-Mt)\exp\left(\int_0^t 2z(\theta_{\tau-t}\omega)\mathrm{d}\tau\right) \\
&\qquad \times \int_0^t \exp(2|z(\theta_{s-t}\omega)|)\mathrm{d}s \\
&\qquad + C_4\left(\int_0^t \exp(2|z(\theta_{s-t}\omega)|)\exp(M(s-t))\exp\left(\int_s^t 2z(\theta_{\tau-t}\omega)\mathrm{d}\tau\right)\mathrm{d}s\right)^2 \\
&\qquad + C_5(\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2)\exp\left(\int_0^t (2z(\theta_{s-t}\omega)-M)\mathrm{d}s\right) \\
&\qquad + C_6\int_0^t \exp(-2z(\theta_{s-t}\omega))\exp\left(\int_s^t (2z(\theta_{\tau-s}\omega)-M)\mathrm{d}\tau\right)\mathrm{d}s \\
&\leq \int_{-\infty}^0 [C_1 + C_2\exp(-2z(\theta_s\omega))]\exp\left(\int_s^0 (2z(\theta_\tau\omega)\mathrm{d}\tau\right)\exp(\delta s)\ \mathrm{d}s \\
&\qquad + C_3(\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2)\exp(-Mt)\exp\left(\int_{-\infty}^0 2z(\theta_\tau\omega)\mathrm{d}\tau\right) \\
&\qquad \times \int_{-\infty}^0 \exp(2|z(\theta_s\omega)|)\mathrm{d}s \\
&\qquad + C_4\left(\int_{-\infty}^0 \exp(2|z(\theta_s\omega)|)\exp(Ms)\exp\left(\int_s^0 2z(\theta_\tau\omega)\mathrm{d}\tau\right)\mathrm{d}s\right)^2 \\
&\qquad + C_5(\|v_1^0(\theta_{-t}\omega)\|_2^2 + \|v_2^0(\theta_{-t}\omega)\|_2^2)\exp(-Mt)\exp\left(\int_{-\infty}^0 2z(\theta_s\omega)\mathrm{d}s\right) \\
&\qquad + C_6\int_{-\infty}^0 \exp(-2z(\theta_s\omega))\exp\left(\int_s^0 2z(\theta_\tau\omega)\mathrm{d}\tau\right)\exp(Ms)\mathrm{d}s.
\end{aligned}
$$

The right hand side can be combined into one tempered random variable $\rho_2(\omega)$. □

With the previous lemma the existence of a random attractor for the random dynamical system ψ follows directly from Theorem 4.19 and by Theorem 4.23 this leads to

Theorem 5.32. *Let Assumptions 5.2 hold. Then the random dynamical system φ generated by (u_1, u_2) has a unique $\mathcal{T}$-random attractor $\mathcal{A}$.*

5.4 Applications

We will give here two applications that fall into the class of systems considered in our analysis above.

5.4.1 FitzHugh-Nagumo system

Let us consider the famous stochastic FitzHugh-Nagumo system, that is, in the additive noise setting

$$\begin{aligned} \mathrm{d}u_1 &= (\nu_1\Delta u_1 - p(x)u_1 - u_1(u_1-1)(u_1-\alpha_1) - u_2)\,\mathrm{d}t + \ B_1\mathrm{d}W_1, \\ \mathrm{d}u_2 &= (\alpha_2 u_1 - \alpha_3 u_2)\,\mathrm{d}t + \ B_2\mathrm{d}W_2, \end{aligned} \tag{5.66}$$

with $D=[0,1]$ and $\alpha_j \in \mathbb{R}$, $j\in\{1,2,3\}$, are fixed parameters and $p\in C^2(D)$. Here, W_1, W_2 are cylindrical Wiener processes and we assume that Assumptions 5.4 are satisfied. In the setting with multiplicative noise the system reads

$$\begin{aligned} \mathrm{d}u_1 &= (\nu_1\Delta u_1 - p(x)u_1 - u_1(u_1-1)(u_1-\alpha_1) - u_2)\,\mathrm{d}t + \ u_1\circ\mathrm{d}B, \\ \mathrm{d}u_2 &= (\alpha_2 u_1 - \alpha_3 u_2)\,\mathrm{d}t + \ u_2\circ\mathrm{d}B, \end{aligned} \tag{5.67}$$

where here B denotes a real-valued Wiener process. As mentioned earlier, this system models the signal propagation in a neuron; in particular, the variable u_1 denotes the *electrical potential* and u_2 is the so-called *recovery variable* that is associated with the local concentration of potassium ions.

Such systems have been considered under various conditions by numerous authors, see the references mentioned in Section 5.1. Our Assumptions 5.2 regarding the reaction terms are satisfied in this example as follows: Identifying the terms in (5.66) and (5.67) with the terms given in (5.3) we have

$$\begin{aligned} &h(x,u_1) = p(x)u_1 + u_1(u_1-1)(u_1-\alpha_1), \quad f(x,u_1,u_2) = u_2, \\ &\sigma(x)u_2 = \alpha_3 u_2, \quad g(x,u_1) = -\alpha_2 u_1. \end{aligned}$$

We have $\sigma(x) = \alpha_3$ and $|f(x,u_1,u_2)| = |u_2|$, i.e., (5.6) and (5.5) are fulfilled. Furthermore, $|\partial_u g(x,u_1)| = |\alpha_2|$ and $|\partial_{x_i} g(x,u_1)| = 0$ for $i=1,\dots,n$, hence (5.7) is satisfied. Finally, as a polynomial with odd degree and negative coefficient for the highest degree, h fulfils (5.4). Thus the analysis above guarantees the existence of global mild solutions and the existence of a random pullback attractor for the stochastic FitzHugh-Nagumo system on a bounded domain in both, the additive and multiplicative linear noise setting.

5.4.2 The driven cubic-quintic Allen-Cahn model

The cubic-quintic Allen-Cahn (or real Ginzburg-Landau) equation is given by

$$\partial_t u = \Delta u + p_1 u + u^3 - u^5, \qquad u = u(t,x), \tag{5.68}$$

where $(t,x) \in [0,T) \times D$, $p_1 \in \mathbb{R}$, is a fixed parameter and we will take D as a bounded open domain with regular boundary. The cubic-quintic polynomial non-linearity frequently occurs in the modelling of Euler buckling [VGVC07], as a re-stabilization mechanism in paradigmatic models for fluid dynamics [MD14], in normal form theory and travelling wave dynamics [KS98, DB94], as well as a test problem for deterministic [Kue15c] and stochastic numerical continuation [Kue15c]. If we want to allow for time-dependent slowly-varying forcing on u and sufficiently regular additive noise, then it is actually very natural to extend the model (5.68) to

$$\begin{aligned} \mathrm{d}u_1 &= \left(\Delta u_1 + p_1 u_1 + u_1^3 - u_1^5 - u_2\right)\ \mathrm{d}t + B_1\ \mathrm{d}W_1, \\ \mathrm{d}u_2 &= \varepsilon(p_2 u_2 - q_2 u_1)\ \mathrm{d}t + B_2\ \mathrm{d}W_2, \end{aligned} \tag{5.69}$$

where p_2, q_2, $0 < \varepsilon \ll 1$ are parameters and B_1, B_2, W_1, W_2 are chosen as in the previous example. One easily checks again that (5.69) fits our general framework as $h(x,u_1) = -p_1 u_1 - u_1^3 + u_1^5$ satisfies the crucial assumption (5.4). The same holds in the linear multiplicative noise setting, when the system reads

$$\begin{aligned} \mathrm{d}u_1 &= \left(\Delta u_1 + p_1 u_1 + u_1^3 - u_1^5 - u_2\right)\ \mathrm{d}t + u_1 \circ \mathrm{d}B, \\ \mathrm{d}u_2 &= \varepsilon(p_2 u_2 - q_2 u_1)\ \mathrm{d}t + u_2 \circ \mathrm{d}B, \end{aligned}$$

with real-valued Wiener process B.

5.5 Non-linear coupling: The stochastic Field-Noyes system

Assumptions 5.2 exclude systems with a non-linear coupling between different components. However, there are many systems relevant for applications that exhibit such a structure. In the following we will analyse one of them, namely the *Field-Noyes system*, in greater detail and show that here as well we can derive the existence of a random attractor.

The *Belousov-Zhabotinsky reaction* is an oscillating reaction that was discovered in the early 1950s by the biochemist B. Belousov (1893-1970), and it was further investigated by the biophysicist A. Zhabotinsky (1938-2008). When performed in a stirred container, oscillations in the concentrations of reactants and products cause a periodic change of color of the solution (*temporal oscillation*). Performed in a shallow, unstirred petri dish also *spatio-temporal oscillations* can be observed. The Belousov-Zhabotinsky reaction is a classical example of non-equilibrium thermodynamics that drives *pattern-formation* in natural systems.

R. Field, E. Körös and R. Noyes developed a mathematical model, today known as the *Field-Noyes model*, to describe the chemical mechanism of the oscillating Belousov-Zhabotinsky reaction in a simplified way [FKN72,FN74]. This model reads as follows

$$\begin{aligned} \frac{\partial u}{\partial t} &= a\triangle u + \varepsilon^{-1}(qw - uw + u - u^2), \\ \frac{\partial v}{\partial t} &= b\triangle v + u - v, \\ \frac{\partial w}{\partial t} &= d\triangle w + \delta^{-1}(-qw - uw + cv), \end{aligned} \tag{5.70}$$

in $(0,\infty)\times D$, where $D \subset \mathbb{R}^3$ is a bounded domain with regular boundary. Here, u denotes the concentration of Bromous acid ($HBrO_2$), v the concentration of Cerium(4+) (Ce^{4+}) and w denotes the concentration of the Bromide ion (Br^{-1}). The letters $\varepsilon, q, \delta, c$ as well as a, b, c, d denote positive constants. This deterministic system has been analysed intensively; see for example [Yag09, Chapter 10], where the existence of global solutions and attractors has been derived.

Again, we would like to analyse a stochastic version of (5.70) and, in particular, investigate the long-term behaviour of solutions in terms of a random attractor. We choose a perturbation by a linear multiplicative noise in the Stratonovich sense. In particular, the influence of the state dependent noise becomes smaller the smaller the concentration of the respective chemical, which seems to be a reasonable modelling assumption for noise induced by internal fluctuations. That is, we consider the following system in $(0,\infty)\times D$ with $D \subset \mathbb{R}^n$ open, bounded with regular boundary

$$\begin{aligned} \mathrm{d}u &= a\Delta u \ \mathrm{d}t + 1/\varepsilon(qw - uw + u - u^2) \ \mathrm{d}t + \sigma u \circ \mathrm{d}B, \\ \mathrm{d}v &= b\Delta v \ \mathrm{d}t + (u - v) \ \mathrm{d}t + \sigma v \circ \mathrm{d}B, \\ \mathrm{d}w &= d\Delta w \ \mathrm{d}t + 1/\delta(-qw - uw + cv) \ \mathrm{d}t + \sigma w \circ \mathrm{d}B, \end{aligned} \tag{5.71}$$

where $(B(t))_{t\in\mathbb{R}}$ is a two-sided, real-valued Wiener process on a probability space $(\Omega, \mathbb{P}, \mathcal{F})$. Furthermore, $\sigma > 0$ controls the intensity of the noise. We equip the system with non-negative initial conditions $u(0,x) = u_0(x) \geq 0$, $v(0,x) = v_0(x) \geq 0$, $w(0,x) = w_0(x) \geq 0$ for $x \in D$ and Neumann boundary conditions on ∂D

$$\frac{\partial u}{\partial n} = \frac{\partial v}{\partial n} = \frac{\partial w}{\partial n} = 0,$$

where n denotes the unit outward normal. As before, the dependence of u, v, w on $(t, x, \omega) \in (0,\infty)\times D \times \Omega$ will often be omitted or be indicated only partially. We choose in this section $n = 3$, as we will need a certain Sobolev embedding later on (see (5.72)) and since this choice also makes sense from a modelling point of view.

Similar systems with multiplicative noise have been analysed for example in [Pha20] (*stochastic Hindmarsh-Rose equations*) and in [TY16] (*stochastic Brusselator system*). In both publications the existence of a random attractor was shown for the respective system.

Remark 5.33. Note that, as we consider here a well stirred solution, we include the Laplace operator in every component as opposed to the partly dissipative structure from the previous sections. However, we anticipate that the partly dissipative case could be handled again by a suitable splitting technique, as it was done in the deterministic setting in [Mar89].

5.5.1 The Cauchy problem

First of all, we will rewrite (5.71) as an abstract Cauchy problem on a Hilbert space. Let us define the spaces

$$H := L^2(D) \times L^2(D) \times L^2(D),\ V := H^1(D) \times H^1(D) \times H^1(D),$$

$$I := \left\{ \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} \in H : h_1, h_2, h_3 \geq 0 \right\}.$$

We denote the norm on H as $\|\cdot\|_H$ and on V as $\|\cdot\|_V$. Furthermore, we set

$$g := (u, v, w)^\top,\ g_0 := (u_0, v_0, w_0)^\top,$$

and we define the following operator in H

$$A := \begin{pmatrix} a\Delta - \frac{1}{\varepsilon} & 0 & 0 \\ 0 & b\Delta - 1 & 0 \\ 0 & 0 & d\Delta - \frac{q}{\delta} \end{pmatrix},$$

understood as the realization under homogeneous Neumann boundary conditions, i.e.

$$\mathcal{D}(A) = \left\{ \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} : h_1, h_2, h_3 \in H^2(D), \frac{\partial h_1}{\partial n} = \frac{\partial h_2}{\partial n} = \frac{\partial h_3}{\partial n} = 0 \text{ on } \partial D \right\}.$$

By classical theory A generates an analytic semigroup on H that we denote as $(T(t))_{t\geq 0}$, see for example [Yag09, Theorem 2.19]. Note that the domains of fractional powers of the sectorial operator $-A$ are given, for $3/4 < \eta < 1$, as

$$\mathcal{D}((-A)^\eta) = \left\{ \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} : h_1, h_2, h_3 \in H^{2\eta}(D), \frac{\partial h_1}{\partial n} = \frac{\partial h_2}{\partial n} = \frac{\partial h_3}{\partial n} = 0 \text{ on } \partial D \right\},$$

and for $0 \leq \eta < 3/4$ as

$$\mathcal{D}((-A)^\eta) = \left\{ \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} : h_1, h_2, h_3 \in H^{2\eta}(D) \right\},$$

see [Yag09, Theorem 16.7]. We now fix some η with $3/4 < \eta < 1$ and define the non-linear operator $f : \mathcal{D}((-A)^\eta) \to H$ as

$$f(g) := (f_1(g), f_2(g), f_3(g))^\top := \begin{pmatrix} 1/\varepsilon(qw - uw + 2u - u^2) \\ u \\ 1/\delta(-uw + cv) \end{pmatrix}.$$

In particular, we have by Sobolev embedding (see Theorem C.1 and recall that the space dimension is $n = 3$) that

$$\mathcal{D}((-A)^\eta) \subset H^{2\eta}(D) \times H^{2\eta}(D) \times H^{2\eta}(D) \subset C(\overline{D}) \times C(\overline{D}) \times C(\overline{D}). \tag{5.72}$$

With these definitions let us rewrite (5.71) as an abstract Cauchy problem on the Hilbert space H

$$\begin{aligned} \mathrm{d}g &= (Ag + f(g))\mathrm{d}t + \sigma g \circ \mathrm{d}B, \quad t > 0, \\ g(0) &= g_0 \in I. \end{aligned} \tag{5.73}$$

5.5.2 Random PDE system

As in the previous sections, we will transform the system of SPDEs (5.71) into a system of random PDEs. Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote the canonical probability space associated to the Brownian motion and we identify $B(t, \omega) = \omega(t)$ for $\omega \in \Omega$. We define the following transformations

$$\begin{aligned} U(t) &:= \exp(-\sigma\omega(t))u(t), \\ V(t) &:= \exp(-\sigma\omega(t))v(t), \\ W(t) &:= \exp(-\sigma\omega(t))w(t). \end{aligned} \tag{5.74}$$

Remark 5.34. For $\sigma > 0$ the stochastic process $X(t) = \exp(-\sigma B(t))$ is a solution to the Stratonovich SDE $\mathrm{d}X(t) = -\sigma X(t) \circ \mathrm{d}B(t)$. This follows immediately from Itô's formula and the conversion between Itô and Stratonovich differentials (see also Subsection 3.1.2)

$$\begin{aligned} \mathrm{d}X(t) &= -\sigma \exp(-\sigma B(t)) \; \mathrm{d}B(t) + \frac{1}{2}\sigma^2 \exp(-\sigma B(t)) \; \mathrm{d}t \\ &= -\sigma X(t) \; \mathrm{d}B(t) + \frac{1}{2}(-\sigma)(-\sigma X(t)) \; \mathrm{d}t \\ &= -\sigma X(t) \circ \mathrm{d}B(t). \end{aligned}$$

Using Remark 5.34 one can easily verify that the transformed random system is given as

$$\begin{aligned} \partial_t U &= a\Delta U - \frac{1}{\varepsilon}U + F_1(U, V, W, t), \\ \partial_t V &= b\Delta V - V + F_2(U, V, W, t), \\ \partial_t W &= d\Delta W - \frac{q}{\delta}W + F_3(U, V, W, t), \end{aligned} \tag{5.75}$$

with the non-autonomous reaction terms

$$
\begin{aligned}
F_1(U,V,W,t) &:= \frac{1}{\varepsilon}(qW - \exp(\sigma\omega(t))UW + 2U - \exp(\sigma\omega(t))U^2),\\
F_2(U,V,W,t) &:= U,\\
F_3(U,V,W,t) &:= \frac{1}{\delta}(-\exp(\sigma\omega(t))UW + cV),
\end{aligned}
$$

and with initial conditions in I

$$
\begin{aligned}
U(0,x) &= \exp(-\sigma\omega(0))u_0(x) = u_0(x),\\
V(0,x) &= \exp(-\sigma\omega(0))v_0(x) = v_0(x),\\
W(0,x) &= \exp(-\sigma\omega(0))w_0(x) = w_0(x),
\end{aligned}
$$

and homogeneous Neumann boundary conditions.

Remark 5.35. Note that for the transformation in (5.74) we used the Brownian motion $B(t,\omega)$ instead of the corresponding Ornstein-Uhlenbeck process as used in the transformation (5.49)-(5.50) in Section 5.3. Using directly the Brownian motion has the advantage that no additional terms appear in the transformed equations, like the $v_1 z(\theta_t\omega)$ in (5.49) and $v_2 z(\theta_t\omega)$ in (5.50).

Defining $G := (U,V,W)^\top$, $G_0 := (U(0,\cdot),V(0,\cdot),W(0,\cdot))^\top$ and $F(G,t) := (F_1(G,t),F_2(G,t),F_3(G,t))^\top$, we can formulate the transformed problem as a *non-autonomous initial value problem* on H

$$
\begin{aligned}
\frac{\mathrm{d}G}{\mathrm{d}t} &= AG + F(G,t), \quad t > 0\\
G(0,\omega;0,G_0) &= G_0 = \exp(-\sigma\omega(0))g_0 = g_0 \in I,
\end{aligned}
\tag{5.76}
$$

where we denote the mild solution at time t with initial condition G_0 at time 0 as $G = G(t,\omega;0,G_0)$. The following two propositions establish the existence of local solutions to (5.76) and their non-negativity.

Proposition 5.36. *For any $G_0 \in H$ problem* (5.76) *possesses a unique local mild solution $G \in C([0,T_{G_0}],H)$, where $T_{G_0} > 0$ depends on $\|G_0\|_H$, satisfying the variation of constants formula*

$$
G(t) = T(t)G_0 + \int_0^t T(t-s)F(G(s),s)\mathrm{d}s, \quad 0 \le t \le T_{G_0}.
$$

Proof. We want to invoke a classical existence result for local solutions of abstract non-autonomous Cauchy problems in Banach spaces as stated in [Yag09, Chapter 4, Section 6]. For that we need to ensure that the following local Lipschitz condition is fulfilled [Yag09, Equation (4.51)]

$$
\begin{aligned}
&\|F(G,t) - F(\tilde{G},s)\|_H\\
&\le \varphi(\|G\|_H + \|\tilde{G}\|_H)\\
&\qquad \Big(\|(-A)^\eta(G-\tilde{G})\|_H + \big(\|(-A)^\eta G\|_H + \|(-A)^\eta \tilde{G}\|_H + 1\big)\big[|t-s| + \|G-\tilde{G}\|_H\big]\Big),
\end{aligned}
$$

for $(G,t),(\tilde{G},s) \in \mathcal{D}((-A)^\eta) \times [0,T]$, where φ is a continuous, increasing function and $T > 0$. We compute

$$
\begin{aligned}
&\|\exp(\sigma\omega(t))U^2 - \exp(\sigma\omega(s))\tilde{U}^2\|_2 \\
&\le \|\exp(\sigma\omega(t))U^2 - \exp(\sigma\omega(t))\tilde{U}^2\|_2 + \|\exp(\sigma\omega(t))\tilde{U}^2 - \exp(\sigma\omega(s))\tilde{U}^2\|_2 \\
&\le \exp\left(\sigma \sup_{s\in[0,T]} |\omega(s)|\right) \|U^2 - \tilde{U}^2\|_2 + |\exp(\sigma\omega(t)) - \exp(\sigma\omega(s))| \, \|\tilde{U}^2\|_2.
\end{aligned}
$$

Now, with $|U^2 - \tilde{U}^2| \le (|U| + |\tilde{U}|)|U - \tilde{U}|$ we have for $U, \tilde{U} \in H^{2\eta}$ invoking (5.72)

$$
\begin{aligned}
\|U^2 - \tilde{U}^2\|_2 &\le (\|U\|_\infty + \|\tilde{U}\|_\infty)\|U - \tilde{U}\|_2 \\
&\le (\|U\|_{H^{2\eta}} + \|\tilde{U}\|_{H^{2\eta}})\|U - \tilde{U}\|_2.
\end{aligned}
$$

Furthermore, using the local Lipschitz continuity of the exponential function and the Hölder continuity of the Brownian motion with some exponent $\delta < \frac{1}{2}$ (recall Remark 3.2) we have

$$
|\exp(\sigma\omega(t)) - \exp(\sigma\omega(s))| \le C|\omega(t) - \omega(s)| \le C'|t-s|^\delta \le C'|t-s|.
$$

Thus we can conclude

$$
\begin{aligned}
&\|\exp(\sigma\omega(t))U^2 - \exp(\sigma\omega(s))\tilde{U}^2\|_2 \\
&\le \exp\left(\sigma \sup_{s\in[0,T]} |\omega(s)|\right) (\|U\|_{H^{2\eta}} + \|\tilde{U}\|_{H^{2\eta}})\|U - \tilde{U}\|_2 + C'|t-s|\|\tilde{U}\|_{H^{2\eta}}\|\tilde{U}\|_2.
\end{aligned}
$$

Furthermore, for $U, \tilde{U}, W, \tilde{W} \in H^{2\eta}$ we can derive similarly

$$
\|UW - \tilde{U}\tilde{W}\|_2 \le C(\|W\|_{H^{2\eta}}\|U - \tilde{U}\|_2 + \|\tilde{U}\|_{H^{2\eta}}\|W - \tilde{W}\|_2),
$$

and thus we obtain

$$
\begin{aligned}
&\|\exp(\sigma\omega(t))UW - \exp(\sigma\omega(s))\tilde{U}\tilde{W}\|_2 \\
&\le \exp\left(\sigma \sup_{s\in[0,T]} |\omega(s)|\right) C(\|W\|_{H^{2\eta}}\|U - \tilde{U}\|_2 + \|\tilde{U}\|_{H^{2\eta}}\|W - \tilde{W}\|_2) \\
&\qquad + C'|t-s|\|\tilde{U}\|_{H^{2\eta}}\|\tilde{W}\|_2.
\end{aligned}
$$

Having these estimates for the non-linear parts of F, we can easily deduce that for $G, \tilde{G} \in \mathcal{D}((-A)^\eta)$, $t, s \in [0,T]$

$$
\begin{aligned}
&\|F(G,t) - F(\tilde{G},s)\|_H \\
&\le C(\|G\|_H + \|\tilde{G}\|_H)(\|(-A)^\eta G\|_H + \|(-A)^\eta \tilde{G}\|_H + 1)(|t-s| + \|G - \tilde{G}\|_H),
\end{aligned}
$$

and thus by [Yag09, Chapter 4, Section 6] for every $G_0 \in H$ there exist a unique local solution up to a time T_{G_0} in $C([0, T_{G_0}], H)$. The theorem also provides the estimate

$$t\| - AG(t)\|_H + \|G(t)\|_H \leq C_{G_0}, \quad 0 < t \leq T_{G_0}, \tag{5.77}$$

for some constant $C_{G_0} > 0$. □

Proposition 5.37. *Let $G_0 \in I$ and let G be the local mild solution of* (5.76)*. Then $G(t) \geq 0$ for all $0 < t \leq T_{G_0}$.*

Proof. The proof works analogously to the proof for the deterministic system, see [Yag09, Chapter 10, Section 2.2]; we provide it for the sake of completeness. Let $G = (U, V, W)^\top$ be the local solution given by Proposition 5.36. We consider the following non-linear operator

$$\tilde{F}(G, t) = \begin{pmatrix} \frac{1}{\varepsilon}(qW - \exp(\sigma\omega(t))UW + 2U - \exp(\sigma\omega(t))U^2) \\ |U| \\ \frac{1}{\delta}(-\exp(\sigma\omega(t))UW + cV) \end{pmatrix}$$

and the corresponding Cauchy problem

$$\begin{aligned} \frac{\mathrm{d}\tilde{G}}{\mathrm{d}t} &= A\tilde{G} + \tilde{F}(\tilde{G}, t), \\ \tilde{G}(0) &= G_0. \end{aligned} \tag{5.78}$$

The existence of a local solution $\tilde{G} = (\tilde{U}, \tilde{V}, \tilde{W})^\top$ up to a time $\tilde{T}_{G_0}$ to this problem follows from the same arguments as outlined in Proposition 5.36. We now verify that $\tilde{G} \in I$.

(i) We define the following non-negative, cut-off function

$$H(v) := \begin{cases} \frac{1}{2}v^2 & , -\infty < v < 0 \\ 0 & , 0 \leq v < \infty \end{cases},$$

then $\varphi(t) := \int_D H(\tilde{V}(t))\mathrm{d}x$ is continuously differentiable with

$$\varphi'(t) = b\int_D H'(\tilde{V})\Delta\tilde{V}\mathrm{d}x + \int_D H'(\tilde{V})(|\tilde{U}| - \tilde{V})\mathrm{d}x.$$

Now note that $H'(\tilde{V}) \leq 0$, $H'(\tilde{V})\tilde{V} \geq 0$ and

$$\int_D H'(\tilde{V})\Delta\tilde{V}\mathrm{d}x = -\int_D \nabla H'(\tilde{V}) \cdot \nabla\tilde{V}\mathrm{d}x = -\int_D |\nabla H'(\tilde{V})|^2\mathrm{d}x \leq 0,$$

thus $\varphi'(t) \leq 0$. Therefore $\varphi(t) \leq \varphi(0)$ and since $\varphi(0) = 0$ we have $\varphi(t) \equiv 0$, i.e. $\tilde{V}(t) \geq 0$ for $0 \leq t \leq \tilde{T}_{G_0}$.

(ii) Let us set $\psi(t) := \int_D H(\tilde{W}(t))\mathrm{d}x$, then

$$\begin{aligned}\psi'(t) &= d\int_D H'(\tilde{W})\Delta\tilde{W}\mathrm{d}x + \int_D H'(\tilde{W})\frac{1}{\delta}(-q\tilde{W} - \exp(\sigma\omega)\tilde{U}\tilde{W} + c\tilde{V})\mathrm{d}x \\ &\leq -\frac{\exp(\sigma\omega)}{\delta}\int_D H'(\tilde{W})\tilde{U}\tilde{W}\mathrm{d}x,\end{aligned}$$

where we have used that

$$\int_D H'(\tilde{W})\Delta\tilde{W}\mathrm{d}x \leq 0,$$

and $H'(\tilde{W}) \leq 0$, $H'(\tilde{W})\tilde{W} \geq 0$, as well as $\tilde{V} \geq 0$. Thus, using that $0 \leq H'(\tilde{W})\tilde{W} \leq 2H(\tilde{W})$,

$$\psi'(t) \leq \frac{\exp(\sigma\omega)}{\delta}\int_D |\tilde{U}| 2H(\tilde{W})\mathrm{d}x \leq \frac{\exp(\sigma\omega)}{\delta}\|\tilde{U}\|_\infty \psi(t).$$

Now making use of (5.77) (which also holds for the local solutions of the modified Cauchy problem (5.78)) we have

$$\|\tilde{U}(t)\|_\infty \leq C\|(-A)^\eta \tilde{G}(t)\|_H \leq C_{G_0} t^{-\eta},$$

and thus by Gronwall's inequality

$$\begin{aligned}\psi(t) &\leq \psi(0)\exp\left(2\int_0^t \exp(\sigma\omega(s))\|\tilde{U}(s)\|_\infty \mathrm{d}s\right) \\ &\leq \psi(0)\exp\left(2C_{G_0}\int_0^t \exp(\sigma\omega(s))s^{-\eta}\mathrm{d}s\right).\end{aligned}$$

Noting the $\mathbb{P}$-a.s. continuity of Brownian motion sample paths, we can conclude from $\psi(0) = 0$ that $\psi(t) \equiv 0$, i.e. $\tilde{W}(t) \geq 0$ for $0 < t \leq \tilde{T}_{G_0}$.

(iii) Finally, consider $\chi(t) := \int_D H(\tilde{U}(t))\mathrm{d}x$. By similar arguments as before we derive

$$\begin{aligned}&\chi'(t) \\ &= a\int_D H'(\tilde{U})\Delta\tilde{U}\mathrm{d}x + \int_D H'(\tilde{U})\frac{1}{\varepsilon}(\tilde{U} + q\tilde{W} - \exp(\sigma\omega)\tilde{U}\tilde{W} - \exp(\sigma\omega)\tilde{U}^2)\mathrm{d}x \\ &\leq \frac{2}{\varepsilon}\int_D (1 + \exp(\sigma\omega)|\tilde{W}| + \exp(\sigma\omega)|\tilde{U}|)H(\tilde{U})\mathrm{d}x \\ &\leq \frac{2}{\varepsilon}(1 + \exp(\sigma\omega)\|\tilde{W}(t)\|_\infty + \exp(\sigma\omega)\|\tilde{U}(t)\|_\infty)\chi(t),\end{aligned}$$

and we can conclude as before that $\chi(t) \equiv 0$, i.e. $\tilde{U}(t) \geq 0$ for $0 < t \leq \tilde{T}_{G_0}$.

Having established the componentwise $\mathbb{P}$-a.s. non-negativity of $\tilde{G}$ we can deduce that $F(\tilde{G}(t),t) = \tilde{F}(\tilde{G}(t),t)$ for $0 \le t \le \tilde{T}_{G_0}$. This implies that $G(t) = \tilde{G}(t)$ for $0 \le t \le \tilde{T}_{G_0}$, where $G(t)$ is the unique solution of (5.76). That is for $\tilde{T}_{G_0} \ge T_{G_0}$ the proof is finished. For $\tilde{T}_{G_0} < T_{G_0}$ consider

$$T_0 = \sup\{0 < T \le T_{G_0} : U(t), V(t), W(t) \ge 0 \text{ for all } 0 < t \le T\}.$$

The continuity of $\varphi(t), \psi(t), \chi(t)$ implies that also $V(T_0), W(T_0), U(T_0) \ge 0$ and if $T_0 = T_{G_0}$ we are finished. If $T_0 < T_{G_0}$ we repeat the procedure with initial time T_0 and initial value $G(T_0)$, this would give us $\tau > 0$ such that $U(t), V(t), W(t) \ge 0$ for $T_0 \le t \le T_0 + \tau$, which is a contradiction, hence $T_0 = T_{G_0}$. □

Remark 5.38. In particular, $\mathcal{M} = [0,\infty) \times [0,\infty) \times [0,\infty)$ can be considered as an invariant set for the system, that is, starting in $\mathcal{M}$, the local solution will stay in $\mathcal{M}$ $\mathbb{P}$-a.s.. This knowledge will allow us to obtain the necessary a-priori estimates that guarantee that solutions exist globally in time, see the following subsection.

Remark 5.39. We also note that the solution continuously depends on the initial data by classical results. Furthermore, we note that the mild and weak solution coincide as in the partly dissipative setting from Subsection 5.2, see [Bal77, Już14]. In the following we refer to solutions always in this mild/weak framework.

5.5.3 Global existence of solutions

In this subsection we will deduce that solutions to the random PDE system exist not only locally, but globally in time. Recall that our transformed system reads as follows with some initial time t_0

$$\frac{\mathrm{d}G}{\mathrm{d}t} = AG + F(G,t), \quad t > t_0, \tag{5.79}$$

$$G(t_0, \omega; t_0, G_{t_0}) = G_{t_0} = \exp(-\sigma\omega(t_0))g_0 \in I, \tag{5.80}$$

and we denote the solution at time t as $G = G(t, \omega; t_0, G_{t_0})$, starting at t_0 with G_{t_0}.

Proposition 5.40. *For any given random variable $\rho(\omega) > 0$, there is a time $-\infty < \tau(\rho,\omega) \le -1$ such that for any $t_0 \le \tau(\rho,\omega)$ and for any initial data $g_0 \in I$ with $\|g_0\|_H \le \rho(\omega)$, the solution $G(t,\omega;t_0,G_{t_0})$ of* (5.79)-(5.80) *uniquely exists on* $[t_0,\infty)$.

Proof. The proof works similarly to the proof of [Pha20, Lemma 2.2 and Lemma 2.3]. Let $\rho(\omega) > 0$ and $\|g_0\|_H \le \rho(\omega)$. Multiplying equation (5.79) componentwise by G and integrating over D yields for the first component

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |U|^2 \mathrm{d}x + a\int_D |\nabla U|^2 \mathrm{d}x + \frac{1}{\varepsilon}\int_D U^2 \mathrm{d}x \\
&= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |U|^2 \mathrm{d}x - a\int_D \Delta U U \mathrm{d}x + \frac{1}{\varepsilon}\int_D U^2 \mathrm{d}x \\
&= \int_D \frac{1}{\varepsilon}(qWU - \exp(\sigma\omega)U^2 W + 2U^2 - \exp(\sigma\omega)U^3)\mathrm{d}x \\
&\leq \int_D \frac{1}{\varepsilon}\left(\frac{q^2}{4}\exp(-\sigma\omega)W + 2U^2 - \exp(\sigma\omega)U^3\right)\mathrm{d}x,
\end{aligned}
$$

where we have used that $qWU \leq \exp(\sigma\omega)W(q^2/4\exp(-2\sigma\omega) + U^2)$. For the second component we obtain

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |V|^2 \mathrm{d}x + b\int_D |\nabla V|^2 \mathrm{d}x + \frac{1}{2}\int_D V^2 \mathrm{d}x \\
&= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |V|^2 \mathrm{d}x - b\int_D \Delta V V \mathrm{d}x + \frac{1}{2}\int_D V^2 \mathrm{d}x \\
&= \int_D (UV - \frac{1}{2}V^2)\mathrm{d}x \\
&\leq \int_D U^2 - \frac{1}{4}V^2 \mathrm{d}x,
\end{aligned}
$$

and finally for the third component

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |W|^2 \mathrm{d}x + d\int_D |\nabla W|^2 \mathrm{d}x + \frac{q}{\delta}\frac{1}{2}\int_D W^2 \mathrm{d}x \\
&= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_D |W|^2 \mathrm{d}x - d\int_D \Delta W W \mathrm{d}x + \frac{q}{\delta}\frac{1}{2}\int_D W^2 \mathrm{d}x \\
&= \int_D -\frac{q}{\delta}\frac{1}{2}W^2 - \frac{1}{\delta}\exp(\sigma\omega)UW^2 + \frac{c}{\delta}VW \mathrm{d}x \\
&\leq \int_D -\frac{q}{\delta}\frac{1}{4}W^2 + \frac{c^2}{q\delta}V^2 \mathrm{d}x,
\end{aligned}
$$

where we have used the non-negativity of the local solution as derived in Proposition

5.37. Now set $\xi = \frac{4c^2}{q\delta}$ and combine the three estimates from above

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_D (U^2 + \xi V^2 + W^2)\mathrm{d}x + \int_D (a|\nabla U|^2 + b|\nabla V|^2 + d|\nabla W|^2)\mathrm{d}x \\
&\qquad + \frac{1}{2}\int_D \left(\frac{2}{\varepsilon}U^2 + \xi V^2 + \frac{q}{\delta}W^2\right)\mathrm{d}x \\
&\le \int_D U^2\left(\frac{2}{\varepsilon} - \frac{1}{\varepsilon}\exp(\sigma\omega)U + \xi\right) + W\left(\frac{q^2}{4\varepsilon}\exp(-\sigma\omega) - \frac{q}{\delta}\frac{1}{4}W\right)\mathrm{d}x \\
&\le \int_D \left(\frac{2}{\varepsilon} + \xi\right)^3 \frac{4}{27}\varepsilon^2 \exp(-2\sigma\omega) + \frac{q^3\delta}{16\varepsilon^2}\exp(-2\sigma\omega)\mathrm{d}x \\
&= |D|\left(\frac{2}{\varepsilon} + \xi\right)^3 \frac{4}{27}\varepsilon^2 \exp(-2\sigma\omega) + |D|\frac{q^3\delta}{16\varepsilon^2}\exp(-2\sigma\omega) \\
&= C\exp(-2\sigma\omega), \qquad (5.81)
\end{aligned}$$

where we have used

$$U^2\left(\frac{2}{\varepsilon} + \xi\right) \le \frac{\exp(\sigma W)}{\varepsilon}U^3 + \left(\frac{2}{\varepsilon} + \xi\right)^3 \frac{4}{27}\varepsilon^2 \exp(-2\sigma\omega),$$

and

$$\frac{1}{\varepsilon}q\exp(-\sigma\omega)W \le \frac{1}{\delta}W^2 + \frac{q^2\delta}{4\varepsilon^2}\exp(-2\sigma\omega).$$

Thus, we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\|U\|_2^2 + \xi\|V\|_2^2 + \|W\|_2^2\right) + 2M(\|\nabla U\|_2^2 + \|\nabla V\|_2^2 + \|\nabla W\|_2^2) \qquad (5.82) \\
&\qquad + \mu\left(\|U\|_2^2 + \xi\|V\|_2^2 + \|W\|_2^2\right) \\
&\le C\exp(-2\sigma\omega),
\end{aligned}$$

where we have set $\mu := \min\{2/\varepsilon, 1, q/\delta\}$ and $M := \min\{a, b, d\}$. Applying Gronwall's inequality yields

$$\begin{aligned}
\|U\|_2^2 + \xi\|V\|_2^2 + \|W\|_2^2 &\le \exp(-\mu(t - t_0))\exp(-2\sigma\omega(t_0))(\|u_0\|_2^2 + \xi\|v_0\|_2^2 + \|w_0\|_2^2) \\
&\qquad + C\int_{t_0}^t \exp(-(t-s)\mu)\exp(-2\sigma\omega(s))\mathrm{d}s,
\end{aligned}$$

that is, with $\underline{\xi} = \min\{1, \xi\}, \overline{\xi} = \max\{1, \xi\}$, we have for $t \ge t_0$

$$\begin{aligned}
\|G(t, \omega; t_0, G_{t_0})\|_H^2 &\le \frac{\overline{\xi}}{\underline{\xi}}\exp(-\mu(t - t_0))\exp(-2\sigma\omega(t_0))\|g_0\|_H^2 \\
&\qquad + C\int_{-\infty}^t \exp(-(t-s)\mu)\exp(-2\sigma\omega(s))\mathrm{d}s. \qquad (5.83)
\end{aligned}$$

Let us take $t = -1$, then

$$\|G(-1,\omega;t_0,G_{t_0})\|_H^2 \leq \frac{\overline{\xi}}{\underline{\xi}} \exp(\mu - \mu|t_0| - 2\sigma\omega(t_0))\rho(\omega)^2 + C\int_{-\infty}^{-1} \exp(\mu+\mu s)\exp(-2\sigma\omega(s))\mathrm{d}s.$$

By the sub-linear growth of the Brownian motion (see Proposition 4.25) the integral on the right hand side converges. Furthermore, for any $\rho(\omega) > 0$ and for a.e. ω there exists a time $\tau(\rho,\omega) \leq -1$ such that for any $t_0 \leq \tau(\rho,\omega)$

$$1 - \frac{2\sigma\omega(t_0)}{\mu t_0} \geq 1/2, \quad \text{and} \quad \exp(\mu(1-1/2|t_0|))\frac{\overline{\xi}}{\underline{\xi}}\rho(\omega)^2 \leq 1.$$

Thus, with $\exp(-\mu|t_0| - 2\sigma\omega(t_0)) = \exp\left(-\mu|t_0|(1-\frac{2\sigma\omega(t_0)}{\mu t_0})\right)$ it follows

$$\|G(-1,\omega;t_0,g_0)\|_H \leq r_0(\omega),$$

where $r_0(\omega) := \sqrt{1 + C\int_{-\infty}^{-1}\exp(\mu+\mu s)\exp(-2\sigma\omega(s))\mathrm{d}s}$.

We can also integrate (5.82) over $[-1,t]$ then we obtain

$$\begin{aligned}\|G(t,\omega;t_0,g_0)\|_H^2 &\leq \frac{\overline{\xi}}{\underline{\xi}}\|G(-1,\omega;t_0,g_0)\|_H^2 + C\int_{-1}^{t}\exp(-2\sigma\omega(s))\mathrm{d}s \qquad (5.84)\\ &\leq \frac{\overline{\xi}}{\underline{\xi}} r_0(\omega)^2 + C\int_{-1}^{t}\exp(-2\sigma\omega(s))\mathrm{d}s,\end{aligned}$$

thus for a.e. $\omega \in \Omega$ and any $T > -1$ the weak solution uniquely exists for $t \in [t_0, T]$ and does not blow up. Uniqueness follows from the uniqueness of the local solution. □

5.5.4 RDS and bounded absorbing set

In the following we will define a random dynamical system associated to the original problem (5.73) and derive the existence of a bounded absorbing set.

Lemma 5.41. *$S(t,\tau,\omega)g_0 = \exp(\sigma\omega(t))G(t,\omega;\tau,G_\tau)$ defines a stochastic flow.*

Proof. Recall Definition 4.14. Proposition 5.40 ensures that $G(t,\omega;\tau,G_\tau)$ uniquely exists on $[\tau,\infty)$, and therefore $S(t,\tau,\omega)g_0$ as well. By the uniqueness of solutions we have for $\tau \leq s \leq t$ and $\omega \in \Omega$

$$\begin{aligned}S(t,s,\omega)S(s,\tau,\omega)g_0 &= S(t,s,\omega)\exp(\sigma\omega(s))G(s,\omega;\tau,G_\tau)\\ &= \exp(\sigma\omega(t))G(t,\omega;s,G(s,\omega,\tau,G_\tau))\\ &= \exp(\sigma\omega(t))G(t,\omega;\tau,G_\tau)\\ &= S(t,\tau,\omega)g_0.\end{aligned}$$

Furthermore, let $(T(t))_{t\geq 0}$ denote the semigroup generated by A. Using the mild formulation of solutions we show that

$$\begin{aligned}
&S(t-\tau,0,\theta_\tau\omega)g_0\\
&=\exp(\sigma\theta_\tau\omega(t-\tau))G(t-\tau,\theta_\tau\omega;0,G_0)\\
&=\exp(\sigma\omega(t)-\sigma\omega(\tau))\,\Big[T(t-\tau)g_0\\
&\qquad+\int_0^{t-\tau}T(t-\tau-r)\exp(-\sigma\theta_\tau\omega(r))f(\exp(\sigma\theta_\tau\omega(r))G(r))\mathrm{d}r\Big]\\
&=\exp(\sigma\omega(t))\,\Big[T(t-\tau)\exp(-\sigma\omega(\tau))g_0\\
&\qquad+\int_0^{t-\tau}T(t-\tau-r)\exp(-\sigma\omega(r+\tau))f(\exp(\sigma\theta_\tau\omega(r))G(r))\mathrm{d}r\Big]\\
&=\exp(\sigma\omega(t))\,\Big[T(t-\tau)\exp(-\sigma\omega(\tau))g_0\\
&\qquad+\int_\tau^{t}T(t-r)\exp(-\sigma\omega(r))f(\exp(\sigma\theta_\tau\omega(r-\tau))G(r-\tau))\mathrm{d}r\Big]\\
&=\exp(\sigma\omega(t))\,\Big[T(t-\tau)\exp(-\sigma\omega(\tau))g_0\\
&\qquad+\int_\tau^{t}T(t-r)\exp(-\sigma\omega(r))f(\exp(\sigma\omega(r)-\sigma\omega(\tau))G(r-\tau))\mathrm{d}r\Big]\\
&=\exp(\sigma\omega(t))G(t,\omega;\tau,G_\tau)\\
&=S(t,\tau,\omega)g_0.
\end{aligned}$$

The measurability condition can be verified easily, since the solution is continuous with respect to the initial conditions and measurable with respect to ω, see also [GLR11, Theorem 1.4]. □

Now we can define a random dynamical system associated to our SPDE system.

Lemma 5.42. *The mapping $\varphi:\mathbb{R}^+\times\Omega\times H\to H$ defined via $\varphi(t,\omega,g_0)=S(t,0,\omega)g_0$ defines a RDS over the MDS $(\Omega,\mathcal{F},\mathbb{P},\theta)$ associated to the original problem* (5.73)*.*

Proof. We have $\varphi(0,\omega,g_0)=S(0,0,\omega)g_0=g_0$ and the cocycle property can be verified as follows

$$\begin{aligned}
\varphi(t,\theta_s\omega,\varphi(s,\omega,g_0)) &= S(t,0,\theta_s\omega)\varphi(s,\omega,g_0)\\
&= S(t,0,\theta_s\omega)S(s,0,\omega)g_0\\
&= S(t+s,s,\omega)S(s,0,\omega)g_0\\
&= S(t+s,0,\omega)g_0\\
&= \varphi(t+s,\omega,g_0).
\end{aligned}$$

□

Note that we work here directly with the RDS associated to the SPDE instead of a conjugated RDS associated to the random PDE, as we did in Subsections 5.2 and 5.3. Let us denote by $\mathcal{T}$ the set of all tempered subsets of H.

Proposition 5.43. *There exists a bounded $\mathcal{T}$-absorbing set for the RDS φ given by the ball $\mathcal{B}(\omega) := B(0, R_0(\omega))$, where $R_0(\omega)$ is a tempered random variable as defined in* (5.85).

Proof. The proof works similarly to the proof of [Pha20, Theorem 2.6]. Let $B(0,\rho(\omega))$ be an arbitrary tempered set and $g_0 \in B(0,\rho(\theta_{-t}\omega))$. Note that

$$\varphi(t,\theta_{-t}\omega,g_0) = S(t,0,\theta_{-t}\omega)g_0 = S(0,-t,\omega)g_0 = G(0,\omega;-t,G_{-t}).$$

Now, by (5.84)

$$\|G(0,\omega;-t,G_{-t})\|_H^2 \leq \frac{\overline{\xi}}{\underline{\xi}}\|G(-1,\omega;-t,G_{-t})\|_H^2 + C\int_{-1}^{0}\exp(-2\sigma\omega(s))\mathrm{d}s,$$

and by (5.83)

$$\begin{aligned}
&\|G(-1,\omega;-t,G_{-t})\|_H^2 \\
&\leq \frac{\overline{\xi}}{\underline{\xi}}\exp(-\mu(-1+t))\exp(-2\sigma\omega(-t))\|g_0\|_H^2 \\
&\qquad + C\int_{-\infty}^{-1}\exp((1+s)\mu)\exp(-2\sigma\omega(s))\mathrm{d}s \\
&\leq \frac{\overline{\xi}}{\underline{\xi}}\exp(\mu)\exp\left(-\frac{\mu t}{2}\left(1-\frac{4\sigma}{\mu}\frac{\omega(-t)}{-t}\right)\right)\exp(-\mu t/2)\rho(\theta_{-t}\omega)^2 \\
&\qquad + C\int_{-\infty}^{-1}\exp((1+s)\mu)\exp(-2\sigma\omega(s))\mathrm{d}s.
\end{aligned}$$

Now, due to the sub-linear growth property of the Brownian motion and the temperedness of ρ, we have

$$\lim_{t\to\infty}\exp\left(-\frac{\mu t}{2}\left(1-\frac{4\sigma}{\mu}\frac{\omega(-t)}{-t}\right)\right) = 0$$

and

$$\lim_{t\to\infty}\exp(-\mu t/2)\rho(\theta_{-t}\omega)^2 = 0.$$

Thus, there exist a finite random time $T(\rho,\omega) > 1$ such that for all $t \geq T(\rho,\omega)$

$$\|G(-1,\omega;-t,G_{-t})\|_H^2 \leq 1 + C\int_{-\infty}^{-1}\exp((1+s)\mu)\exp(-2\sigma\omega(s))\mathrm{d}s = r_0^2(\omega).$$

Therefore for $t \geq T(\rho, \omega)$

$$\begin{aligned}\|\varphi(t, \theta_{-t}\omega, g_0)\|_H^2 &= \|G(0, \omega; -t, G_{-t})\|_H^2 \\ &\leq \frac{\overline{\xi}}{\underline{\xi}} r_0^2(\omega) + C \int_{-1}^0 \exp(-2\sigma\omega(s))\mathrm{d}s =: R_0(\omega)^2. \end{aligned} \tag{5.85}$$

Thus $\mathcal{B}(\omega) := B(0, R_0(\omega))$ is a bounded tempered absorbing set for the RDS φ. □

Let us set $T(\omega) := \sup_{B(0,\rho(\omega))\in\mathcal{T}} T(\rho, \omega)$ as the total absorption time for the universe $\mathcal{T}$.

5.5.5 Compact absorbing set

Having the Laplace operator in every component of (5.75) we can, in contrast to the partly dissipative systems from Sections 5.2 and 5.3, derive in this setting an absorbing set in V for the full RDS φ, without invoking any splitting technique. Due to compact embedding this allows us to directly derive the existence of a compact absorbing set.

Proposition 5.44. *There exists a random variable $R(\omega) > 0$ such that for any given random variable $\rho(\omega) > 0$ there is a finite time $0 < \hat{T}(\rho, \omega)$ such that if $g_0 \in I$ with $\|g_0\|_H \leq \rho(\omega)$, then*

$$\|\varphi(t, \theta_{-t}\omega, g_0)\|_V \leq R(\omega), \qquad \text{for } t > \hat{T}(\rho, \omega). \tag{5.86}$$

Proof. The proof works similarly to the proof of [Pha20, Lemma 3.1 and 3.2]. Multiplying (5.79) by $-\triangle G$ and integrating over D yields for the first component

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla U\|_2^2 + a\|\triangle U\|_2^2 \\ &= -\frac{1}{\varepsilon}\int_D qW\triangle U\mathrm{d}x + \frac{1}{\varepsilon}\int_D \exp(\sigma\omega)UW\triangle U\mathrm{d}x - \frac{1}{\varepsilon}\int_D U\triangle U\mathrm{d}x \\ &\qquad + \frac{1}{\varepsilon}\int_D \exp(\sigma\omega)U^2\triangle U\mathrm{d}x \\ &\leq \frac{a}{8}\int_D \triangle U^2\mathrm{d}x + \frac{2}{a}\left(\frac{q}{\varepsilon}\right)^2\int_D W^2\mathrm{d}x + \frac{a}{8}\int_D \triangle U^2\mathrm{d}x + \frac{2}{a}\left(\frac{\exp(\sigma\omega)}{\varepsilon}\right)^2\int_D U^2W^2\mathrm{d}x \\ &\qquad + \frac{a}{8}\int_D \triangle U^2\mathrm{d}x + \frac{2}{a}\frac{1}{\varepsilon^2}\int_D U^2\mathrm{d}x + \frac{a}{8}\int_D \triangle U^2\mathrm{d}x + \frac{2}{a}\left(\frac{\exp(\sigma\omega)}{\varepsilon}\right)^2\int_D U^4\mathrm{d}x,\end{aligned}$$

and thus

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla U\|_2^2 + a\|\triangle U\|_2^2 \\
&\leq \frac{4}{a}\left(\frac{q}{\varepsilon}\right)^2\|W\|_2^2 + \frac{4}{a}\frac{1}{\varepsilon^2}\|U\|_2^2 + \frac{4}{a}\left(\frac{\exp(\sigma\omega)}{\varepsilon}\right)^2\|U\|_4^4 + \frac{2}{a}\left(\frac{\exp(\sigma\omega)}{\varepsilon}\right)^2\|U\|_4^4 \\
&\quad + \frac{2}{a}\left(\frac{\exp(\sigma\omega)}{\varepsilon}\right)^2\|W\|_4^4 \\
&= \frac{4}{a}\left(\frac{q}{\varepsilon}\right)^2\|W\|_2^2 + \frac{4}{a}\frac{1}{\varepsilon^2}\|U\|_2^2 + \frac{6}{a}\left(\frac{\exp(\sigma\omega)}{\varepsilon}\right)^2\|U\|_4^4 + \frac{2}{a}\left(\frac{\exp(\sigma\omega)}{\varepsilon}\right)^2\|W\|_4^4.
\end{aligned}$$

For the second component we obtain

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla V\|_2^2 + b\|\triangle V\|_2^2 &= -\int_D U\triangle V\mathrm{d}x + \int V\triangle V\mathrm{d}x \\
&\leq \frac{b}{2}\|\triangle V\|_2^2 + \frac{1}{2b}\|U\|_2^2 - \|\nabla V\|_2^2,
\end{aligned}$$

and thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla V\|_2^2 + b\|\triangle V\|_2^2 \leq \frac{1}{b}\|U\|_2^2 - 2\|\nabla V\|_2^2.$$

Finally, for the third component we compute

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla W\|_2^2 + d\|\triangle W\|_2^2 \\
&= \frac{q}{\delta}\int_D W\triangle W\mathrm{d}x + \frac{\exp(\sigma\omega)}{\delta}\int_D UW\triangle W\mathrm{d}x - \frac{c}{\delta}\int_D V\triangle W\mathrm{d}x \\
&\leq -\frac{q}{\delta}\|\nabla W\|_2^2 + \frac{d}{4}\|\triangle W\|_2^2 + \frac{1}{d}\left(\frac{\exp(\sigma\omega)}{\delta}\right)^2\int_D U^2W^2\mathrm{d}x + \frac{d}{4}\|\triangle W\|_2^2 \\
&\quad + \frac{1}{d}\left(\frac{c}{\delta}\right)^2\|V\|_2^2,
\end{aligned}$$

that is

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla W\|_2^2 + d\|\triangle W\|_2^2 \\
&\leq \frac{2}{d}\left(\frac{c}{\delta}\right)^2\|V\|_2^2 - 2\frac{q}{\delta}\|\nabla W\|_2^2 + \frac{2}{d}\left(\frac{\exp(\sigma\omega)}{\delta}\right)^2\|U\|_4^4 + \frac{2}{d}\left(\frac{\exp(\sigma\omega)}{\delta}\right)^2\|W\|_4^4.
\end{aligned}$$

Combining these three inequalities we arrive at

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(\|\nabla U\|_2^2+\|\nabla V\|_2^2+\|\nabla W\|_2^2)+(a\|\triangle U\|_2^2+b\|\triangle V\|_2^2+d\|\triangle W\|_2^2)\\
&\quad +2\|\nabla V\|_2^2+2\left(\frac{q}{\delta}\right)\|\nabla W\|_2^2\\
&\le\left(\frac{1}{b}+\frac{4}{a\varepsilon^2}\right)\|U\|_2^2+\frac{2}{d}\left(\frac{c}{\delta}\right)^2\|V\|_2^2+\frac{4}{a}\left(\frac{q}{\varepsilon}\right)^2\|W\|_2^2\\
&\quad +\left(\frac{6}{a}\left(\frac{\exp(\sigma\omega)}{\varepsilon}\right)^2+\frac{2}{d}\left(\frac{\exp(\sigma\omega)}{\delta}\right)^2\right)\|U\|_4^4\\
&\quad +\left(\frac{2}{a}\left(\frac{\exp(\sigma\omega)}{\varepsilon}\right)^2+\frac{2}{d}\left(\frac{\exp(\sigma\omega)}{\delta}\right)^2\right)\|W\|_4^4.
\end{aligned}$$

Now we have the Sobolev embedding $H^1(D)\hookrightarrow L^4(D)$ (cf. Theorem C.1), i.e. there exists $\chi>0$ such that

$$\|U\|_4^4\le\chi(\|U\|_2^2+\|\nabla U\|_2^2)^2\le 2\chi(\|U\|_2^4+\|\nabla U\|_2^4),$$

and

$$\|W\|_4^4\le 2\chi(\|W\|_2^4+\|\nabla W\|_2^4).$$

Thus we obtain the following inequality

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla G(t,\omega;\tau,G_\tau)\|_H^2\\
&\le C_1\|G(t,\omega;\tau,G_\tau)\|_H^2+C_2\exp(2\sigma\omega)\|G(t,\omega;\tau,G_\tau)\|_H^4\\
&\quad +C_2\exp(2\sigma\omega)\|\nabla G(t,\omega;\tau,G_\tau)\|_H^2\|\nabla G(t,\omega;\tau,G_\tau)\|_H^2,
\end{aligned}$$

where $C_1=\max\left\{\left(\frac{1}{b}+\frac{4}{a\varepsilon^2}\right),\frac{2}{d}\left(\frac{c}{\delta}\right)^2,\frac{4}{a}\left(\frac{q}{\varepsilon}\right)^2\right\}$, $C_2=\left(\frac{6}{a}\left(\frac{1}{\varepsilon}\right)^2+\frac{2}{d}\left(\frac{1}{\delta}\right)^2\right)2\chi$. We want to apply the uniform Gronwall lemma to this inequality. We therefore note the following. Let $t^*\in[-2,-1]$. Let $g_0\in\mathcal{B}(\omega)$ (that is we can assume $\rho(\omega)=R_0(\omega)$ due to the absorption property) and recall that $G_\tau=\exp(-\sigma\omega(\tau))g_0$.

(i) We prove that there exists a time $T^*(R_0(\omega))<-2$ such that for any $\tau\le T^*(R_0(\omega))$ and $t\in[-2,0]$

$$\|G(t,\omega;\tau,G_\tau)\|_H\le\rho_1(\omega), \tag{5.87}$$

where $\rho_1(\omega)$ is a positive random variable defined below. From (5.83) we have for $t=-2$

$$\begin{aligned}
\|G(-2,\omega;\tau,G_\tau)\|_H^2\le{}&\frac{\overline{\xi}}{\underline{\xi}}\exp(2\mu-\mu|\tau|-2\sigma\omega(\tau))R_0^2(\omega)\\
&+C\int_{-\infty}^{-2}\exp(2\mu+s\mu)\exp(-2\sigma\omega(s))\mathrm{d}s.
\end{aligned}$$

As in the proof of Proposition 5.40, we note that by the sub-linear growth property of the Wiener process and the temperedness of $R_0(\omega)$ there exist a $T^*(R_0(\omega)) \leq -2$ such that for any $\tau \leq T^*(R_0(\omega))$,

$$1 - \frac{2\sigma\omega(\tau)}{\mu\tau} \geq 1/2, \qquad \exp(\mu(2 - 1/2|\tau|))\frac{\overline{\xi}}{\underline{\xi}} R_0^2(\omega) \leq 1.$$

Thus for $\tau \leq T^*(R_0(\omega))$

$$\begin{aligned}\|G(-2,\omega;\tau,G_\tau)\|_H^2 &\leq \frac{\overline{\xi}}{\underline{\xi}} \exp\left(2\mu - \mu|\tau|\left[1 - \frac{2\sigma\omega(\tau)}{\mu\tau}\right]\right) R_0^2(\omega) \\ &\quad + C\int_{-\infty}^{-2} \exp(2\mu + s\mu)\exp(-2\sigma\omega(s))\mathrm{d}s \\ &\leq 1 + C\int_{-\infty}^{-1} \exp(\mu + s\mu)\exp(-2\sigma\omega(s))\mathrm{d}s \\ &= r_0(\omega)^2.\end{aligned}$$

Now integrating (5.82) over $[-2, t]$ for $t \in [-2, 0]$ yields

$$\begin{aligned}&\|G(t,\omega;\tau,G_\tau)\|_H^2 + \frac{2M}{\underline{\xi}}\int_{-2}^{t} \|\nabla G(s,\omega;\tau,G_\tau)\|_H^2 \mathrm{d}s \\ &\leq \frac{\overline{\xi}}{\underline{\xi}}\|G(-2,\omega;\tau,G_\tau)\|_H^2 + C\int_{-2}^{t} \exp(-2\sigma\omega(s))\mathrm{d}s \\ &\leq \frac{\overline{\xi}}{\underline{\xi}} r_0(\omega)^2 + C\int_{-2}^{0} \exp(-2\sigma\omega(s))\mathrm{d}s \\ &=: \rho_1(\omega).\end{aligned}$$

This establishes (5.87).

(ii) Now, by (5.81), for $t \in [t^*, -1]$, we can deduce that for any $\tau < T^*(R_0(\omega))$

$$\begin{aligned}\int_t^{t+1} \|\nabla G(t,\omega;\tau,G_\tau)\|_H^2 \mathrm{d}s &\leq \frac{\overline{\xi}}{\underline{\xi}2M}\rho_1(\omega) + \frac{C}{2M\underline{\xi}}\exp\left(2\sigma \sup_{s\in[-2,0]} |\omega(s)|\right) \\ &=: \kappa_1(\omega)\end{aligned}$$

(iii) Furthermore, for $t \in [t^*, -1]$, we can deduce that for any $\tau < T^*(R_0(\omega))$

$$\begin{aligned}&\int_t^{t+1} C_2 \exp(2\sigma\omega(s))\|\nabla G(t,\omega;\tau,G_\tau)\|_H^2 \mathrm{d}s \\ &\leq C_2 \exp\left(2\sigma \sup_{s\in[-2,0]} |\omega(s)|\right)\kappa_1(\omega) =: \kappa_2(\omega).\end{aligned}$$

(iv) And finally for $t \in [t^*, -1]$, we can deduce that for any $\tau < T^*(R_0(\omega))$

$$\int_t^{t+1} C_1 \|G(t,\omega;\tau,G_\tau)\|_H^2 + C_2 \exp(2\sigma\omega)\|G(t,\omega;\tau,G_\tau)\|_H^4 \mathrm{d}s$$
$$\leq C_1\rho_1(\omega) + C_2 \exp\left(2\sigma \sup_{s\in[-2,0]} |\omega(s)|\right) \rho_1(\omega)^2 =: \kappa_3(\omega).$$

Now, applying the uniform Gronwall lemma yields for all $t \in [t^* + 1, 0]$ and $\tau < T^*(R_0(\omega))$

$$\|\nabla G(t,\omega;\tau,G_\tau)\|_H^2 \leq (\kappa_1(\omega) + \kappa_3(\omega))e^{\kappa_2(\omega)} =: R_1(\omega)^2.$$

Therefore for $t > \hat{T}(\rho,\omega) := \max\{|T^*(R_0(\omega))|, T(\rho,\omega)\}$

$$\begin{aligned}\|\varphi(t,\theta_{-t}\omega,g_0)\|_V &= \|G(0,\omega;-t,G_{-t})\|_V \\ &= \sqrt{\|G(0,\omega;-t,G_{-t})\|_H^2 + \|\nabla G(0,\omega;-t,G_{-t})\|_H^2} \\ &= \sqrt{R_1(\omega)^2 + R_0(\omega)^2} =: R(\omega).\end{aligned}$$

□

The existence of a random attractor is now a straightforward consequence.

Theorem 5.45. *There exists a unique $\mathcal{T}$-random attractor for the random dynamical system φ.*

Proof. By Proposition 5.44 there exist a bounded absorbing set in V and due to the compact embedding $V \hookrightarrow H$ (Theorem C.2) this implies that there exists a compact absorbing set. The statement thus follows by Theorem 4.19. □

5.6 Outlook

In this chapter we have derived the existence of random attractors for different stochastic (partly dissipative) reaction-diffusion systems with different types of Wiener noise, i.e. additive and multiplicative and different forms of couplings, i.e. linear and non-linear. There are several directions in which this research could be extended in the future.

As mentioned before, in [Mar89, Mar87] the existence of random attractors for general partly/fully dissipative autonomous reaction-diffusion systems that possess an invariant region has been derived. It would be desirable to extend this theory to the stochastic setting, as started in Subsection 5.5 (see Remark 5.38). More specifically, in the deterministic setting the invariant region is used in order to derive a bounded absorbing set. That is, let us for simplicity look at a one component reaction-diffusion equation on $D \subset \mathbb{R}^n$

$$\partial_t u = d\triangle u + f(x,u), \tag{5.88}$$

where f is sufficiently regular and assume that $\mathcal{M} \subset \mathbb{R}$ is a closed convex region that is positively invariant and $c = \sup_{(x,u)\in D\times\mathcal{M}} |f(x,u)|$ is finite. Assume that $u(0,x) = u_0(x)$ and $u_0 \in L^2(D,\mathcal{M})$. Then multiplying (5.88) by u and integrating over the spatial domain D yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_2^2 + d\|\nabla u\|_2^2 = \int_D f(x,u)u\mathrm{d}x \le c|D|^{1/2}\|u\|_2.$$

Using the Poincaŕe inequality and Young's inequality one can infer that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_2^2 + C_1\|u\|_2^2 \le C_2,$$

for some constants $C_1, C_2 > 0$. Thus by Gronwall a bounded absorbing set in the form of a closed ball in $L^2(D,\mathcal{M})$ can be derived easily. Now, consider the following stochastic equation, where B is a real-valued Wiener process

$$\mathrm{d}u = (d\triangle u + f(u))\mathrm{d}t + u \circ \mathrm{d}B, \tag{5.89}$$

which can be converted by a standard Doss-Sussmann transformation into the random equation

$$\partial_t U = d\triangle U + \exp(-B(t))f(\exp(B(t))U). \tag{5.90}$$

Suppose that the solution remains non-negative for non-negative initial conditions and that f is bounded on $[0,\infty)$, then we could do a similar analysis as above, i.e. we would arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\|U\|_2^2 + C_1\|U\|_2^2 \le C_2\exp(-2B(t)),$$

that is, we would have the analogue of equation (5.82) that we encountered for the Field-Noyes system and just like there we could infer the existence of a bounded absorbing set for (5.90). Note that for the Field-Noyes system we only showed that a priori the solution will remain non-negative for non-negative initial conditions, however a priori the non-linearity is not bounded on $[0,\infty)^3$. Thus we had to make explicitly use of the structure of the non-linearity in order to derive (5.82). Nevertheless, one may be able to generalize the result that we have obtained for the Field-Noyes system to a larger class of stochastic reaction-diffusion equations whose solutions preserve the non-negativity of the initial condition and whose non-linearity obeys a certain structure, see for example [Kot92].

Concerning systems with general invariant regions let us note that there are few results available on *deterministic invariant regions* for stochastic equations.

In [CV04] the authors considered systems of semi-linear stochastic equations, whose deterministic counterpart possesses a bounded invariant set. They showed through a *Wong-Zakaï* type approximation that there exist uniformly bounded approximate solutions to their stochastic systems, what then allowed them to employ deterministic techniques to derive the existence of an invariant region for the stochastic system. One may be able to use this result for the derivation of a random attractor for these systems by invoking similar thoughts as described above. Note however, that the restrictions on the non-linear terms in [CV04] are rather strict and do not allow for non-linear couplings between components. That is, this result could have not been applied for the Field-Noyes system. Another starting point might be to consider *random invariant regions*, however, in this case, a different approach compared to the one above will probably be needed.

Of course, it would be desirable to not only establish the existence of random attractors for the systems covered in this chapter but also to characterise the attractors further. That is, first of all, it would be useful to find bounds on the dimensions of the random attractors and to compare them to the determinisitc results, see Remark 4.21. Another possibility would be to show that a random exponential attractor exists. These are compact subsets of the phase space that are attracting at an exponential rate and, since they are of finite fractal dimension and contain the global random attractor, their existence would already imply that the global random attractor has finite fractal dimension as well, see [CS17]. Furthermore, a finer resolution of the dynamics might be possible by analysing invariant manifolds of the system, see for example [DLS03, CS10]. In the following chapter, we will explore this possibility in so-called fast-slow SPDE systems, where the dynamics evolve on two well-separated time scales. More precisely, we will characterise the fluctuations of sample paths around an hyperbolic attracting slow manifold of the corresponding deterministic system.

Chapter 6

Fast-slow stochastic partial differential equations

In this chapter we consider one of the simplest representatives of *fast-slow SPDEs*, namely a linear SPDE with a slowly varying parameter. We prove that for a short period of time the probability for the sample paths to leave a uniform neighbourhood around the stable slow manifold (which is simply the zero solution in this case) is exponentially small. This chapter is based on joint work with Manuel Gnann and Christian Kuehn, which was published in [GKP19].

6.1 Introduction

Fast-slow systems consist of a pair of coupled ordinary differential equations (ODE's), where one of the equations contains a very small scaling parameter ε in the derivative. This leads to the phenomenon that the coupled sub-processes evolve on *well-separated time scales*. As this kind of behaviour can be observed in many physical and biological systems, fast-slow systems naturally arise as an important tool for mathematical modelling. Real-world examples include

- *Ocean-atmosphere systems* in climate models, where the slow variable describes the state of the ocean and the fast variable describes the state of the atmosphere.

- *Predator-prey systems* in ecology when the reproduction rates of predator and prey differ strongly.

- *Enzymatic reactions*, where the enzyme's concentration evolves much faster than the concentration of other involved reactants.

The mathematical analysis of fast-slow systems is well-established and several asymptotic and geometric techniques have been developed throughout the years.

We refer the reader to the following monographs for a thorough background [Eck11, Kue15b, KC12, O'm91, Ver05]. One of the most interesting dynamical behaviours that can be observed in non-linear fast-slow systems is that of relaxation oscillations. These are periodic motions of the system due to an alternation in fast and slow phases, see Example 6.8 below for more details.

As outlined in the Introduction, by adding a noise term to a differential or partial differential equation a physical system can often be modelled more realistically. It is thus not surprising that *stochastic fast-slow systems* were intensively analysed as well. Applications range again from neuroscience [LSG99, SRT04], over climate science [BPSV82, MTvdE01, Has76], to ecology [SMSG07, SK18], among many other areas. In particular, the *sample path viewpoint*, developed by Berglund and Gentz [BG06], yields a very successful technique to analyse these systems and to characterise noise-induced phenomena. In their theory they focus on random perturbations given by a real-valued Brownian motion or a finite sum of Brownian motions, see Section 6.3 for more details. In [EKN20] the authors started to extend this theory to stochastic fast-slow systems driven by *fractional Brownian motion* (fBM). A fBM is a centred Gaussian process that is parametrized by the so-called Hurst parameter $H \in (0,1)$. For $H \neq 1/2$ the increments of the corresponding process are no longer independent but correlated, allowing to model long-range dependencies.

It would be very desirable to have a generalization of the theory by Berglund and Gentz to *fast-slow SPDEs*. Examples of such systems arising in applications are the FitzHugh-Nagumo SPDE [GOS12, BK16], slowly-driven amplitude/modulation equations [Blö07, GK15], and degenerate controlled SPDEs [LPS15, LSP18]. There are certainly many other important examples as most PDEs arising in applications have parameters, which quite often are *slow variables*, and those PDEs should frequently have *noise terms*, e.g. due to internal or external fluctuations. We formulate an important class of such equations as follows

$$\begin{aligned} \mathrm{d}u &= [Au + f(u,v,\varepsilon)]\ \mathrm{d}t + \sigma\ \mathrm{d}W, \\ \mathrm{d}v &= \varepsilon g(u,v,\varepsilon)\ \mathrm{d}t, \end{aligned} \tag{6.1}$$

where A is a differential operator that generates a strongly continuous semigroup on a Hilbert space space H and W is a H-valued Q-Wiener process. As a first step towards such a generalization, we consider the situation where the system is reduced to a *scalar linear non-autonomous SPDE*. This is already the theoretical analogue to the key step in the SODE theory of Berglund/Gentz; see the linearized parts of the estimates in [BG06, Section 5.1.2].

We investigate the SPDE on a bounded interval so that the solution can be expressed by a *Fourier series* and the SPDE is naturally reformulated as an infinite-dimensional system of SODEs. The linear reaction term consists of a *time-dependent coefficient* and a *non-local operator* that generates linear couplings between the first

$k_* - 1$ modes. Thus, it is natural to split the system into two parts: the first part consisting of the first $k_* - 1$ coupled *low-frequency modes* and the second part consisting of infinitely many decoupled *high-frequency modes*. Both components are estimated by *Bernstein-type inequalities* and the limiting process in the second part is approached by an iteration argument. The probability to exit a suitably chosen neighbourhood around the stable slow manifold of the system can then be estimated by convolving the corresponding probabilities for finite and high frequencies. In particular, we are able to show that for a short period of time this probability is exponentially small with respect to the size of the neighbourhood. Systems of the form presented here arise not only directly in numerical spectral Galerkin methods for SPDEs [Kue15c] but also in the context of inertial manifolds defined via a finite number of effective Fourier modes; see [Tem12] for the classical deterministic setting.

The remaining part of this chapter is structured as follows: In Section 6.2 we will formally define fast-slow systems and introduce the corresponding terminology. Subsequently, in Section 6.3 we will consider the stochastic counterpart in finite dimensions and, in particular, present relevant parts of the theory by Berglund and Gentz. In Section 6.4 we consider the extension to SPDE's as described above. Results in this section were published in [GKP19] (joint work with Manuel Gnann and Christian Kuehn). More precisely, in Subsection 6.4.1 we define the setting at hand in detail and we derive the finite-dimensional approximation in Subsection 6.4.2. The main result and technical contributions are contained in Subsections 6.4.3-6.4.7. Lastly, we provide a summary and an outlook on future research in Section 6.5.

6.2 Deterministic fast-slow systems

In this section we briefly introduce the terminology of fast-slow systems. We keep this section short and we refer the reader to [Kue15b] and the references mentioned therein for a detailed presentation of the theory.

Definition 6.1. A *fast-slow system* is a system of ODE's of the following form

$$\begin{aligned} \varepsilon\frac{\mathrm{d}u}{\mathrm{d}\tau} &= f(u,v,\varepsilon), \\ \frac{\mathrm{d}v}{\mathrm{d}\tau} &= g(u,v,\varepsilon), \end{aligned} \tag{6.2}$$

where u respectively v are unknown $\mathbb{R}^m$- respectively $\mathbb{R}^n$-valued functions, $f:\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}^m$ and $g:\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}^n$ are sufficiently regular vector fields and $0<\varepsilon\ll 1$ is the so-called *time-scale parameter*. We refer to u as the *fast variable*, to v as the *slow variable* and to τ as the *slow time scale*. By defining the *fast time scale* $t:=\tau/\varepsilon$ we can transform (6.2) into the equivalent system

$$\begin{aligned} \frac{\mathrm{d}u}{\mathrm{d}t} &= f(u,v,\varepsilon), \\ \frac{\mathrm{d}v}{\mathrm{d}t} &= \varepsilon g(u,v,\varepsilon). \end{aligned} \tag{6.3}$$

A first approach towards analysing the dynamics of a fast-slow system is to take the so-called *singular limit*, that is setting $\varepsilon = 0$. Depending on the time scale this leads to two different subsystems:

Definition 6.2. The singular limit of (6.2) yields a differential algebraic equation, the so-called *slow subsystem*

$$\begin{aligned} 0 &= f(u,v,0), \\ \frac{\mathrm{d}v}{\mathrm{d}\tau} &= g(u,v,0). \end{aligned} \tag{6.4}$$

The flow generated by this system is called the *slow flow*. Taking the singular limit in (6.3) results in the so-called *fast subsystem*

$$\begin{aligned} \frac{\mathrm{d}u}{\mathrm{d}t} &= f(u,v,0), \\ \frac{\mathrm{d}v}{\mathrm{d}t} &= 0. \end{aligned} \tag{6.5}$$

This is a parameterized ODE and the generated flow is called the *fast flow*. Note that the two subsystems (6.4) and (6.5) are not equivalent any more.

The main idea is to analyse invariant objects of the two limiting subsystems and then use perturbation methods to characterise the dynamics of the full systems (6.2) and (6.3) for $\varepsilon > 0$ sufficiently small. This approach is often termed *Geometric Singular Perturbation Theory* (GSPT). A key object in this theory is defined in the following.

Definition 6.3. The set

$$C_0 = \{(u,v) \in \mathbb{R}^m \times \mathbb{R}^n : f(u,v,0) = 0\},$$

is called the *critical set*. If C_0 is a manifold, we call it the *critical manifold*.

The critical set consists of equilibria of the fast subsystem and also determines the phase space of solutions of the slow subsystem. It can be further characterised in the following way.

Definition 6.4. We call a point $p = (u,v) \in C_0$ *hyperbolic* if the matrix $(\mathrm{D}_u f)(p,0)$ has no eigenvalues with vanishing real part. Here D_u denotes the total derivative with respect to u. We call a subset $S \subset C_0$ *normally hyperbolic* if all $p \in S$ are hyperbolic.

If all eigenvalues of $(\mathrm{D}_u f)(p,0)$ have negative real part for all $p \in S$, we call S *attracting*. Similarly, if all eigenvalues of $(\mathrm{D}_u f)(p,0)$ have positive real part for all $p \in S$, we call S *repelling*. A normally hyperbolic subset S where the eigenvalues of $(\mathrm{D}_u f)(p,0)$ have both positive and negative real parts, is called of *saddle-type*.

Here, we focus on the case where the critical manifold is given locally by a graph of the slow variable, that is for an open subset $D_0 \subset \mathbb{R}^n$ we can write

$$C_0 = \{(u^*(v),v) \in \mathbb{R}^m \times \mathbb{R}^n : u^* : D_0 \to \mathbb{R}^m, f(u^*(v),v,0) = 0, v \in D_0\}. \tag{6.6}$$

Remark 6.5. Note that for a normally hyperbolic critical manifold such a local representation is always guaranteed by the *implicit function theorem*.

Assume that C_0 is normally hyperbolic attracting. In this case solutions to (6.2) starting close to the critical manifold will approach C_0 in a time of order $\varepsilon|\log\varepsilon|$. During this time interval the fast flow provides a reasonable approximation for the dynamics. For larger times solutions remain in an ε-neighbourhood of C_0 and they are well approximated by the slow flow. This result of an exponentially fast convergence towards an *ε-neighbourhood* of a normally hyperbolic attracting critical manifold is due to Tihonov [Tik52] and Gradshtein [Gra53]. Furthermore, Fenichel provided by means of a geometrical approach insight into the dynamics within this neighbourhood in terms of an *invariant manifold*. Recall that a manifold M in a phase space is invariant with respect to a flow ϕ_t if $\phi_t(m) \in M$ for all $m \in M$ and $t \in \mathbb{R}$.

Theorem 6.6 (Fenichel's theorem [Fen79])**.** *Let the critical manifold of the fast-slow system* (6.2) *be given by* (6.6)*. Assume that C_0 is normally hyperbolic and that $f, g \in C^k(\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R})$, $1 \le k < \infty$. Then there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon \le \varepsilon_0$ there exists a locally invariant C^k-smooth manifold, called the slow manifold,*

$$C_\varepsilon = \{(u,v) \in \mathbb{R}^m \times \mathbb{R}^n : u = \overline{u}(v,\varepsilon) := u^*(v) + \mathcal{O}(\varepsilon), v \in D_0\}. \tag{6.7}$$

Furthermore, C_ε has the same local stability properties with respect to the fast variable as C_0.

Remark 6.7. In the normally hyperbolic case *Fenichel's theorem* is one of the main tools for analysing dynamical phenomena in fast-slow systems. In settings with a non-hyperbolic critical manifold other methods have been developed such as the so-called *blow up method*, see for instance [Kue15b] and [EK20] for a recent result in the infinite-dimensional setting.

Example 6.8 (Van der Pol equation)**.** *A classical example of a fast-slow system is generated by the famous Van der Pol equation*

$$u'' + \mu(u^2 - 1)u' + u = a, \tag{6.8}$$

where $\mu \gg 1$ is a parameter and $a > 0$ is a constant forcing term. Transforming this second order ODE into a system of first order equations

$$\begin{aligned} \varepsilon u' &= v - \frac{u^3}{3} + u, \\ v' &= a - u, \end{aligned} \tag{6.9}$$

where we have set $\varepsilon := 1/\mu^2$, reveals the typical fast-slow structure. Let us consider the unforced setting with $a = 0$. The critical manifold is given by the cubic curve $C_0 = \{(u,v) \in \mathbb{R}^2 : v = \frac{u^3}{3} - u\}$ with its two fold points $F_1 = (-1, +2/3)$ and $F_2 =$

$(1, -2/3)$, *see the blue graph on the right side in Figure 6.1. We can furthermore identify a normally hyperbolic repelling part of the critical manifold, namely* $C_0^r = C_0 \cap \{-1 < u < 1\}$, *and two normally hyperbolic attracting parts, that is* $C_0^1 = C_0 \cap \{u < -1\}$ *and* $C_0^2 = C_0 \cap \{u > 1\}$.

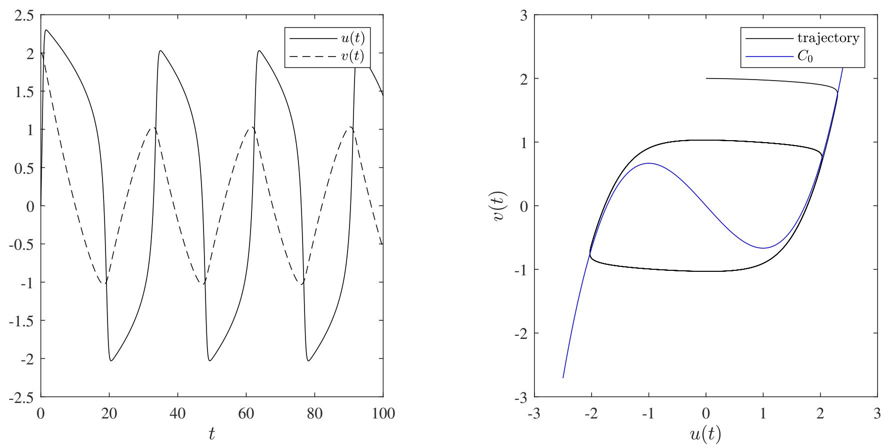

Figure 6.1: Numerical solution of (6.9) with $\varepsilon = 0.1$, $a = 0$ and initial conditions $u(0) = 0$, $v(0) = 2$ using a simple Euler method. We observe *relaxation oscillations* of the fast and slow variables.

In the singular limit of the fast subsystem

$$\begin{aligned} \frac{\mathrm{d}u}{\mathrm{d}t} &= v - \frac{u^3}{3} + u, \\ \frac{\mathrm{d}v}{\mathrm{d}t} &= 0, \end{aligned}$$

v *is a constant parameter, and depending on its value the dynamics of the fast variable* u *exhibits one, two or three equilibrium points lying on the critical manifold. Thus trajectories start with a fast motion towards the attracting branches* C_0^1 *or* C_0^2. *The slow subsystem, given by*

$$\begin{aligned} 0 &= v - \frac{u^3}{3} + u, \\ \frac{\mathrm{d}v}{\mathrm{d}\tau} &= -u, \end{aligned}$$

describes the flow on the critical manifold in the singular limit $\varepsilon = 0$. *That is trajectories starting on* C_0 *move upward on* C_0^1 *and downward on* C_0^2 *according to* $\frac{\mathrm{d}u}{\mathrm{d}\tau} = -\frac{u}{u^2-1}$. *Approaching one of the fold points, let us say* F_1, *the trajectory jumps in the fast subsystem to the drop point* $(2, 2/3) \in C_0^2$. *Then it continues along* C_0^2 *towards* F_2, *where it jumps back to the drop point on* C_0^1. *Thus we observe a periodic orbit in the singular limits* $\varepsilon = 0$. *This alternation between slow dynamics and fast* *dynamics in the form of jumps between generic fold points and normally hyperbolic drop points, is called a relaxation oscillation. If the full dynamics converge in the singular limit $\varepsilon \to 0$ to such a behaviour, or conversely the periodic orbit for $\varepsilon = 0$ persists as an attracting limit cycle for $\varepsilon > 0$, we call the full system a relaxation oscillator. For the Van der Pol system this is indeed the case, which can be observed in simulations of the full dynamics, see Figure 6.1 left.*

6.3 Stochastic fast-slow systems

Definition 6.9. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space. A general *stochastic fast-slow system* has the form

$$\begin{aligned} \mathrm{d}u &= \tfrac{1}{\varepsilon} f(u,v)\ \mathrm{d}\tau + \tfrac{\sigma_f}{\sqrt{\varepsilon}} F(u,v)\ \mathrm{d}B, \\ \mathrm{d}v &= g(u,v)\ \mathrm{d}\tau + \sigma_g G(u,v)\ \mathrm{d}B, \end{aligned} \tag{6.10}$$

where $(u,v) = (u(\tau,\omega), v(\tau,\omega)) \in \mathbb{R}^m \times \mathbb{R}^n$, $\omega \in \Omega$ and $B = B(\tau,\omega)$ is a k-dimensional vector of iid Brownian motions defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Furthermore, $0 < \varepsilon, \sigma_f, \sigma_g \ll 1$ are small parameters, the maps f, g, F, G are assumed to be sufficiently smooth and all maps have suitable domains and ranges, that is $F : \mathbb{R}^{m+n} \times \mathbb{R} \to \mathbb{R}^{m\times k}$, $G : \mathbb{R}^{m+n} \times \mathbb{R} \to \mathbb{R}^{m\times k}$, $f : \mathbb{R}^{m+n} \to \mathbb{R}^m$ and $g : \mathbb{R}^{m+n} \to \mathbb{R}^n$. The parameter ε controls the time scale separation between the fast u variables and the slow v variables, while the parameters σ_f and σ_g regulate the noise level. We further define the ratio $\rho := \sigma_g/\sigma_f$.

The solution of (6.10) is a stochastic process $(u(\tau,\omega), v(\tau,\omega))$ depending both on time $\tau \in \mathbb{R}^+$ and the realisation $\omega \in \Omega$. There are two major viewpoints for analysing solutions: Either one is interested in the *distribution* of the corresponding random variables for fixed times τ or the goal is to characterise the behaviour of *sample paths*, that is the temporal evolution of solutions for fixed realisation ω. We will focus here on the latter one.

Remark 6.10. Note that as we are mainly interested in sample paths we often omit the dependency on $\omega \in \Omega$ in our notations, e.g. we write $(u(\tau), v(\tau))$ for solutions of (6.10).

Example 6.11 (Stochastic Van der Pol equation)**.** *We pick up once more the standard example of the Van der Pol system* (6.9) *and perturb the fast variable by a Brownian motion $B(t)$*

$$\begin{aligned} \mathrm{d}u &= (v - \tfrac{u^3}{3} + u)\ \mathrm{d}t + \sigma\ \mathrm{d}B, \\ \mathrm{d}v &= -\varepsilon u\ \mathrm{d}t. \end{aligned} \tag{6.11}$$

Using a simple Euler-Maruyama method we can derive a numerical solution of (6.11) *and we plot a corresponding sample path, see Figure 6.2.*

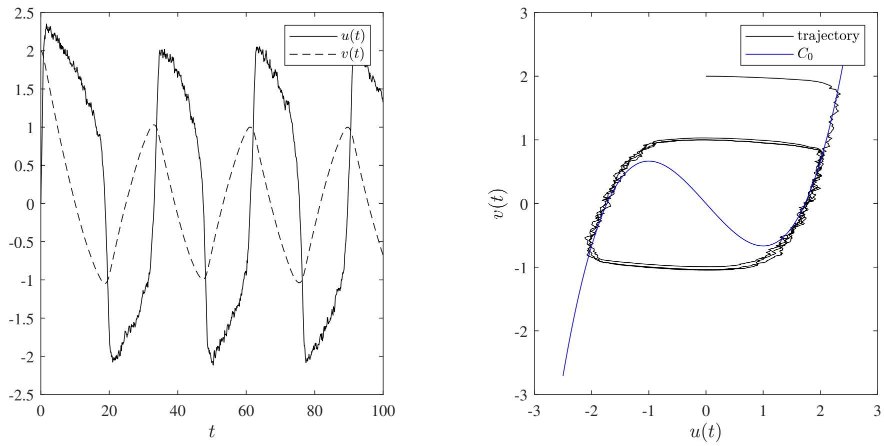

Figure 6.2: Numerical solution of (6.11) with $\varepsilon = 0.1$ and initial conditions $u(0) = 0$, $v(0) = 2$ using a simple Euler-Maruyama method.

A fundamental theory to rigourously analyse the behaviour of sample paths for stochastic fast-slow systems was developed by Berglund and Gentz [BG06, BG03, BG02]. In their work they not only analyse the dynamics near an attracting normally hyperbolic manifolds, but they also treat the cases where the fast subsystem admits *bifurcation points* or *periodic orbits*. Here, we focus on the normally hyperbolic attracting case and in the following we will briefly summarize the main result by Berglund and Gentz in this setting, following [BG06, Section 5].

In the case where the critical set of the associated deterministic system is a normally hyperbolic attracting manifold, Fenichel's Theorem yields the existence of a locally invariant slow manifold $\mathcal{C}_\varepsilon$ that is normally hyperbolic attracting as well. In this setting and under sufficiently small noise, it can be observed that a typical sample path of (6.10) starting near $\mathcal{C}_\varepsilon$ is going to fluctuate around $\mathcal{C}_\varepsilon$ and also slowly drifts according to a stochastic perturbation of the slow subsystem. We will make this precise in the following, whereby we closely follow [BG06, Section 5].

Let us fix the following assumptions for the system (6.10).

Assumptions 6.12.

(i) *Regularity*: Let $D \subset \mathbb{R}^m \times \mathbb{R}^n$ be open and $f \in C^2(D, \mathbb{R}^m)$, $g \in C^2(D, \mathbb{R}^n)$, $F \in C^1(D, \mathbb{R}^{m\times k})$ and $G \in C^1(D, \mathbb{R}^{n\times k})$. Furthermore, f, g, F, G and all their existing derivatives are uniformly bounded in D by a constant M.

(ii) *Critical manifold*: There exists a connected open subset $D_0 \subset \mathbb{R}^n$ and $u^* \in C(D_0, \mathbb{R}^m)$ such that

$$C_0 = \{(u, v) \in D : u = u^*(v), v \in D_0\}$$

is a critical manifold of the corresponding deterministic system.

(iii) *Stability*: The critical manifold C_0 is normally hyperbolic attracting.

(iv) *Non-degeneracy*: The matrix $FF^\top$ is positive definite.

By Fenichel's Theorem there exists a locally invariant manifold of the deterministic system

$$C_\varepsilon = \{(u,v) \in D : u = \overline{u}(v,\varepsilon) := u^*(v) + \mathcal{O}(\varepsilon), v \in D_0\}.$$

In the following we will construct a suitable neighbourhood around C_ε. Using Itô's formula one can derive a SODE for the evolution of the deviation $\xi = u - \overline{u}(v,\varepsilon)$ of sample paths from the slow manifold [BG06, Equation (5.1.5)]. The following is a linear approximation of this SODE, where v is replaced by its deterministic version v^d

$$\begin{array}{rcl} \mathrm{d}\xi^0 &=& \frac{1}{\varepsilon}A(v^{\mathrm{d}},\varepsilon)\xi^0 \,\mathrm{d}\tau + \frac{\sigma_f}{\sqrt{\varepsilon}}F_0(v^{\mathrm{d}},\varepsilon)\,\mathrm{d}B, \\ \mathrm{d}v^{\mathrm{d}} &=& g(\bar{u}(v^{\mathrm{d}},\varepsilon),v^{\mathrm{d}})\,\mathrm{d}\tau, \end{array} \tag{6.12}$$

with

$$A(v,\varepsilon) = \mathrm{D}_u f(\bar{u}(v,\varepsilon),v) - \varepsilon \mathrm{D}_v \bar{u}(v,\varepsilon)\mathrm{D}_u g(\bar{u}(v,\varepsilon),v),$$
$$F_0(v,\varepsilon) = F(\bar{u}(v,\varepsilon),v) - \rho\sqrt{\varepsilon}\mathrm{D}_v \bar{u}(v,\varepsilon)G(\bar{u}(v,\varepsilon),v).$$

With an initial condition $(\xi^0(0), v^{\mathrm{d}}(0)) = (0, v^{\mathrm{d}}(0))$ the solution to the first equation in (6.12) is given by the stochastic convolution

$$\xi^0(\tau) = \frac{\sigma_f}{\sqrt{\varepsilon}} \int_0^\tau U(\tau,s)F_0(v^{\mathrm{d}},\varepsilon)\,\mathrm{d}B(s),$$

where $U(\tau,s)$ denotes the principal solution to $\dot{\xi} = \frac{1}{\varepsilon}A(v^{\mathrm{d}},\varepsilon)\xi$. Now, $\xi^0(\tau)$ is Gaussian with zero mean and covariance $\mathrm{Cov}(\xi^0(\tau))$. In order to define a suitable neighbourhood within which the solution stays with high probability, it seems reasonable to use this covariance as an indicator for admissible deviations from the invariant manifold. In particular, it can be shown that the scaled covariance $X(\tau) := \frac{1}{\sigma_f^2}\mathrm{Cov}(\xi^0(\tau))$ is a solution to the deterministic fast-slow system

$$\begin{array}{rcl} \varepsilon\frac{\mathrm{d}X}{\mathrm{d}\tau} &=& A(v,\varepsilon)X + XA(v,\varepsilon)^\top + F_0(v,\varepsilon)F_0(v,\varepsilon)^\top, \\ \frac{\mathrm{d}v}{\mathrm{d}\tau} &=& g(\bar{u}(v,\varepsilon),v). \end{array} \tag{6.13}$$

As, for ε sufficiently small, the eigenvalues of $A(v,\varepsilon)$ have strictly negative real parts by Assumptions 6.12 (iii), [BG06, Lem. 5.12] ensures that the critical manifold of (6.13) is normal hyperbolic attracting and it is given by $\{X^*(v,\varepsilon), v \in D_0\}$ with

$$X^*(v,\varepsilon) = \int_0^\infty \exp(sA(v,\varepsilon))F_0(v,\varepsilon)F_0(v,\varepsilon)^\top \exp(sA(v,\varepsilon)^\top)\,\mathrm{d}s.$$

Again, Fenichel's theorem implies the existence of an invariant manifold

$$\{\bar{X}(v,\varepsilon), v \in D_0\},$$

with

$$\bar{X}(v,\varepsilon) = X^*(v,\varepsilon) + \mathcal{O}(\varepsilon).$$

Assumption 6.12 (iv) guarantees that $\bar{X}(v,\varepsilon)$ is invertible and thus we can define the following set, describing a neighbourhood around the slow manifold

$$\mathcal{B}(r) := \left\{(u,v) \in D : \langle u - \bar{u}(v,\varepsilon), \bar{X}(v,\varepsilon)^{-1}(u - \bar{u}(v,\varepsilon))\rangle < r^2, v \in D_0\right\}. \quad (6.14)$$

Definition 6.13. We define the following *exit times*

$$\begin{aligned} \tau_{\mathcal{B}(r)} &:= \inf\{\tau > 0 : (u(\tau), v(\tau)) \notin \mathcal{B}(r)\}, \\ \tau_{D_0} &:= \inf\{\tau > 0 : v(\tau) \notin D_0\}. \end{aligned}$$

Berglund and Gentz proved that sample paths are likely to stay for exponential long time spans within $\mathcal{B}(r)$, see Figure 6.3 for an illustration. More precisely, they have proved the following theorem.

Theorem 6.14 (cf. [BG06, Theorem 5.1.6])**.** *Let $u(0) = \overline{u}(v(0),\varepsilon)$ and denote by $\mathbb{P}^{u(0),v(0)}$ the law of the process $(u(\tau), v(\tau))_{\tau \geq 0}$ starting in $(u(0), v(0))$ at time $\tau = 0$. Under Assumptions 6.12 there exist constants $\varepsilon_0, \Delta_0, r_0, c, c_1, L > 0$ such that for all $\varepsilon \leq \varepsilon_0$, $\Delta \leq \Delta_0$, $r \leq r_0$, all $\gamma \in (0,1)$ and all $\tau \geq 0$*

$$\mathbb{P}^{u(0),v(0)}\left(\tau_{\mathcal{B}(r)} < \tau \wedge \tau_{\mathcal{D}_0}\right) \leq C_{r/\sigma_f,m,n,\gamma,\Delta}(\tau,\varepsilon) e^{-\kappa r^2/(2\sigma_f^2)}, \quad (6.15)$$

where

$$\kappa := \gamma\left[1 - c_1(r + \Delta + n\varepsilon\rho^2\sigma_f^2/r^2 + e^{-c/\varepsilon}/(1-\gamma))\right],$$

and

$$C_{r/\sigma_f,m,n,\gamma,\Delta}(\tau,\varepsilon) := L\frac{(1+\tau)^2}{\Delta\varepsilon}\left[(1-\gamma)^{-m} + e^{m/4} + e^{n/4}\right]\left(1 + r^2/\sigma_f^2\right).$$

Remark 6.15. The proof of Theorem 6.14 starts with a change of variables in (6.10) to ξ and v^d and a Taylor expansion of the resulting system. Then, precise estimates for the linear terms of the representation of the solution are derived, while the non-linear terms are treated as small perturbations. Another key ingredient is to consider the dynamics first on small time intervals where stochastic convolutions can be approximated by Gaussian martingales before glueing short time approximations to larger scales via the Markov property.

Remark 6.16. Clearly, estimate (6.15) is useless for an infinite-dimensional system, as the constant $C_{r/\sigma_f,m,n,\gamma,\triangle}$ would explode for n or m going to infinity. Thus, a direct transfer of this result to an SPDE interpreted as an infinite-dimensional SODE system fails.

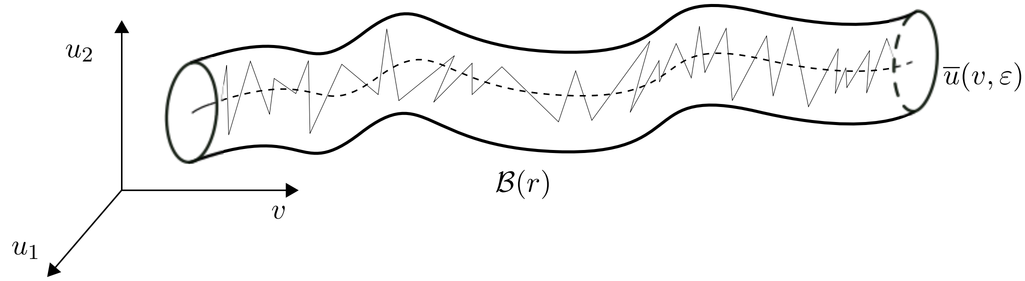

Figure 6.3: Sketch of a sample path inside an ellipsoidal neighbourhood $\mathcal{B}(r)$ in the finite-dimensional SODE setting with two fast and one slow variable. In the SPDE setting, we are going to view the u-direction as infinite-dimensional.

A key estimate in the proof of Berglund and Gentz is given by the following *Bernstein-type inequality*, which we will also need in the SPDE setting.

Proposition 6.17 (see [BG06, Lemma B.1.3])**.** *Let $\varphi:\mathbb{R}\to\mathbb{R}$ be Borel measurable and such that*

$$\Phi(t) := \int_0^t \varphi(\tau)^2 \, \mathrm{d}\tau$$

exists. Furthermore, let $B(\tau)$ be a real-valued Brownian motion. Then the following estimate holds:

$$\mathbb{P}\left(\sup_{0\le s\le t}\int_0^s \varphi(\tau)\mathrm{d}B(\tau)\ge c\right)\le \exp\left(-\frac{c^2}{2\Phi(t)}\right). \tag{6.16}$$

Proof. Using Itô's formula one can derive that for any $\gamma>0$

$$M_s = \exp\left(\int_0^s \varphi(\tau)\mathrm{d}B(\tau) - \frac{1}{2}\int_0^s \gamma^2\varphi(\tau)^2\mathrm{d}\tau\right),$$

is an exponential martingale with $\mathbb{E}\left[M_t\right]=1$. Then Doob's martingale inequality (see Theorem A.4) yields

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\le s\le t}\int_0^s \varphi(\tau)\mathrm{d}B(\tau)\ge c\right)\\
&\le \mathbb{P}\left(\sup_{0\le s\le t} M_s \ge \exp\left(c\gamma - \frac{1}{2}\int_0^t \gamma^2\varphi(\tau)^2\mathrm{d}\tau\right)\right)\\
&\le \exp\left(-c\gamma + \frac{1}{2}\int_0^t \gamma^2\varphi(\tau)^2\mathrm{d}\tau\right),
\end{aligned}$$

and the right hand side is maximal for the choice $\gamma = c/\Phi(t)$. □

6.4 Sample path estimates for fast-slow SPDE systems

As mentioned in Section 6.1, we would like to extend the sample path theory from the previous section to fast-slow SPDE systems. We will consider a linear setting as described below. Note that this section is based on [GKP19], which is joint work with Manuel Gnann and Christian Kuehn, and all the technical calculations are copied from there.

6.4.1 The linear setting

We consider the Hilbert space $\mathcal{H} := L^2([0,L])$ for some $L > 0$ and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We will focus on the following special case of equation (6.1): we choose $A = \frac{\partial^2}{\partial x^2}$, $g \equiv 1$ and $f(u,s,\varepsilon) = a(s)u(s) + Bu(s)$, with $a : \mathbb{R} \to \mathbb{R}$ being bounded and measurable and B being a non-local operator as defined below. After changing to the *slow time scale* s, we arrive at the following equation

$$\mathrm{d}u(s) = \frac{1}{\varepsilon}\left[\frac{\partial^2}{\partial x^2}u(s) + a(s)u(s) + Bu(s)\right] \,\mathrm{d}s + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W(s), \tag{6.17}$$

which is interpreted as a linear evolution equation in $\mathcal{H}$. Note that W is a $\mathcal{H}$-valued Q-Wiener process. We equip the equation with homogeneous Dirichlet boundary conditions, i.e, we have $u(s,0) = u(s,L) = 0$ for all $s \in [0,t]$, for some $t > 0$. Furthermore, we assume as initial condition $u(0,x) = 0$ for all $x \in [0,L]$.

The differential operator The domain of the operator $A = \frac{\partial^2}{\partial x^2}$ in $L^2([0,L])$ under homogeneous Dirichlet boundary conditions is given by

$$\mathcal{D}(A) = H^2([0,L]) \cap H_0^1([0,L]).$$

$\mathcal{D}(A)$ is dense in $L^2([0,L])$ and A has a complete orthonormal set of eigenfunctions

$$\{\phi_k\}_k := \left\{x \mapsto \sqrt{\frac{2}{L}}\sin\left(\frac{k\pi}{L}x\right)\right\}_k, \quad k \geq 1,$$

in $L^2([0,L])$ with associated eigenvalues $\mu_k := -\frac{k^2\pi^2}{L^2}$ for $k \geq 1$. In particular, A generates a strongly continuous semigroup $(S(s))_{s\geq 0}$, $S(s) = \mathrm{e}^{sA}$, in $L^2([0,L])$.

The noise term Let $\{e_k\}_k$ denote the eigenfunctions of the trace-class covariance operator $Q : \mathcal{H} \to \mathcal{H}$ with associated non-negative eigenvalues $\{\lambda_k\}_k$, so that

$$Qe_k = \lambda_k e_k \quad \text{for all } k \geq 1. \tag{6.18}$$

We assume that the eigenvalues are ordered as $\lambda_1 \geq \lambda_2 \geq ....$ The eigenfunctions form a complete orthonormal system of $\mathcal{H}$. By Proposition 3.13 we have the representation

$$W(s) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} B_k(s) e_k, \tag{6.19}$$

where $\{B_k\}_k$ is a sequence of independent real-valued Brownian motions on $(\Omega, \mathcal{F}, \mathbb{P})$. In the following we assume that there exist $c > 0$ and $p > 1$ such that

$$\lambda_k \leq ck^{2-p}.$$

In particular, the relation between the eigenvalues of A ($|\mu_k| \sim k^2$) and the eigenvalues of the trace-class covariance operator Q (which decrease monotonically to zero as $k \to \infty$, since the trace is finite) guarantees that the deterministic decay towards zero of the drift term dominates the noisy fluctuations in the higher modes.

The non-local operator B The operator B is defined via its action on the eigenfunctions $\{\phi_k\}_k$ in the following way for some fixed k_*

$$B\phi_k(x) = \begin{cases} \sum_{\ell=1}^{k_*-1} b_k^\ell \phi_\ell(x) & , k \leq k_* - 1, \\ 0 & , k \geq k_*. \end{cases} \tag{6.20}$$

Mild solution The operator $\mathcal{L}(s) := \frac{1}{\varepsilon}\left[\frac{\partial^2}{\partial x^2} + a(s) + B\right]$ generates a *strongly continuous evolution family* $(R(s,r))_{0 \leq r \leq s \leq t}$ and (6.17) admits a unique mild solution in $L^2([0,L])$ (cf. [Ver10], where a cylindrical Wiener process is considered) given by the stochastic convolution

$$u(s) = \frac{\sigma}{\sqrt{\varepsilon}} \int_0^s R(s,r) \, \mathrm{d}W(r). \tag{6.21}$$

Remark 6.18. Note that due to the explicit time dependence of the integrand, (6.21) is not a martingale. Thus one can not directly apply suitable martingale inequalities to bound the solution. In the one-dimensional setting by Berglund and Gentz [BG06, Chapter 3] this problem is tackled by approximating the solution locally by Gaussian martingales, see [BG06, Proposition 3.1.5]. However, their procedure is based on a certain splitting of the one-dimensional semigroup, which is not valid for a generic operator-valued evolution family. Thus, a direct transfer of the one-dimensional result to the Hilbert space norm of the SPDE solution is not possible.

6.4.2 Spectral Galerkin approximation

In this subsection, we derive a finite-dimensional approximation of (6.17). For this we make the following additional assumption.

Assumptions 6.19. The operators A and Q commute.

We use the *spectral Galerkin approximation* as introduced in Subsection 3.3. That is, we consider the orthogonal projections onto the the m-dimensional space $V_m := \mathrm{span}\{\phi_1, ..., \phi_m\}$ for some $m \geq k_*$. Note that by Assumptions 6.19 we can assume $\phi_k = e_k$ for all k. According to (3.17) the m-dimensional Galerkin approximation $u_m(s)$ of (6.21) satisfies

$$\mathrm{d}u_m = \frac{1}{\varepsilon} P_m \left[\frac{\partial^2}{\partial x^2} u_m + a(s) u_m + B u_m \right] \mathrm{d}s + \frac{\sigma}{\sqrt{\varepsilon}} P_m \, \mathrm{d}W. \tag{6.22}$$

Then, under suitable assumptions, $u_m(s)$ converges in the L^∞-topology to $u(s)$ as $m \to \infty$ (see [BJ13] for the details regarding this convergence). We calculate for

$$P_m u(s) = \sum_{k=1}^{m} \hat{u}_k(s) \phi_k, \qquad \hat{u}_k(s) := \int_0^L \phi_k(x) u(s,x) \, \mathrm{d}x,$$

by using integration by parts

$$\begin{aligned} \int_0^L \phi_k(x) \frac{\partial^2}{\partial x^2} u(s,x) \, \mathrm{d}x &= \int_0^L \frac{\partial^2}{\partial x^2} \phi_k(x) u(s,x) \, \mathrm{d}x \\ &= \mu_k \int_0^L \phi_k(x) u(s,x) \, \mathrm{d}x = \mu_k \hat{u}_k(s), \end{aligned}$$

hence we obtain

$$P_m \frac{\partial^2}{\partial x^2} u(s) = \sum_{k=1}^{m} \mu_k \hat{u}_k(s) \phi_k.$$

Similarly, we also project the noise term using (6.19)

$$P_m W(s) = \sum_{k=1}^{m} \sqrt{\lambda_k} B_k(s) \phi_k,$$

as

$$\begin{aligned} \int_0^L \phi_k(x) \sum_{j=1}^{\infty} \sqrt{\lambda_j} B_j(s) \phi_j(x) \, \mathrm{d}x &= \sum_{j=1}^{\infty} \sqrt{\lambda_j} B_j(s) \int_0^L \phi_k(x) \phi_j(x) \, \mathrm{d}x \\ &= \sqrt{\lambda_k} B_k(s). \end{aligned}$$

Furthermore, the non-autonomous part of the drift term gives

$$P_m a(s) u(s) = \sum_{k=1}^{m} a(s) \hat{u}_k(s) \phi_k,$$

and

$$P_m B u(s) = \sum_{\ell=1}^{k_*-1} \sum_{k=1}^{k_*-1} b_k^\ell \hat{u}_k(s) \phi_\ell.$$

In total, (6.22) is equivalent to the following finite-dimensional system of SODEs

$$\begin{pmatrix} \mathrm{d}U_1(s) \\ \mathrm{d}U_2(s) \end{pmatrix} = \frac{1}{\varepsilon} \underbrace{\begin{pmatrix} J_1(s) & 0_1 \\ 0_2 & J_2(s) \end{pmatrix}}_{=:J(s)} \underbrace{\begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix}}_{=:U_m(s)} \mathrm{d}s + \frac{\sigma}{\sqrt{\varepsilon}} \begin{pmatrix} F_1 & 0_1 \\ 0_2 & F_2 \end{pmatrix} \begin{pmatrix} \mathrm{d}\mathbf{B}_1(s) \\ \mathrm{d}\mathbf{B}_2(s) \end{pmatrix}, \tag{6.23}$$

where $0_1 \in \mathbb{R}^{(k_*-1)\times(m-k_*+1)}$ and $0_2 \in \mathbb{R}^{(m-k_*+1)\times(k_*-1)}$ are matrices filled with zeros,

$$J_1(s) := \begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mu_{k_*-1} \end{pmatrix} + a(s)\, \mathrm{id}_{k_*-1} + \underbrace{\begin{pmatrix} b_1^1 & b_2^1 & \dots & b_{k_*-1}^1 \\ b_1^2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_{k_*-1}^{k_*-2} \\ b_1^{k_*-1} & \dots & b_{k_*-2}^{k_*-1} & b_{k_*-1}^{k_*-1} \end{pmatrix}}_{=:\mathbf{B}}, \tag{6.24}$$

$$J_2(s) := \begin{pmatrix} \mu_{k_*} + a(s) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mu_m + a(s) \end{pmatrix}. \tag{6.25}$$

Furthermore, $U_1(s) := (\hat{u}_1(s), \dots, \hat{u}_{k_*-1}(s))^\top$, $U_2(s) := (\hat{u}_{k_*}(s), \dots, \hat{u}_m(s))^\top$,

$$F_1 := \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{k_*-1}}), \quad F_2 := \mathrm{diag}(\sqrt{\lambda_{k_*}}, \dots, \sqrt{\lambda_m}),$$

and

$$\mathbf{B}_1(s) = (B_1(s), \dots, B_{k_*-1}(s))^\top, \quad \mathbf{B}_2(s) = (B_{k_*}(s), \dots, B_m(s))^\top.$$

Here, $\mathrm{diag}(a_1, ..., a_n)$ denotes the $n \times n$ diagonal matrix with entries $a_1, ..., a_n$ on the diagonal and id_n is the $n \times n$ identity matrix.

From now on let the following assumptions hold.

Assumptions 6.20.

(i) $J_{i,j} \in C^1([0,t], \mathbb{R})$, for all $i, j = 1, ..., m$ and the derivatives are uniformly bounded by a constant M.

(ii) $a_- < a(s) < a_+$ for all $s \in [0,t]$ where $a_-, a_+ \in \mathbb{R}$.

(iii) $\mu_1 + a_+ + \|\mathbf{B}\|_{\mathrm{op}} =: -\overline{\kappa} < 0$.

(iv) $\lambda_k \neq 0$ for all $k = 1, ..., m$.

Note that the deterministic attracting slow manifold of (6.23) is given by $\mathcal{C}_\varepsilon = \{U_m(s) = 0\}$ for $s \in [0,t]$ since $U_m(s) \equiv 0$ solves the problem without noise for any $\varepsilon > 0$ and any $s \in [0,t]$, and Assumption 6.20 (iii) guarantees that $\mathcal{C}_\varepsilon$ is attracting.

6.4.3 Main result

We consider equation (6.17) on the spatial interval $[0, L]$, $L > 0$, and time interval $[0, t]$, where we assume for $\Lambda > 0$

$$t = \Lambda\varepsilon,$$

together with homogeneous Dirichlet boundary conditions and initial condition $u(0) = 0$, as discussed in Section 6.4.1. Let u be the mild solution to this problem and let u_m be the m-dimensional Galerkin approximation, as discussed in Section 6.4.2. Furthermore, let Assumptions 6.19 and 6.20 hold.

Notations It is helpful to introduce some notation to deal with various constants appearing in the following arguments. Let

$$C_1 := C_1(\gamma) := C\frac{\underline{\kappa} + \beta}{(k_* - 1)^3\sigma^2\lambda_1} \exp\left(2\Lambda(\overline{\kappa} - \underline{\kappa} - 2\beta)\right) \exp\left(-\frac{2\gamma}{\underline{\kappa}}(\underline{\kappa} + \beta)\right), \quad (6.26)$$

where $\gamma > 0$ is chosen arbitrarily, C, β are constants depending on the particular form of $\mathbf{B}$ and $\underline{\kappa}$, $\overline{\kappa}$ are lower and upper bounds on the eigenvalues of $J_1(s)$ (see Section 6.4.5 for details). Likewise, for arbitrary $\tilde{\gamma} > 0$ we define

$$C_2 = C_2(\tilde{\gamma}) := \frac{c\tilde{c}\exp\left(-2\tilde{\gamma}\right)\pi^2}{\sigma^2 L^2}, \quad (6.27)$$

where $c > 0$ is the constant such that $\frac{k^2}{\lambda_k} \geq ck^p$, and $\tilde{c} > 0$ is chosen such that $|\mu_k + a_+| \geq \tilde{c}|\mu_k|$. We also introduce the notation

$$H_*(k) := \frac{\ln\left(2\left\lceil\frac{\Lambda}{\tilde{\gamma}}|a_- + \mu_k|\right\rceil\right)}{C_2 k^p}, \quad (6.28)$$

and we note that, defining $H_*^m := H_*(k_* + m)$ for $m \in \mathbb{N}$, we have $\sum_{m=0}^{\infty} H_*^m < \infty$. Furthermore, we define

$$\eta_* := \sum_{m=0}^{\infty} H_*^m + \frac{2p}{(p^2 - 1)C_2} + 2\delta, \quad (6.29)$$

where δ is chosen such that $\frac{1}{C_2\delta} \geq 1$, and

$$\zeta_* := \frac{\ln\left(2\lceil\Lambda\underline{\kappa}/\gamma\rceil\right)}{C_1}, \quad (6.30)$$

as well as

$$\xi_* := \frac{1}{C_2\delta}\frac{\exp\left(-C_2 k_*^p\delta\right)}{1 - \exp\left(-C_2 p k_*^{p-1}\delta\right)}. \quad (6.31)$$

Main result Our main result then reads as follows

Theorem 6.21. *Let $\gamma, \tilde{\gamma} > 0$ be arbitrary. For $H \geq \eta_* + \zeta_*$ we have*

$$\begin{aligned}\mathbb{P}&\left(\sup_{0\leq s\leq t} \|u(s)\|^2 \geq H\right)\\ &\leq \exp\left(-C_1\left(H-\eta_*-\zeta_*\right)\right)\\ &\quad + \xi_* C_1 \frac{|\exp\left(-C_1(H-\eta_*-\zeta_*)\right) - \exp\left(-C_2 k_*^p (H-\eta_*-\zeta_*)\right)|}{|C_1 - C_2 k_*^p|},\end{aligned} \tag{6.32}$$

where $\|\cdot\| = \|\cdot\|_{L^2([0,L])}$ and all the constants are defined as above. The case $C_1 = C_2 k_*^p$ is to be understood in the sense of taking the derivative.

Remark 6.22. Theorem 6.21 tells us the following: For H large enough, the probability that the solution to equation (6.17) deviates more than H from the deterministic slow manifold $\mathcal{C}_\varepsilon$ within an ε-small time interval, is exponentially small in H. The larger p, i.e., the faster the eigenvalues of the covariance operator Q decrease, the smaller is the lower bound on H for which we can guarantee this exponential decay.

Remark 6.23. The left hand side of (6.32) is equivalent to the exit time probability $\mathbb{P}\left(\tau_{\mathcal{B}(H)} < t\right)$, where

$$\mathcal{B}(H) = \left\{u \in \mathcal{H} : \|u\|^2 < H\right\}.$$

Since we have not rescaled our neighbourhood by the noise process, as compared to the approach by Berglund and Gentz (see (6.14)), the principal eigenvalue λ_1 of the covariance operator appears on the right hand side of (6.32) (in the constant C_1).

Proof strategy Let us briefly outline the strategy of the proof. Note that by Parseval's identity we have for $H > 0$

$$\mathbb{P}\left(\sup_{0\leq s\leq t} \|u(s)\|^2 \geq H\right) = \mathbb{P}\left(\sup_{0\leq s\leq t} \sum_{k=1}^{\infty} |\hat{u}_k(s)|^2 \geq H\right). \tag{6.33}$$

For readability we write $u_k(s)$ for $\hat{u}_k(s)$ from now on. The main idea to prove Theorem 6.21 is to split the infinite sum in (6.33) into two parts, one containing the first k_*-1 components and the other one containing the last $m-k_*+1$ components, where we let m tend to infinity. We call the first sum the *finite-frequency part* and the second sum the *high-frequency part*. The two parts can be estimated as follows

Proposition 6.24. *For arbitrary $\gamma > 0$ we have for $H \geq \zeta_*$*

$$\mathbb{P}\left(\sup_{0\leq s\leq t} \sum_{k=1}^{k_*-1} |u_k(s)|^2 \geq H\right) \leq \exp\left(-C_1(H-\zeta_*)\right). \tag{6.34}$$

Proposition 6.25. *For arbitrary $\tilde{\gamma} > 0$ we have for $H \geq \eta_*$*

$$\mathbb{P}\left(\sum_{k=k_*}^{\infty} \sup_{0\leq s\leq t} |u_k(s)|^2 \geq H\right) \leq \xi_* \exp\left(-C_2 k_*^p (H - \eta_*)\right). \tag{6.35}$$

Proposition 6.24 will be proved in Section 6.4.5. In order to prove Proposition 6.25 we will use one-dimensional estimates for each component, which we will then combine iteratively, see Section 6.4.6. Finally, to prove Theorem 6.21 we will concatenate the estimates for the finite-frequency part and the high-frequency part, i.e., Propositions 6.24 and 6.25, see Section 6.4.7.

Remark 6.26. Note that similar estimates for one-dimensional and (finite) multi-dimensional SODE systems have been proved in [BG06]. We use a similar strategy for the proofs, however, in a way which is tailor-made for the setting at hand.

6.4.4 Auxiliary results

Before proving Propositions 6.24, 6.25 and Theorem 6.21 we provide a couple of auxiliary results. The following proposition will become crucial for obtaining estimates on the distribution of the sum of two random variables when exponential estimates for each individual random variable are given.

Proposition 6.27. *Let X, Y be two independent non-negative random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with absolutely continuous cumulative distribution functions. Assume that the following two estimates hold*

$$\mathbb{P}(X \geq H) \leq \sum_{i=0}^{n} \xi_X(i) \exp\left(-\kappa_X(i)(H - \eta_X)\right), \qquad \text{for all } H \geq \eta_X, \tag{6.36a}$$

$$\mathbb{P}(Y \geq H) \leq \xi_Y \exp\left(-\kappa_Y (H - \eta_Y)\right), \qquad \text{for all } H \geq \eta_Y, \tag{6.36b}$$

where $n \in \mathbb{N}$, $\xi_X(i), \xi_Y, \kappa_X(i), \kappa_Y > 0$, $\eta_X, \eta_Y \geq 0$ for $i = 0, \ldots, n$. Then, for $H \geq \eta_X + \eta_Y$ we have

$$\begin{aligned}
&\mathbb{P}(X + Y \geq H) \\
&\leq \mathbb{P}(Y \geq H - \eta_X)\left(1 - \sum_{i=0}^{n} \xi_X(i)\right) + \sum_{i=0}^{n} \xi_X(i) \exp\left(-\kappa_X(i)(H - \eta_X - \eta_Y)\right) \\
&\quad - \sum_{i=0}^{n} \xi_Y \xi_X(i) \kappa_X(i) \frac{\exp\left(-\kappa_X(i)(H - \eta_X - \eta_Y)\right) - \exp\left(-\kappa_Y (H - \eta_X - \eta_Y)\right)}{\kappa_X(i) - \kappa_Y},
\end{aligned} \tag{6.37}$$

where the case $\kappa_X(i) = \kappa_Y$ for an $i \in \{1, \ldots, n\}$ is to be understood in the sense of taking the derivative.

Proof. For simplicity we assume $\kappa_X(i) \neq \kappa_Y$ for all $i \in \{1, \dots, n\}$ in what follows. Further assume $H \geq \eta_X + \eta_Y$. As X and Y are independent we can use the convolution formula for the cumulative distribution function of the sum of two independent random variables, that is

$$\begin{aligned}
&\mathbb{P}(X+Y \geq H) \\
&= 1 - \mathbb{P}(X+Y < H) \\
&= 1 - \int_0^H \left(\frac{\mathrm{d}}{\mathrm{d}H_1} (1 - \mathbb{P}(Y \geq H_1)) \right) (1 - \mathbb{P}(X \geq H - H_1))\, \mathrm{d}H_1 \\
&= 1 + \int_0^H \frac{\mathrm{d}}{\mathrm{d}H_1} \mathbb{P}(Y \geq H_1)\, \mathrm{d}H_1 - \int_0^H \left(\frac{\mathrm{d}}{\mathrm{d}H_1} \mathbb{P}(Y \geq H_1) \right) \mathbb{P}(X \geq H - H_1)\, \mathrm{d}H_1 \\
&= \mathbb{P}(Y \geq H) - \int_0^H \left(\frac{\mathrm{d}}{\mathrm{d}H_1} \mathbb{P}(Y \geq H_1) \right) \mathbb{P}(X \geq H - H_1)\, \mathrm{d}H_1.
\end{aligned}$$

Using that $\frac{\mathrm{d}}{\mathrm{d}H_1} \mathbb{P}(Y \geq H_1) \leq 0$ and (6.36a), we calculate further

$$\begin{aligned}
&\mathbb{P}(X+Y \geq H) \\
&= 1 - \mathbb{P}(X+Y < H) \\
&\leq \mathbb{P}(Y \geq H) \\
&\quad - \int_0^{H-\eta_X} \left(\frac{\mathrm{d}}{\mathrm{d}H_1} \mathbb{P}(Y \geq H_1) \right) \left[\sum_{i=0}^n \xi_X(i) \exp\left(-\kappa_X(i)(H - H_1 - \eta_X)\right) \right] \mathrm{d}H_1 \\
&\quad - \int_{H-\eta_X}^H \frac{\mathrm{d}}{\mathrm{d}H_1} \mathbb{P}(Y \geq H_1)\ \mathrm{d}H_1 \\
&= \mathbb{P}(Y \geq H - \eta_X) \\
&\quad \underbrace{- \int_0^{H-\eta_X} \left(\frac{\mathrm{d}}{\mathrm{d}H_1} \mathbb{P}(Y \geq H_1) \right) \left[\sum_{i=0}^n \xi_X(i) \exp\left(-\kappa_X(i)(H - H_1 - \eta_X)\right) \right] \mathrm{d}H_1}_{=:I_1}.
\end{aligned} \tag{6.38}$$

With integration by parts and by applying equation (6.36b), we can estimate I_1 further as follows

$$
\begin{aligned}
I_1 = &- \left[\mathbb{P}\left(Y \geq H_1\right) \sum_{i=0}^{n} \xi_X(i) \exp\left(-\kappa_X(i)(H - H_1 - \eta_X)\right) \right]_{H_1=0}^{H-\eta_X} \\
&+ \int_0^{H-\eta_X} \mathbb{P}\left(Y \geq H_1\right) \left[\sum_{i=0}^{n} \xi_X(i)\kappa_X(i) \exp\left(-\kappa_X(i)(H - H_1 - \eta_X)\right) \right] \, \mathrm{d}H_1 \\
\leq &-\mathbb{P}\left(Y \geq H - \eta_X\right) \sum_{i=0}^{n} \xi_X(i) + \sum_{i=0}^{n} \xi_X(i) \exp\left(-\kappa_X(i)(H - \eta_X)\right) \\
&+ \int_0^{\eta_Y} \sum_{i=0}^{n} \xi_X(i)\kappa_X(i) \exp\left(-\kappa_X(i)(H - H_1 - \eta_X)\right) \mathrm{d}H_1 \\
&+ \int_{\eta_Y}^{H-\eta_X} \xi_Y \exp\left(-\kappa_Y(H_1 - \eta_Y)\right) \\
&\quad \left[\sum_{i=0}^{n} \xi_X(i)\kappa_X(i) \exp\left(-\kappa_X(i)(H - H_1 - \eta_X)\right) \right] \, \mathrm{d}H_1 \\
= &-\mathbb{P}\left(Y \geq H - \eta_X\right) \sum_{i=0}^{n} \xi_X(i) + \sum_{i=0}^{n} \xi_X(i) \exp\left(-\kappa_X(i)(H - \eta_X - \eta_Y)\right) + I_2,
\end{aligned}
$$

with

$$
\begin{aligned}
I_2 := &\int_{\eta_Y}^{H-\eta_X} \xi_Y \exp\left(-\kappa_Y(H_1 - \eta_Y)\right) \\
&\quad \left[\sum_{i=0}^{n} \xi_X(i)\kappa_X(i) \exp\left(-\kappa_X(i)(H - H_1 - \eta_X)\right) \right] \, \mathrm{d}H_1 \\
= &\sum_{i=0}^{n} \xi_Y \xi_X(i)\kappa_X(i) \exp\left(\kappa_Y \eta_Y - \kappa_X(i) H + \kappa_X(i)\eta_X\right) \\
&\quad \int_{\eta_Y}^{H-\eta_X} \exp\left(H_1(\kappa_X(i) - \kappa_Y)\right) \, \mathrm{d}H_1 \\
= &-\sum_{i=0}^{n} \xi_Y \xi_X(i)\kappa_X(i) \frac{\exp\left(-\kappa_X(i)(H - \eta_X - \eta_Y)\right) - \exp\left(-\kappa_Y(H - \eta_X - \eta_Y)\right)}{\kappa_X(i) - \kappa_Y}.
\end{aligned}
$$

Inserting I_2 into I_1 and estimating I_1 as above in equation (6.38) concludes the proof. □

We will further need the following results for the iteration step in the high-frequency estimate.

Lemma 6.28. *For $n \in \mathbb{N}$ let $\{x_k\}_{k=0}^n$ be distinct non-negative real numbers. Then*

$$\sum_{i=0}^{n} \prod_{m=0,m\neq i}^{n} \frac{x_m}{x_m - x_i} = 1, \tag{6.39a}$$

$$\sum_{i=0}^{n} \prod_{m=0,m\neq i}^{n} \frac{1}{x_m - x_i} = 0. \tag{6.39b}$$

Proof of (6.39a). Define the auxiliary function $f : \mathbb{R} \to \mathbb{R}$

$$f(x) := \sum_{i=0}^{n} \prod_{m=0,m\neq i}^{n} \frac{1 - x/x_m}{1 - x_i/x_m}.$$

Then for $k = 0, ..., n$ we have

$$\begin{aligned} f(x_k) &= \sum_{i=0}^{n} \prod_{m=0,m\neq i}^{n} \frac{1 - x_k/x_m}{1 - x_i/x_m} \\ &= \sum_{i=0,i\neq k}^{n} \prod_{m=0,m\neq i}^{n} \frac{1 - x_k/x_m}{1 - x_i/x_m} + \prod_{m=0,m\neq k}^{n} \frac{1 - x_k/x_m}{1 - x_k/x_m} \\ &= \sum_{i=0,i\neq k}^{n} \left(\prod_{m=0,m\neq i,m\neq k}^{n} \frac{1 - x_k/x_m}{1 - x_i/x_m} \right) \underbrace{\frac{1 - x_k/x_k}{1 - x_i/x_p}}_{=0} +1 = 1. \end{aligned}$$

Now, $f(x) - 1$ is a polynomial of degree n with $n+1$ roots, i.e., $f(x) - 1 \equiv 0$. Hence,

$$1 = f(0) = \sum_{i=0}^{n} \prod_{m=0,m\neq i}^{n} \frac{1}{1 - x_i/x_m} = \sum_{i=0}^{n} \prod_{m=0,m\neq i}^{n} \frac{x_m}{x_m - x_i}.$$

Proof of (6.39b). We prove the second identity by induction. For the base case $n = 1$ we have

$$\sum_{i=0}^{n} \prod_{m=0,m\neq i}^{n} \frac{1}{x_m - x_i} = \frac{1}{x_1 - x_0} + \frac{1}{x_0 - x_1} = 0.$$

Now, let (6.39b) hold for arbitrary but fixed $n \in \mathbb{N}$ (inductive hypothesis). Then

$$\begin{aligned}
&\sum_{i=0}^{n+1} \prod_{m=0,m\neq i}^{n+1} \frac{1}{x_m - x_i} \\
&= \frac{1}{x_{n+1} - x_n} \left[\sum_{i=0}^{n+1} \frac{x_{n+1} - x_i}{\prod_{m=0,m\neq i}^{n+1}(x_m - x_i)} - \sum_{i=0}^{n+1} \frac{x_n - x_i}{\prod_{m=0,m\neq i}^{n+1}(x_m - x_i)} \right] \\
&= \frac{1}{x_{n+1} - x_n} \left[\sum_{i=0}^{n} \frac{1}{\prod_{m=0,m\neq i}^{n}(x_m - x_i)} - \sum_{i=0,i\neq n}^{n+1} \frac{1}{\prod_{m=0,m\neq i,m\neq n}^{n+1}(x_m - x_i)} \right] \\
&= \frac{1}{x_{n+1} - x_n} [0 - 0] = 0,
\end{aligned}$$

where we have used the inductive hypothesis in the last line. □

Corollary 6.29. *For $k_* \in \mathbb{N}$, $a, b, c \in \mathbb{N}$ let us define the following quotient*

$$Q^a_{b,c} := \frac{(k_* + a)^p}{(k_* + b)^p - (k_* + c)^p}, \tag{6.40}$$

which will appear in the estimate for the high-frequency part. For $n \in \mathbb{N}$ it holds

$$\sum_{i=0}^{n} \prod_{n=0,m\neq i}^{n} Q^m_{m,i} = 1, \tag{6.41a}$$

$$\sum_{i=0}^{n+1} \prod_{m=0,m\neq i}^{n+1} Q^0_{m,i} = 0, \tag{6.41b}$$

$$1 - Q^i_{i,n+1} = Q^{n+1}_{n+1,i}, \tag{6.41c}$$

$$\sum_{i=0}^{n} \left(\prod_{m=0,m\neq i}^{n} Q^m_{m,i} \right) Q^i_{i,n+1} = \prod_{m=0}^{n} Q^m_{m,n+1}. \tag{6.41d}$$

Proof. Identities (6.41a) and (6.41b) follow directly from Lemma 6.28. The third identity (6.41c) can be easily verified by direct calculation

$$\begin{aligned}
1 - Q^i_{i,n+1} &= \frac{(k_* + i)^p - (k_* + n + 1)^p}{(k_* + i)^p - (k_* + n + 1)^p} - \frac{(k_* + i)^p}{(k_* + i)^p - (k_* + n + 1)^p} \\
&= \frac{-(k_* + n + 1)^p}{(k_* + i)^p - (k_* + n + 1)^p} = Q^{n+1}_{n+1,i}.
\end{aligned}$$

The identity (6.41d) follows from

$$\begin{aligned}
&\sum_{i=0}^{n}\left(\prod_{m=0,m\neq i}^{n} Q_{m,i}^{m}\right) Q_{i,n+1}^{i} \\
&= \prod_{m=0}^{n}(k_*+m)^p\left[-\sum_{i=0}^{n}\prod_{m=0,m\neq i}^{n+1}\frac{1}{(k_*+m)^p-(k_*+i)^p}\right] \\
&= \prod_{m=0}^{n}(k_*+m)^p\Bigg[\underbrace{-\sum_{i=0}^{n+1}\prod_{m=0,m\neq i}^{n+1}\frac{1}{(k_*+m)^p-(k_*+i)^p}}_{=0 \text{ by } (6.41b)} \\
&\qquad + \prod_{m=0}^{n}\frac{1}{(k_*+m)^p-(k_*+n+1)^p}\Bigg] \\
&= \prod_{m=0}^{n}\frac{(k_*+m)^p}{(k_*+m)^p-(k_*+n+1)^p} = \prod_{m=0}^{n} Q_{m,n+1}^{m}.
\end{aligned}$$

□

6.4.5 Finite-frequency estimate

Proof of Proposition 6.24. We begin by estimating the eigenvalues of the matrix $J_1(s)$. Let $\psi(s)$ be an eigenvalue of $J_1(s) = \operatorname{diag}(\mu_1,\dots,\mu_{k_*-1}) + a(s)\mathrm{id}_{k_*-1} + \mathbf{B}$ with corresponding normalized eigenvector w ($\|w\|_2 = 1$), i.e. $J_1(s)w = \psi(s)w$. Note that $\|\cdot\|_2$ denotes the Euclidean norm of vectors in $\mathbb{R}^{k_*-1}$ and $\|\cdot\|_{\mathrm{op}}$ is the operator norm on the space $\mathbb{R}^{k_*-1\times k_*-1}$. We have

$$\begin{aligned}
\|\mathbf{B}\|_{\mathrm{op}} &\geq \|\mathbf{B}w\|_2 \\
&= \|\operatorname{diag}(\psi(s)-a(s)-\mu_1,\dots,\psi(s)-a(s)-\mu_{k_*-1})w\|_2 \\
&\geq \min_{k=1,\dots,k_*-1}|\psi(s)-a(s)-\mu_k|.
\end{aligned}$$

This estimate yields an upper and a lower bound on $\psi(s)$:

$$\begin{aligned}
\psi(s) &\leq a(s) + \max_{k=1,\dots,k_*-1}\mu_k + \|\mathbf{B}\|_{\mathrm{op}} \leq a_+ + \mu_1 + \|\mathbf{B}\|_{\mathrm{op}} =: -\overline{\kappa}, \\
\psi(s) &\geq a(s) + \min_{k=1,\dots,k_*-1}\mu_k - \|\mathbf{B}\|_{\mathrm{op}} \geq a_- + \mu_{k_*-1} - \|\mathbf{B}\|_{\mathrm{op}} =: -\underline{\kappa},
\end{aligned}$$

with $0 < \overline{\kappa} < \underline{\kappa}$ (cf. Assumption 6.20 (iii)). Now, let $U_1(s)$ be the solution to the $k_* - 1$-dimensional system

$$U_1(s) = \frac{1}{\varepsilon}J_1(s)U_1(s) + \frac{\sigma}{\sqrt{\varepsilon}}F_1\,\mathrm{d}\mathbf{B}_1(s).$$

Using Duhamel's principle the solution can be represented as follows

$$U_1(s) = \frac{\sigma}{\sqrt{\varepsilon}} \int_0^s \exp\left(\frac{1}{\varepsilon}\alpha(s,\tau)\right) F_1 \;\mathrm{d}\mathbf{B}_1(\tau),$$

with $\alpha(s,\tau) := \int_\tau^s J_1(r)\;\mathrm{d}r$. Furthermore, define $\alpha(s) := \alpha(s,0)$. Since we have an upper and a lower bound for the eigenvalues of $J_1(s)$, we can obtain the following estimates

$$\left\|\exp\left(\frac{1}{\varepsilon}\alpha(s)\right)\right\|_{\mathrm{op}} \leq \overline{C}\exp\left(-\frac{s}{\varepsilon}(\overline{\kappa}-\beta)\right), \tag{6.42}$$

$$\left(\exp\left(-\frac{1}{\varepsilon}\alpha(\tau)\right)\right)_{i,j} \leq \underline{C}\exp\left(\frac{\tau}{\varepsilon}(\underline{\kappa}+\beta)\right), \tag{6.43}$$

where the constant $\beta \geq 0$ comes from the polynomial part appearing in non-diagonalizable matrices and $\underline{C}, \overline{C}$ are time-independent constants.

Let us now introduce a partition of the time interval $[0,t]$ by $0 = s_0 < s_1 < \ldots < s_N = t$ with step size $s_{j+1} - s_j = \frac{\varepsilon\gamma}{\underline{\kappa}}$ and $N = \left\lceil \frac{t\underline{\kappa}}{\varepsilon\gamma} \right\rceil$, for some $\gamma > 0$. Using (6.42) we can estimate the probability as follows

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\leq s\leq t}\sum_{k=1}^{k_*-1}|u_k(s)|^2 \geq H\right)\\
&\leq \sum_{j=0}^{N-1}\mathbb{P}\left(\sup_{s_j\leq s\leq s_{j+1}}\left\|\frac{\sigma}{\sqrt{\varepsilon}}\int_0^s \exp\left(\frac{1}{\varepsilon}\alpha(s,\tau)\right)F_1\;\mathrm{d}\mathbf{B}_1(\tau)\right\|_2^2 \geq H\right)\\
&\leq \sum_{j=0}^{N-1}\mathbb{P}\left(\sup_{s_j\leq s\leq s_{j+1}}\frac{(k_*-1)^2\sigma^2}{\varepsilon}\max_{1\leq k\leq k_*-1}\lambda_k\left\|\exp\left(\frac{1}{\varepsilon}\alpha(s)\right)\right\|_{\mathrm{op}}^2\right.\\
&\qquad\left.\left\|\int_0^s\exp\left(-\frac{1}{\varepsilon}\alpha(\tau)\right)e_k\;\mathrm{d}B_k(\tau)\right\|_2^2 \geq H\right)\\
&\leq \sum_{j=0}^{N-1}\mathbb{P}\left(\sup_{s_j\leq s\leq s_{j+1}}\frac{(k_*-1)^3\sigma^2\lambda_1\overline{C}^2}{\varepsilon\exp\left(\frac{2}{\varepsilon}s(\overline{\kappa}-\beta)\right)}\right.\\
&\qquad\left.\max_{1\leq k,\ell\leq k_*-1}\left|\int_0^s\left(\exp\left(-\frac{1}{\varepsilon}\alpha(\tau)\right)e_k\right)_\ell\;\mathrm{d}B_k(\tau)\right|^2 \geq H\right)\\
&\leq \sum_{j=0}^{N-1}\max_{1\leq k,\ell\leq k_*-1}\mathbb{P}\left(\sup_{s_j\leq s\leq s_{j+1}}\left|\int_0^s\left(\exp\left(-\frac{1}{\varepsilon}\alpha(\tau)\right)e_k\right)_\ell\;\mathrm{d}B_k(\tau)\right|^2\right.\\
&\qquad\left.\geq \frac{H\varepsilon\exp\left(\frac{2}{\varepsilon}s_j(\overline{\kappa}-\beta)\right)}{(k_*-1)^3\sigma^2\lambda_1\overline{C}^2}\right).
\end{aligned}$$

Applying the Bernstein inequality from Proposition 6.17 and using (6.43), we calculate further

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\le s\le t}\sum_{k=1}^{k_*-1}|u_k(s)|^2\ge H\right)\\
&\le\sum_{j=0}^{N-1}\max_{1\le k,\ell\le k_*-1}2\exp\left(-\frac{H\varepsilon}{(k_*-1)^3\sigma^2\lambda_1\overline{C}^2}\frac{\exp\left(\frac{2}{\varepsilon}s_j(\overline{\kappa}-\beta)\right)}{2\int_0^{s_{j+1}}\left(\exp\left(-\frac{1}{\varepsilon}\alpha(\tau)\right)e_k\right)_\ell^2\,\mathrm{d}\tau}\right)\\
&\le\sum_{j=0}^{N-1}2\exp\left(-C\frac{H\varepsilon}{(k_*-1)^3\sigma^2\lambda_1}\exp\left(\frac{2}{\varepsilon}s_j(\overline{\kappa}-\beta)\right)\frac{1}{2\int_0^{s_{j+1}}\exp\left(\frac{2\tau}{\varepsilon}(\underline{\kappa}+\beta)\right)\mathrm{d}\tau}\right)\\
&\le\sum_{j=0}^{N-1}2\exp\left(-C\frac{H(\underline{\kappa}+\beta)}{(k_*-1)^3\sigma^2\lambda_1}\exp\left(\frac{2}{\varepsilon}s_j(\overline{\kappa}-\underline{\kappa}-2\beta)\right)\exp\left(-\frac{2\gamma}{\underline{\kappa}}(\underline{\kappa}+\beta)\right)\right)\\
&\le 2\left\lceil\frac{\Lambda\underline{\kappa}}{\gamma}\right\rceil\exp\left(-C\frac{H(\underline{\kappa}+\beta)}{(k_*-1)^3\sigma^2\lambda_1}\exp\left(2\Lambda(\overline{\kappa}-\underline{\kappa}-2\beta)\right)\exp\left(-\frac{2\gamma}{\underline{\kappa}}(\underline{\kappa}+\beta)\right)\right)\\
&=\exp\left(-C_1\left(H-\zeta_*\right)\right),
\end{aligned}$$

where $C=\frac{1}{\underline{C}^2\overline{C}^2}$ and C_1, ζ_* are defined in (6.26) and (6.30). □

6.4.6 High-frequency estimate

To obtain an estimate for the high-frequency part we are going to derive estimates for each component $u_k(s)$ with $k\ge k_*$ and then concatenate them via Proposition 6.27. First note that we have the following estimate for one single mode

Lemma 6.30. *For all $k\ge k_*$ we have*

$$\mathbb{P}\left(\sup_{0\le s\le t}|u_k(s)|^2\ge H\right)\le\exp\left(-C_2k^p(H-H_*(k))\right),\tag{6.44}$$

where $C_2, H_*(k)$ are defined in (6.27) and (6.28).

Proof. Let $k\ge k_*$, the equation for the k-th component reads

$$\mathrm{d}u_k(s)=\frac{1}{\varepsilon}(\mu_k+a(s))u_k(s)\,\mathrm{d}s+\frac{\sigma}{\sqrt{\varepsilon}}\sqrt{\lambda_k}\,\mathrm{d}B_k(s),\tag{6.45}$$

and using Duhamel's principle its solution can be represented as

$$u_k(s)=\frac{\sigma}{\sqrt{\varepsilon}}\int_0^s\exp\left(\frac{\alpha_k(s,\tau)}{\varepsilon}\right)\sqrt{\lambda_k}\,\mathrm{d}B_k(\tau),\tag{6.46}$$

with $\alpha_k(s,\tau) = \int_\tau^s (\mu_k + a(r))\ \mathrm{d}r$. We have the following estimate

$$\begin{aligned}
&\int_0^s \exp\left(\frac{2\alpha_k(s,\tau)}{\varepsilon}\right)\ \mathrm{d}\tau \\
&\leq \int_0^s \exp\left(\frac{2}{\varepsilon}\int_\tau^s (\mu_k + a_+)\ \mathrm{d}r\right)\ \mathrm{d}\tau \\
&= \int_0^s \exp\left(\frac{2}{\varepsilon}(\mu_k + a_+)\tau\right)\ \mathrm{d}\tau \\
&= \frac{\varepsilon}{2}\frac{1}{\mu_k + a_+}\left[\exp\left(\frac{1}{\varepsilon}(\mu_k + a_+)s\right) - 1\right] \\
&\leq \frac{\varepsilon}{2|\mu_k + a_+|}.
\end{aligned} \tag{6.47}$$

Now fix $\tilde{\gamma} > 0$ and we introduce a k-dependent partition $0 = s_0^k < s_1^k < \ldots < s_N^k = t$ of $[0,t] = [0, \varepsilon\Lambda]$ with $-\alpha_k(s_{j+1}^k, s_j^k) = \varepsilon\tilde{\gamma}$ for $0 \leq j < N_k = \left\lceil \frac{|\alpha_k(t)|}{\varepsilon\tilde{\gamma}} \right\rceil$. Then, using the Bernstein inequality (Proposition 6.17) and estimate (6.47), we obtain

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\leq s\leq t} |u_k(s)|^2 \geq H\right) \\
&= \mathbb{P}\left(\sup_{0\leq s\leq t}\left|\frac{\sigma}{\sqrt{\varepsilon}}\int_0^s \exp\left(\frac{\alpha_k(s,\tau)}{\varepsilon}\right)\sqrt{\lambda_k}\ \mathrm{d}B_k(\tau)\right| \geq \sqrt{H}\right) \\
&\leq \sum_{j=0}^{N_k-1} \mathbb{P}\left(\sup_{s_j^k\leq s\leq s_{j+1}^k}\left|\int_0^s \exp\left(-\frac{\alpha_k(\tau)}{\varepsilon}\right)\ \mathrm{d}B_k(\tau)\right|\right. \\
&\qquad\qquad \left.\geq \frac{\sqrt{H\varepsilon}}{\sigma\sqrt{\lambda_k}} \inf_{s_j^k\leq s\leq s_{j+1}^k} \exp\left(-\frac{\alpha_k(s)}{\varepsilon}\right)\right) \\
&\leq \sum_{j=0}^{N_k-1} 2\exp\left(-\frac{H\varepsilon}{\sigma^2\lambda_k}\frac{\inf_{s_j^k\leq s\leq s_{j+1}^k}\exp\left(-2\alpha_k(s)/\varepsilon\right)}{2\int_0^{s_{j+1}^k}\exp\left(-2\alpha_k(\tau)/\varepsilon\right)\ \mathrm{d}\tau}\right) \\
&\leq \sum_{j=0}^{N_k-1} 2\exp\left(-\frac{H\varepsilon}{2\sigma^2\lambda_k}\frac{\exp\left(2\alpha_k(s_{j+1}^k, s_j^k)/\varepsilon\right)}{\int_0^{s_{j+1}^k}\exp\left(2\alpha_k(s_{j+1}^k,\tau)/\varepsilon\right)\ \mathrm{d}\tau}\right) \\
&\leq \sum_{j=0}^{N_k-1} 2\exp\left(-\frac{H\varepsilon}{2\sigma^2\lambda_k}\exp\left(-2\tilde{\gamma}\right)\frac{2|\mu_k + a_+|}{\varepsilon}\right) \\
&\leq 2\left\lceil \frac{|\alpha_k(t)|}{\varepsilon\tilde{\gamma}} \right\rceil \exp\left(-\frac{H}{\sigma^2}\exp\left(-2\tilde{\gamma}\right)\frac{\tilde{c}\pi^2}{L^2}ck^p\right) \\
&= \exp\left(-C_2 k^p (H - H_*(k))\right)
\end{aligned}$$

where C_2 and $H_*(k)$ have been defined in (6.27) and (6.28). □

We are now going to prove an estimate on a finite sum of components with index $k \geq k_*$. This will be used to prove Proposition 6.25 by finding a bound independent of the number of addends n.

Proposition 6.31. *Let $n \in \mathbb{N}$. For $H \geq \sum_{m=0}^{n} H_*^m$ we have*

$$\mathbb{P}\left(\sum_{i=0}^{n} \sup_{0\leq s\leq t} |u_{k_*+i}(s)|^2 \geq H\right) \leq \sum_{i=0}^{n}\left[\exp\left(-C_2(k_*+i)^p\left(H-\sum_{m=0}^{n} H_*^m\right)\right)\prod_{m=0,m\neq i}^{n} Q_{m,i}^m\right], \tag{6.48}$$

where $Q_{m,i}^m$ has been defined in Corollary 6.29.

Proof. We prove the statement inductively. The base case $n = 0$ directly follows from Lemma 6.30. Now, let (6.48) hold for arbitrary but fixed n (inductive hypothesis). Note that $\sum_{i=0}^{n} \sup_{0\leq s\leq t} |u_{k_*+i}(s)|^2$ and $\sup_{0\leq s\leq t} |u_{k_*+n+1}(s)|^2$ are independent. Furthermore, by the inductive hypothesis we have for the sum the estimate given in equation (6.48) and for the (k_*+n+1)th component we have by Lemma 6.30

$$\mathbb{P}\left(\sup_{0\leq s\leq t} |u_{k_*+n+1}(s)|^2 \geq H\right) \leq \exp\left(-C_2(k_*+n+1)^p(H-H_*^{n+1})\right). \tag{6.49}$$

Now, applying Proposition 6.27 with $\xi_X(i) = \prod_{m=0,m\neq i}^{n} Q_{m,i}^m$, $\xi_Y = 1$, $\kappa_X(i) = C_2(k_*+i)^p$, $\kappa_Y = C_2(k_*+n+1)^p$, $\eta_X = \sum_{m=0}^{n} H_*^m$ and $\eta_Y = H_*^{n+1}$, where $i = 0...n$, yields for $H \geq \sum_{m=0}^{n} H_*^m + H_*^{n+1} = \sum_{m=0}^{n+1} H_*^m$

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=0}^{n+1} \sup_{0\leq s\leq t} |u_{k_*+i}(s)|^2 \geq H\right) \\
&= \mathbb{P}\left(\sum_{i=0}^{n} \sup_{0\leq s\leq t} |u_{k_*+i}(s)|^2 + \sup_{0\leq s\leq t} |u_{k_*+n+1}(s)|^2 \geq H\right) \\
&\leq \mathbb{P}\left(\sup_{0\leq s\leq t} |u_{k_*+n+1}(s)|^2 \geq H - H_*\right)\underbrace{\left(1-\sum_{i=0}^{n}\prod_{m=0,m\neq i}^{n} Q_{m,i}^m\right)}_{=0 \text{ by } (6.41a)} \\
&\quad + \exp\left(-C_2(k_*+n+1)^p\left(H-\sum_{m=0}^{n} H_*^m - H_*^{n+1}\right)\right)\sum_{i=0}^{n}\frac{\prod_{m=0,m\neq i}^{n} Q_{m,i}^m C_2(k_*+i)^p}{C_2(k_*+i)^p - C_2(k_*+n+1)^p} \\
&\quad + \sum_{i=0}^{n}\exp\left(-C_2(k_*+i)^p\left(H-\sum_{m=0}^{n} H_*^m - H_*^{n+1}\right)\right) \\
&\qquad\qquad \left(\prod_{m=0,m\neq i}^{n} Q_{m,i}^m - \frac{\prod_{m=0,m\neq i}^{n} Q_{m,i}^m C_2(k_*+i)^p}{C_2(k_*+i)^p - C_2(k_*+n+1)^p}\right).
\end{aligned}$$

The last line can be estimated further as follows

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=0}^{n+1}\sup_{0\le s\le t}|u_{k_*+i}(s)|^2\ge H\right)\\
&=\exp\left(-C_2(k_*+n+1)^p\left(H-\sum_{m=0}^{n+1}H_*^m\right)\right)\underbrace{\sum_{i=0}^{n}\prod_{m=0,m\neq i}^{n}Q_{m,i}^mQ_{i,n+1}^i}_{=\prod_{m=0}^{n}Q_{m,n+1}^m\text{ by (6.41}d)}\\
&\quad+\sum_{i=0}^{n}\exp\left(-C_2(k_*+i)^p\left(H-\sum_{m=0}^{n+1}H_*^m\right)\right)\prod_{m=0,m\neq i}^{n}Q_{m,i}^m\underbrace{\left(1-Q_{i,n+1}^i\right)}_{=Q_{n+1,i}^{n+1}\text{ by (6.41}c)}\\
&=\exp\left(-C_2(k_*+n+1)^p\left(H-\sum_{m=0}^{n+1}H_*^m\right)\right)\prod_{m=0,m\neq n+1}^{n+1}Q_{m,n+1}^m\\
&\quad+\sum_{i=0}^{n}\exp\left(-C_2(k_*+i)^p\left(H-\sum_{m=0}^{n+1}H_*^m\right)\right)\prod_{m=0,m\neq i}^{n+1}Q_{m,i}^m\\
&=\sum_{i=0}^{n+1}\exp\left(-C_2(k_*+i)^p\left(H-\sum_{m=0}^{n+1}H_*^m\right)\right)\prod_{m=0,m\neq i}^{n+1}Q_{m,i}^m,
\end{aligned}$$

where we have used results of Corollary 6.29. □

We are now able to prove Proposition 6.25.

Proof of Proposition 6.25. Applying Proposition 6.31 yields

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=0}^{n}\sup_{0\le s\le t}|u_{k_*+i}(s)|^2\ge H\right)\\
&\le\exp\left(-C_2k_*^pH\right)\\
&\quad\sum_{i=0}^{n}\left[\exp\left(-C_2[(k_*+i)^p-k_*^p]H+C_2(k_*+i)^p\sum_{m=0}^{n}H_*^m\right)\prod_{m=0,m\neq i}^{n}Q_{m,i}^m\right].
\end{aligned}$$

Note that

$$\prod_{m=0,m\neq i}^{n}Q_{m,i}^m=\prod_{m=0}^{i-1}Q_{m,i}^m\prod_{m=i+1}^{n}Q_{m,i}^m=(-1)^i\prod_{m=0}^{i-1}Q_{i,m}^m\prod_{m=i+1}^{n}Q_{m,i}^m.$$

By monotone convergence we have

$$\mathbb{P}\left(\sum_{k=k_*}^{\infty}\sup_{0\le s\le t}|u_k(s)|^2\ge H\right)\le\exp\left(-C_2k_*^pH\right)\lim_{n\to\infty}\sum_{i=0}^{n}\left[(-1)^is_i^n\right],\tag{6.50}$$

where

$$s_i^n = \exp\left(-C_2[(k_*+i)^p - k_*^p]H + C_2(k_*+i)^p \sum_{m=0}^{n} H_*^m\right) \prod_{m=0}^{i-1} Q_{i,m}^m \prod_{m=i+1}^{n} Q_{m,i}^m$$
$$\geq 0.$$

In what follows, we derive an upper bound for s_i^n being uniform in n

$$\begin{aligned}
\prod_{m=0}^{i-1} Q_{i,m}^m &= \exp\left(-\sum_{m=0}^{i-1} \ln\left(\left(\frac{k_*+i}{k_*+m}\right)^p \left(1-\left(\frac{k_*+m}{k_*+i}\right)^p\right)\right)\right) \\
&\leq \exp\left(\int_0^i \ln\left(\frac{1}{1-\left(\frac{k_*+m}{k_*+i}\right)^p}\right) \,\mathrm{d}m\right) \\
&= \exp\left(-(k_*+i)\int_{k_*/(k_*+i)}^{1} \ln(1-x^p) \,\mathrm{d}x\right) \\
&\leq \exp\left((k_*+i)\int_{k_*/(k_*+i)}^{1} x^p \,\mathrm{d}x\right) \\
&= \exp\left(\frac{k_*+i}{p+1} - \frac{k_*^{p+1}}{(p+1)(k_*+i)^p}\right)
\end{aligned} \tag{6.51}$$

and

$$\begin{aligned}
\prod_{m=i+1}^{n} Q_{m,i}^m &= \prod_{m=i+1}^{n} \left(1+\frac{(k_*+i)^p}{(k_*+m)^p-(k_*+i)^p}\right) \\
&\leq \exp\left(\sum_{m=1}^{n} \ln\left(1+\left(\frac{k_*+i}{m}\right)^p\right)\right) \\
&= \exp\left(\ln\left(1+(k_*+i)^p\right) + \sum_{m=2}^{n} \ln\left(1+\left(\frac{k_*+i}{m}\right)^p\right)\right) \\
&\leq (1+(k_*+i)^p)\exp\left(\int_1^n \ln\left(1+\left(\frac{k_*+i}{m}\right)^p\right) \,\mathrm{d}m\right) \\
&\leq (1+(k_*+i)^p)\exp\left((k_*+i)\int_{1/(k_*+i)}^{n/(k_*+i)} \frac{1}{y^p} \,\mathrm{d}y\right) \\
&= (1+(k_*+i)^p)\exp\left(\frac{1}{1-p}\left(\frac{(k_*+i)^p}{n^{p-1}} - (k_*+i)^p\right)\right) \\
&\leq (1+(k_*+i)^p)\exp\left(\frac{1}{p-1}(k_*+i)^p\right) \qquad \text{for all } n \geq 1.
\end{aligned} \tag{6.52}$$

Now, let $\delta > 0$ such that $\frac{1}{C_2\delta} \geq 1$. Then

$$(1 + (k_* + i)^p) \leq \frac{1}{C_2\delta} \exp\left(C_2\delta(k_* + i)^p\right). \tag{6.53}$$

By inserting the estimates (6.51) and (6.52) into s_i^n and applying (6.53) we obtain uniformly in $n \geq 0$

$$\begin{aligned}
s_i^n \leq\ & \exp\left(-C_2[(k_*+i)^p - k_*^p]H + C_2(k_*+i)^p \sum_{m=0}^{\infty} H_*^m\right) \\
& \exp\left(\frac{k_*+i}{p+1} - \frac{k_*^{p+1}}{(p+1)(k_*+i)^p}\right)(1+(k_*+i)^p)\exp\left((k_*+i)^p\frac{1}{p-1}\right) \\
\leq\ & \exp\left(-C_2[(k_*+i)^p - k_*^p]H + C_2(k_*+i)^p \sum_{m=0}^{\infty} H_*^m\right) \\
& \frac{1}{C_2\delta}\exp\left(C_2\delta(k_*+i)^p\right)\exp\left(\frac{(k_*+i)^p}{p-1} + \frac{k_*+i}{p+1} - \frac{k_*^{p+1}}{(p+1)(k_*+i)^p}\right) \\
\leq\ & \frac{1}{C_2\delta}\exp\left(C_2k_*^pH\right)\exp\left(-\frac{k_*^{p+1}}{(p+1)(k_*+i)^p}\right) \\
& \exp\left(-C_2(k_*+i)^p\left(H - \sum_{m=0}^{\infty} H_*^m - \frac{1}{(p-1)C_2} - \frac{1}{(p+1)C_2} - \delta\right)\right) \\
\leq\ & \frac{1}{C_2\delta}\exp\left(C_2k_*^p\left(\sum_{m=0}^{\infty} H_*^m + \frac{1}{(p-1)C_2} + \frac{1}{(p+1)C_2} + \delta\right)\right) \\
& \exp\left(-C_2pk_*^{p-1}i\left(H - \sum_{m=0}^{\infty} H_*^m - \frac{1}{(p-1)C_2} - \frac{1}{(p+1)C_2} - \delta\right)\right),
\end{aligned}$$

where we have used $(k_*+i)^p \geq k_*^p + pk_*^{p-1}i$ in the last line. Consequently, we get for $H \geq \eta_*$

$$\begin{aligned}
&\lim_{n\to\infty}\sum_{i=0}^{n} s_i^n \\
&\leq \frac{1}{C_2\delta}\exp\left(C_2k_*^p\left(\sum_{m=0}^{\infty} H_*^m + \frac{1}{(p-1)C_2} + \frac{1}{(p+1)C_2} + \delta\right)\right)\sum_{i=0}^{\infty}\exp\left(-C_2pk_*^{p-1}i\delta\right) \\
&= \frac{1}{C_2\delta}\exp\left(C_2k_*^p\left(\sum_{m=0}^{\infty} H_*^m + \frac{2p}{(p^2-1)C_2} + \delta\right)\right)\frac{1}{1-\exp\left(-C_2pk_*^{p-1}\delta\right)} \\
&=: \xi_*\exp\left(C_2k_*^p\eta_*\right),
\end{aligned}$$

where ξ_* and η_* are defined in (6.31) and (6.29). Together with (6.50) this completes the proof. □

6.4.7 Combining finite- and high-frequency estimates

Proof of Theorem 6.21. As outlined before, we split the sum of the components into the finite- and the high-frequency part and obtain

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\leq s\leq t}\|u(s)\|^2 \geq H\right)\\
&= \mathbb{P}\left(\sup_{0\leq s\leq t}\sum_{k=1}^{\infty}|u_k(s)|^2 \geq H\right)\\
&\leq \mathbb{P}\left(\sup_{0\leq s\leq t}\sum_{k=1}^{k_*-1}|u_k(s)|^2 + \sup_{0\leq s\leq t}\sum_{k=k_*}^{\infty}|u_k(s)|^2 \geq H\right)\\
&\leq \mathbb{P}\left(\sup_{0\leq s\leq t}\sum_{k=1}^{k_*-1}|u_k(s)|^2 + \sum_{k=k_*}^{\infty}\sup_{0\leq s\leq t}|u_k(s)|^2 \geq H\right). \qquad (6.54)
\end{aligned}$$

Now using Proposition 6.24 and 6.25 we can once more apply Proposition 6.27 with $n = 0$, $\xi_X(0) = 1$, $\xi_Y = \xi_*$, $\kappa_X(0) = C_1$, $\kappa_Y = C_2 k_*^p$, $\eta_X = \zeta_*$, $\eta_Y = \eta_*$, and we obtain for $H \geq \eta_* + \zeta_*$

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\leq s\leq t}\|u(s)\|^2 \geq H\right)\\
&\leq \exp\left(-C_1\left(H-\eta_*-\zeta_*\right)\right)\\
&\quad - \xi_* C_1 \frac{\exp\left(-C_1(H-\eta_*-\zeta_*)\right) - \exp\left(-C_2 k_*^p(H-\eta_*-\zeta_*)\right)}{C_1 - C_2 k_*^p},
\end{aligned}$$

where the case $C_1 = C_2 k_*^p$ is to be understood in the sense of derivatives. □

6.5 Outlook

In our main result, Theorem 6.21, we have established that it is possible in a simplified setting to extend finite-dimensional fast-slow SODE bounds [BG06] near normally hyperbolic slow manifolds to the infinite-dimensional SPDE setting. In particular we have obtained exponential bounds on the probability to stay near a slow manifold. Our proof has shown that it is possible to naturally extend finite-dimensional results to the SPDE (6.17) using a splitting approach into finitely many ('low') frequency modes as stated in Proposition 6.24 and infinitely many ('high') frequency modes as covered by Proposition 6.25. Furthermore, the key idea is to make use of the growth relation between the eigenvalues coming from the deterministic drift term and the eigenvalues of the covariance operator of the noise. The splitting and the iterative treatment of the high-frequency modes are the key steps in the proof. Those steps could be directly converted into a numerical method. Indeed, just keeping the low-frequency modes corresponds to a Galerkin truncation.

Yet, our approach is only a first step towards providing a detailed theory of multiple time scale SPDEs. There are a few direct possible generalizations. For example, it is evident that the decay in the eigenvalues of the operator A and the spectrum of Q are the key objects, which have to be balanced, to obtain exponential error estimates. Hence, we can allow for more general linear operators A with suitable spectra.

Another next natural step would be to allow linear couplings between the fast and the slow variable, i.e., systems of the form (where we set $B \equiv 0$)

$$\begin{cases} \mathrm{d}u &= \frac{1}{\varepsilon}\left[Au + p_1 u + p_2 v\right]\ \mathrm{d}s + \frac{\sigma}{\sqrt{\varepsilon}}\ \mathrm{d}W, \\ \mathrm{d}v &= \left[p_3 u + p_4 v\right]\ \mathrm{d}s, \end{cases} \tag{6.55}$$

with parameters $p_1, p_2, p_3, p_4 \in \mathbb{R}$. In this case one obtains in the Galerkin approximation 2×2-blocks along the diagonal and the eigenvalues of this block-structured matrix can easily be computed. Under certain assumptions on the eigenvalues and with an iterative scheme similar to the one presented here, we expect to obtain exponential bounds on the sample paths as well. In addition, it is natural to conjecture that suitable regular perturbations of order $\mathcal{O}(\varepsilon)$ of the coefficients are not going to alter the results presented here.

However, there are also several extensions, which are substantially more technical. In particular, dealing with non-linear terms and introducing a general slow SODE including non-linear terms.

One might also ask, whether a more direct approach that avoids the Galerkin approximation technique applied above may lead to fruitful results regarding the concentration of sample paths for fast-slow SPDEs. A first step would be to define the slow manifold of the corresponding deterministic fast-slow PDE system. For that one might be able to use the recent work [HK20], where the authors lift the classical Fenichel theory to the infinite-dimensional setting. Another step would be to derive suitable Bernstein-type estimates in the infinite-dimensional setting.

Notations

Symbol	Description/Definition
$\mathbb{N}$	The natural numbers.
$\mathbb{Z}$	The integers.
$\mathbb{R}$	The real numbers.
$\mathbb{R}^+$	$\mathbb{R}^+ := \{x \in \mathbb{R} : x \geq 0\}$.
$\mathbb{C}$	The complex numbers.
i	The complex number $\sqrt{-1}$.
$\log^{(3)}(t)$	$\log^{(3)}(t) := \log\log\log(t)$.
$\lvert x\rvert$	Euclidean norm of a vector x or graph distance of a vertex x in a graph.
$x^\top$	Transpose of a vector x.
$\otimes$	Tensor product.
$\mathbf{1}_A(x)$, $x \in X$	The indicator function of a subset $A \subset X$.
$\mathbb{1}\{E\}$	Notation for indicator random variable in Chapter 2.
$\delta(x)$	The Dirac delta-function.
$B(x,r)$	In a metric space (M,d) the symbol for a ball of radius $r > 0$ centred at $x \in M$, i.e. $B(x,r) := \{m \in M : d(m,x) \leq r\}$.
$B(r), B_r$	$B_r = B(r) := B(0,r)$.
2^X	Power set of X.
$\mathrm{dist}(A,B)$	Hausdorff semi-distance between two sets A and B.
$\mathcal{B}(X)$	Set of all Borel sets of a topological space X.
Id	Identity.
A^*	Adjoint of an operator A.
$\mathsf{Tr}\, Q$	Trace of an operator Q.
$\mathcal{D}(A)$	Domain of an operator A.
$L(U,H)$	Space of bounded linear operators from U to H.
$L(H)$	$L(H) := L(H,H)$.
$L_2(H,H')$	Space of Hilbert-Schmidt operators from H to H'.
$L_1(H,H')$	Space of trace class operators from H to H'.
$C(D)$	Space of continuous functions on D.
$\lVert\cdot\rVert_\infty$	The supremum norm of a real-valued bounded function f on a domain D given by $\lVert f\rVert_\infty := \sup\{\lvert f(x)\rvert : x \in D\}$.

Symbol	Description/Definition
$C^p(D)$	Space of continuous functions on D that have continuous first p derivatives, $p \in \mathbb{N}$.
$C^\gamma(D)$	For $\gamma \in (0,1]$ space of γ-Hölder continuous functions on D, that is for $f \in C^\gamma(D)$ there exists $C > 0$ such that for all $x, y \in D$: $\lvert f(x) - f(y)\rvert \le C\lvert x-y\rvert^\gamma$.
$C^\infty(D)$	Space of smooth functions on D.
$C_0^\infty(D)$	Space of smooth functions on D that vanish at infinity.
$C_0(\mathbb{R}, U)$	U separable Hilbert space, continuous functions on $\mathbb{R}$ with values in U that vanish at the origin.
$\ell^2(I)$	Space of square summable sequences with index set I, $\ell^2(I) := \big\{(x_i)_{i\in I} \in \mathbb{R}^I : \sum_{i\in I} \lvert x_i\rvert^2 < \infty\big\}$.
$L^p(D)$	The Lebesgue spaces on a domain D with norm $\lVert\cdot\rVert_p$, $p \ge 1$.
$L^p([0,T];H)$	Space of L^p-functions on $[0,T]$ with values in H.
$W^{k,p}(D)$	Sobolev space of order $k \in \mathbb{N}$, defined as $$W^{k,p}(D) := \{u \in L^p(D) : D^\alpha u \in L^p(D)\ \forall \lvert\alpha\rvert \leqslant k\},$$ with multi-index α, where the norm is given by $$\lVert u\rVert_{W^{k,p}(D)} := \begin{cases} \Big(\sum_{\lvert\alpha\rvert\leqslant k} \lVert D^\alpha u\rVert^p_{L^p(D)}\Big)^{\frac{1}{p}} & 1 \leqslant p < \infty; \\ \max_{\lvert\alpha\rvert\leqslant k} \lVert D^\alpha u\rVert_{L^\infty(D)} & p = \infty. \end{cases}$$
$W_0^{k,p}(D)$	Functions in $W^{k,p}(D)$ that vanish at the boundary ∂D in the sense of traces.
$H^k(D)$	$H^k(D) := W^{k,2}(D)$.
$H_0^k(D)$	Functions in $H^k(D)$ that vanish at the boundary ∂D in the sense of traces.
$H^2_{\text{per}}([0,L])$	$H^2_{\text{per}}([0,L]) := \{u \in H^2([0,L]) : u(0) = u(L)\}$
$\mathcal{M}_T^2(K)$	Space of K-valued continuous, square integrable martingales $M(t)$, $t \in [0,T]$.
iid	independent and identically distributed.
a.s.	almost surely.
i.o.	infinitely often.
τ_A	The first exit time of a stochastic process $(X_t)_{t\ge0}$ from a set A, i.e. $\tau_A = \inf\{t \ge 0 : X_t \notin A\}$.
H_v	The first hitting time of a site v by a stochastic process $(X_t)_{t\ge0}$, i.e. $H_v = \inf\{t \ge 0 : X_t = v\}$.
$X \overset{\mathcal{D}}{=} Y$	The random variables X and Y are equal in distribution.
$X_n \overset{\mathcal{D}}{\to} X$	The sequence of random variables $(X_n)_{n\ge1}$ converges in distribution to X for $n \to \infty$.

Symbol	Description/Definition
$X \sim \mathsf{Dist}$	The random variable X is distributed according to Dist.
$\mathcal{L}(X)$	Law of a random variable X.
$\mathcal{N}(\mu, \sigma^2)$	Gaussian law with mean μ and variance σ^2.
$\mathsf{Bernoulli}(p)$	Bernoulli distribution with parameter p.
$\mathsf{Binom}(n, p)$	Binomial distribution with parameters n and p.
$\mathsf{Exp}(\lambda)$	Exponential distribution with parameter λ.
$\mathsf{Poi}(\lambda)$	Poisson distribution with parameter λ.
$\mathsf{Dom}(\mu)$	Domain of attraction of a stable law μ.
$f(x) \asymp g(x)$	There exist constants c_1, c_2 such that $\lvert f(x)\rvert \le c_1 \lvert g(x)\rvert$ and $\lvert g(x)\rvert \le c_2 \lvert f(x)\rvert$ for all x.
$f(x) = O(g(s))$	There exist a constant c such that $\lvert f(x)\rvert \le c\lvert g(x)\rvert$ for all x. Also denoted as $\mathcal{O}$ in Chapter 6.
$f(x) \sim g(x)$	The functions f and g are asymptotically equal as $x \to \infty$, that is $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.
$f(x) = o(g(x))$	$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$.
$\mathbb{T}$	Set of all discrete, plane, rooted trees.
$\mathbb{T}_\infty$	Set of all discrete, plane, rooted trees with a single (one-ended) path to infinity.
$v \sim w$	For vertices v, w in a graph means that v and w are nearest neighbours.
$v \prec w$	Fort vertices v, w is a tree means that v is an ancestor of w.
O	Root of a tree.
$V(\Lambda)$	Volume, i.e. total number of vertices, of $\Lambda \subset G$, for a graph G.
T_u	Subtree $T_u := \{v \in T : u \prec v\}$, where $T \in \mathbb{T}$ and $u \in T$.
$\mathsf{Height}(T)$	$\mathsf{Height}(T) := \sup\{\lvert u\rvert, u \in T\}$ for $T \in \mathbb{T}$, where $\lvert \cdot \rvert$ is the graph distance.
$\mathsf{Diam}(\Lambda)$	$\mathsf{Diam}(\Lambda) := \sup\{\lvert u - v\rvert : u, v \in \Lambda\}$, where $\Lambda \subset G$, for a graph G with graph distance $\lvert \cdot \rvert$.
$\deg(v)$	Degree of a vertex v.
$\mathrm{Off}(v)$	Number of offspring of a vertex v in a tree, also denoted as k_v.
T_∞	Critical GWT conditioned to survive with offspring distribution in the domain of attraction of a β-stable law, see Subsection 1.5.
$s_0, s_1, s_2, \ldots$	Infinite backbone of T_∞.
A_r	Connected component containing O obtained after removing the vertex s_{r+1} from T_∞.
$\tilde{T}_\infty$	$\tilde{T}_\infty := \{v \in T_\infty : \deg v \le 4\}$.
$\tilde{A}_r$	$\tilde{A}_r := \{v \in A_r : \deg v \le 4\}$.
Z_B	$Z_B := \arg\max_{v \in B} \xi(v)$, see Chapter 2.

Symbol	Description/Definition
$\tilde{Z}_B$	$\tilde{Z}_B := \arg\max_{v\in B}[\xi(v) - \deg(v)]$, see Chapter 2.
$\hat{Z}_t = \hat{Z}_t^{(1)}$	$\hat{Z}_t := \arg\max_{v\in T_\infty} \psi_t(v)$, see Chapter 2.
g_Λ	$g_\Lambda := \xi(Z_\Lambda) - \max_{z\in\Lambda, z\neq Z_\Lambda}\{\xi(z)\}$.
$\tilde{g}_\Lambda$	$\tilde{g}_\Lambda := \xi(\tilde{Z}_\Lambda) - \deg(\tilde{Z}_\Lambda) - \max_{z\in\Lambda, z\neq \tilde{Z}_\Lambda}\{\xi(z) - \deg(z)\}$.
$r(t)$	$r(t) := \left(\frac{t}{\log t}\right)^{q+1}$, $q = \frac{d}{\alpha-d}$, $d = \frac{\beta}{\beta-1}$.
$a(t)$	$a(t) := \left(\frac{t}{\log t}\right)^{q+1} = r(t)^{d/\alpha}$, $q = \frac{d}{\alpha-d}$, $d = \frac{\beta}{\beta-1}$.

Appendix A

Probability theory

For completeness we state here a couple of classical results from probability theory, which we apply or refer to at certain points within this thesis.

Lemma A.1 (Borel-Cantelli Lemma I, see [Fel68])**.** *Let $E_1, E_2, ...$ be a sequence of events in some probability space. If*

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty,$$

then

$$\mathbb{P}\left(\limsup_{n\to\infty} E_n\right) = 0.$$

Lemma A.2 (Borel-Cantelli Lemma II, see [Fel68])**.** *Let $E_1, E_2, ...$ be a sequence of independent events in some probability space. If*

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty,$$

then

$$\mathbb{P}\left(\limsup_{n\to\infty} E_n\right) = 1.$$

Lemma A.3 (Chernoff bound for binomial random variable)**.** *Let $X \sim \mathsf{Binom}(n, p)$ and $\theta > 0$, then*

$$\mathbb{P}(X \geq a) \leq \frac{e^{np(e^{\theta}-1)}}{e^{\theta a}}.$$

Proof. Let $X_1, ..., X_n$ be independent random variables with $\mathsf{Bernoulli}(p)$ distribution. Then

$$\mathbb{E}[e^{\theta X_1}] = (1-p)e^0 + pe^{\theta} = 1 + p(e^{\theta} - 1) \leq e^{p(e^{\theta}-1)},$$

hence using Markov's inequality and $X = \sum_{i=1}^n X_i$

$$\mathbb{P}(X \geq a) = \mathbb{P}(e^{\theta X} \geq e^{\theta a}) \leq \frac{\mathbb{E}[e^{\theta X}]}{e^{\theta a}} = \frac{\prod_{i=1}^n \mathbb{E}[e^{\theta X_i}]}{e^{\theta a}} = \frac{e^{np(e^\theta - 1)}}{e^{\theta a}}.$$

□

Theorem A.4 (Doob's martingale inequality, [BG06, Lemma B.1.2]). *Suppose that $(M_t)_{t\geq 0}$ is a positive sub-martingale with continuous paths. Then, for any $L > 0$ and $t > 0$ it holds*

$$\mathbb{P}\left(\sup_{0\leq s\leq t} M_s \geq L\right) \leq \frac{1}{L}\mathbb{E}[M_t].$$

Theorem A.5 (Kolmogorov's lemma, see [LG16, Lemma 2.9]). *Let $X = (X(t))_{t\in I}$ be a stochastic process indexed by a bounded interval $I \subset \mathbb{R}$, and taking values in a complete metric space (E, d). Assume that there exist $q, \varepsilon, C > 0$ such that for every $s, t \in I$*

$$\mathbb{E}\left[d(X(t), X(s))^q\right] \leq C|t - s|^{1+\varepsilon}.$$

Then there exists a modification $\tilde{X}$ of X whose sample paths are Hölder continuous with exponent α for every $\alpha \in (0, \varepsilon/q)$. This means that for every $\omega \in \Omega$ and every $\alpha \in (0, \varepsilon/q)$, there exists a finite constant $C_\alpha(\omega)$ such that for every $s, t \in I$ we have

$$d(\tilde{X}(t,\omega), \tilde{X}(s,\omega)) \leq C_\alpha(\omega)|t - s|^\alpha.$$

In particular, X has a continuous modification.

Appendix B

Analysis

B.1 Useful inequalities

Lemma B.1 (ε-Young inequality)**.** *For $x, y \in \mathbb{R}$, $\varepsilon > 0$, $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, we have*

$$|xy| \leq \varepsilon |x|^p + \frac{(p\varepsilon)^{1-q}}{q} |y|^q.$$

Lemma B.2 (Minkowski's inequality)**.** *Let $p > 1$ and $x, y \in \mathbb{R}$, then*

$$|x + y|^p \leq 2^{p-1}(|x|^p + |y|^p).$$

Lemma B.3 (Gronwall lemma)**.** *Let φ, α and β be real-valued functions on (t_0, ∞). Furthermore, α, β are continuous and φ is differentiable. If*

$$\varphi'(t) \leq \alpha(t) + \beta(t)\varphi(t),$$

then

$$\varphi(t) \leq \varphi(t_0) \exp\left(\int_{t_0}^{t} \beta(\tau) d\tau\right) + \int_{t_0}^{t} \alpha(s) \exp\left(\int_{s}^{t} \beta(\tau) d\tau\right) ds, \quad \text{for all } t \geq t_0.$$

Lemma B.4 (Uniform Gronwall lemma [Tem12, Ch. 3, Lemma 1.1])**.** *Let φ, α, β be positive locally integrable functions on (t_0, ∞) such that φ' is locally integrable on (t_0, ∞) and which satisfy*

$$\varphi'(t) \leq \alpha(t) + \beta(t)\varphi(t), \qquad \text{for } t \geq t_0,$$

$$\int_t^{t+r} \beta(s) \mathrm{d}s \leq a_1, \qquad \int_t^{t+r} \alpha(s) \mathrm{d}s \leq a_2, \qquad \int_t^{t+r} \varphi(s) \mathrm{d}s \leq a_3, \qquad \text{for } t \geq t_0,$$

where r, a_1, a_2, a_3 are positive constants. Then

$$\varphi(t + r) \leq \left(\frac{a_3}{r} + a_2\right) \exp\left(a_1\right), \qquad \text{for all } t \geq t_0.$$

Remark B.5. The uniform Gronwall lemma provides, in contrast to the normal Gronwall lemma, an estimate that is uniform in $t \geq t_0$ and thus guarantees boundedness for $t \to \infty$.

Lemma B.6 (Poincaré's inequality)**.** *Let $1 \leq p < \infty$ and let $D \subset \mathbb{R}^n$ be a bounded open subset. Then there exists a constant $c = c(D, p)$ such that for every function $u \in W_0^{1,p}(D)$*

$$\|u\|_p \leq c\|\nabla u\|_p.$$

B.2 Slowly and regularly varying functions

Definition B.7 (Slowly and regularly varying functions)**.** A measurable function $L : (0, \infty) \to (0, \infty)$ is called *slowly varying* at infinity respectively at zero if for all $a > 0$

$$\lim_{x\to\infty} \frac{L(ax)}{L(x)} = 1 \quad \text{respectively} \quad \lim_{x\to 0^+} \frac{L(ax)}{L(x)} = 1.$$

A measurable function $R : (0, \infty) \to (0, \infty)$ is called *regularly varying* at infinity respectively at zero if for all $a > 0$ the limit

$$\lim_{x\to\infty} \frac{R(ax)}{R(x)} \quad \text{respectively} \quad \lim_{x\to 0^+} \frac{R(ax)}{R(x)}$$

is finite but non-zero.

Remark B.8. Every regularly varying function R is of the form

$$R(x) = x^\beta L(x),$$

where $\beta \in \mathbb{R}$ and L is a slowly varying function.

Appendix C

Functional Analysis

C.1 Embedding theorems

Theorem C.1 (Sobolev embeddings, see [Tri78]). *Let $D \subset \mathbb{R}^n$ be bounded with Lipschitz boundary. Let $1 < p < \infty$ and $0 \leq s < \infty$.*

(i) If $0 \leq s < n/p$, then we have the continuous embedding

$$W^{s,p}(D) \subset L^r(D),$$

where $p \leq r \leq \frac{pn}{n-ps}$.

(ii) If $s = n/p$, then we have the continuous embedding

$$W^{n/p,p}(D) \subset L^r(D),$$

where $p \leq r < \infty$.

(iii) If $s > n/p$, then we have the continuous embedding

$$W^{s,p}(D) \subset C(\overline{D}).$$

Theorem C.2 (Rellich-Kondrachov compactness theorem). *Let $D \subset \mathbb{R}^n$ be bounded with Lipschitz boundary. Then $H^1(D)$ is compactly embedded in $L^2(D)$, i.e.*

$$H^1(D) \subset\subset L^2(D).$$

C.2 Nuclear and Hilbert-Schmidt operators

We restrict ourselves to the Hilbert space setting. Thus, let $(H, \|\cdot\|_H)$ and $(U, \|\cdot\|_U)$ be two separable Hilbert spaces. The following standard definitions and propositions can be found for example in [PR07, Appendix B] or [DPZ92, Appendix C].

Definition C.3. (Nuclear operator) $T \in L(U, H)$ is called *nuclear* or *trace class* if there exists a sequence $(a_j)_{j\in\mathbb{N}}$ in H and a sequence $(b_j)_{j\in\mathbb{N}}$ in U such that T has the representation

$$Tx = \sum_{j=1}^{\infty} a_j \langle b_j, x\rangle_U, \qquad \text{for all } x \in U,$$

and

$$\sum_{j=1}^{\infty} \|a_j\|_H \|b_j\|_U < \infty.$$

Remark C.4. A nuclear operator is a compact operator.

Proposition C.5. *The space of all nuclear operators from U to H is denoted by $L_1(U, H)$. Endowed with the norm*

$$\|T\|_{L_1(U,H)} := \inf \left\{ \sum_{j=1}^{\infty} \|a_j\|_H \|b_j\|_U : Tx = \sum_{j=1}^{\infty} a_j \langle b_j, x\rangle_U, x \in U \right\},$$

this is a Banach space. We also use the abbreviation $L_1(U) := L_1(U, U)$.

Definition C.6. (Trace) Let $T \in L(U)$ and let $(e_k)_{k\in\mathbb{N}}$ be a complete orthonormal system of U. We define

$$\mathsf{Tr}\, T := \sum_{k=1}^{\infty} \langle Te_k, e_k\rangle_U,$$

if the series is convergent. $\mathsf{Tr}\, T$ is called the *trace* of the operator.

Proposition C.7.

(i) If $T \in L_1(U)$ then $\mathsf{Tr}\, T$ is well-defined and independent of the choice of the orthonormal basis $(e_k)_{k\in\mathbb{N}}$.

(ii) We have $|\mathsf{Tr}\, T| \le \|T\|_{L_1(U)}$ for all $T \in L_1(U)$.

(iii) If $T \in L_1(U)$, $S \in L(U)$, then $TS \in L_1(U)$ and

$$\mathsf{Tr}\, TS = \mathsf{Tr}\, ST \le \|T\|_{L_1(U)} \|S\|_{L(U)}.$$

Proposition C.8. *A non-negative operator $T \in L(U)$ is of trace class if and only if for an orthonormal basis $(e_k)_{k\in\mathbb{N}}$ on U it holds*

$$\sum_{k=1}^{\infty} \langle Te_k, e_k\rangle_U < \infty.$$

In this case $\mathsf{Tr}\, T = \|T\|_{L_1(U)}$.

Definition C.9. (Hilbert-Schmidt operator) $T \in L(U,H)$ is called *Hilbert-Schmidt operator* if

$$\sum_{k=1}^{\infty} \|Te_k\|_H^2 < \infty,$$

where $(e_k)_{k\in\mathbb{N}}$ is an orthonormal basis of U.

Proposition C.10. *We denote the set of all Hilbert-Schmidt operators from U to H as $L_2(U,H)$. Equipped with the norm*

$$\|T\|_{L_2(U,H)} := \sum_{k=1}^{\infty} \|Te_k\|_H^2,$$

this is a Hilbert space.

Proposition C.11. *Let G be a separable Hilbert space. Let $S \in L(H,G)$ and $T \in L_2(U,H)$, then $ST \in L_2(U,G)$.*

Proposition C.12. *Let G be a separable Hilbert space. Let $S \in L_2(H,G)$ and $T \in L_2(U,H)$, then $ST \in L_1(U,G)$ and*

$$\|ST\|_{L_1(U,G)} \leq \|S\|_{L_2(H,G)}\|T\|_{L_2(U,H)}.$$

Remark C.13. Clearly, we have the inclusion $L_1(U,H) \subset L_2(U,H) \subset L(U,H)$.

C.3 Semigroups of operators

We refer to the book [Paz12] for a comprehensive presentation of the theory of semigroups of linear operators. All definitions and propositions listed below can be found therein or in [RR06, Chapter 12]. In the following, we always assume that X is a Banach space with norm $\|\cdot\|_X$ and $\|\cdot\|$ denotes the operator norm on $L(X)$.

Definition C.14 (Semigroup)**.** Let X be a Banach space. A family of bounded linear operators $\{T(t)\}_{t\geq 0}$ in X is called a *semigroup* if the following two properties hold

(i) $T(0) = \mathsf{Id}$,

(ii) $T(t+s) = T(t)T(s)$, for all $t, s \geq 0$.

Definition C.15 (Infinitesimal generator)**.** Let X be a Banach space. Let $\{T(t)\}_{t\geq 0}$ be a semigroup of bounded linear operators on X. The linear operator A defined by

$$Au = \lim_{t\downarrow 0} \frac{T(t)u - u}{t},$$

on the domain $D(A) = \left\{u \in X : \lim_{t\downarrow 0} \frac{T(t)u-u}{t} \text{ exists}\right\}$, is called the *infinitesimal generator* of the semigroup.

C.3.1 C_0-semigroups

Definition C.16 (C_0-semigroup)**.** Let X be a Banach space. A family of bounded linear operators $\{T(t)\}_{t\geq 0}$ in X is called a *strongly continuous semigroup* or C_0-semigroup, if

(i) $\{T(t)\}_{t\geq 0}$ is a semigroup,

(ii) $\lim_{t\downarrow 0} T(t)u = u$ for every $u \in X$, i.e. $t \mapsto T(t)u$ is continuous at $t = 0$.

Theorem C.17. *For each C_0-semigroup $\{T(t)\}_{t\geq 0}$, there are constants $M \geq 1$ and $\omega \geq 0$ such that*

$$\|T(t)\| \leq M \exp(\omega t), \qquad \text{for all } t \geq 0, \tag{C.1}$$

where $\|\cdot\|$ denotes the operator norm.

Definition C.18. ((Quasi-)contraction semigroup)
A C_0-semigroup $\{T(t)\}_{t\geq 0}$ is called

1. a quasicontraction semigroup if $\|T(t)\| \leq \exp(\omega t)$ for some ω, i.e. the growth estimate (C.1) is satisfied with $M = 1$.
2. a contraction semigroup if $\|T(t)\| \leq 1$, i.e. the growth estimate (C.1) is satisfied with $M = 1$ and $\omega = 0$.

Remark C.19. If $\{T(t)\}_{t\geq 0}$ is a quasicontraction semigroup with growth estimate $\|T(t)\| \leq \exp(\omega t)$, then $\{S(t)\}_{t\geq 0}$ given by $S(t) := \exp(-\omega t)\, T(t)$ for all $t \geq 0$ is a contraction semigroup. Hence, every quasicontraction semigroup can be transformed into a contraction semigroup.

C.3.2 Analytic semigroups

Definition C.20 (Analytic semigroup)**.** Let X be a Banach space. A C_0-semigroup $\{T(t)\}_{t\geq 0}$ in X is called an *analytic semigroup* if the following holds:

(i) For some $\delta \in (0, \pi/2)$, $T(t) \in L(X)$ can be extended to all $t \in \Delta_\delta := \{t \in \mathbb{C} : |\arg t| < \delta\} \cup \{0\}$ and for $t \in \Delta_\delta$ the conditions (i)-(ii) of Definition C.16 hold.

(ii) For $t \in \Delta_\delta \backslash \{0\}$, the mapping $t \mapsto T(t)x$ is analytic for every $x \in X$.

Some of the key properties of analytic semigroups are summarized in the following lemmas.

Lemma C.21. *(cf. [RR06, Lemma 12.36]) Let A be the generator of an analytic semigroup $\{\exp(At)\}_{t\geq 0}$ on a Banach space X and assume that the spectrum of A lies entirely in the open left half-plane. Then there exists $\delta > 0$ and constants M, M_1, M_n such that*

$$\|\exp(At)\| \leq Me^{-\delta t},$$

$$\|A\exp(At)\| \le M_1 \exp(-\delta t)/t,$$

$$\|A^n \exp(At)\| \le M_n \exp(-\delta t)/t^n, \text{ for } n \in \mathbb{N}.$$

Furthermore, let $\alpha > 0$, then for every $x \in \mathcal{D}((-A)^\alpha)$, we have

$$\exp(At)(-A)^\alpha x = (-A)^\alpha \exp(At)x,$$

and

$$\|(-A)^\alpha \exp(At)\| \le M_\alpha t^{-\alpha} e^{-\delta t}.$$

Lemma C.22 (Perturbations of analytic semigroups, see [RR06, Thm 12.37])**.** *Let A be the generator of an analytic semigroup on X. Then there exists $\delta > 0$ such that, if B is any operator such that*

(i) B is closed and $\mathcal{D}(B) \supset \mathcal{D}(A)$,

(ii) $\|Bu\|_X \le a\|Au\|_X + b\|u\|_X$ for $u \in \mathcal{D}(A)$ and $a \le \delta$,

then $A + B$ also generates an analytic semigroup.

Bibliography

[AD14a] R. Abraham and J.-F. Delmas. Local limits of conditioned Galton-Watson trees: the condensation case. *Electron. J. Probab.*, 19, 2014.

[AD14b] R. Abraham and J.-F. Delmas. Local limits of conditioned Galton-Watson trees: the infinite spine case. *Electron. J. Probab.*, 19, 2014.

[AGH20] L. Avena, O. Gün, and M. Hesse. The parabolic Anderson model on the hypercube. *Stoch. Process. Their Appl.*, 130(6):3369–3393, 2020.

[AN12] K. B. Athreya and P. E. Ney. *Branching Processes*. Grundlehren der mathematischen Wissenschaften. Springer, 2012.

[And58] P. W. Anderson. Absence of diffusion in certain random lattices. *Phys. Rev.*, 109(5):1492–1505, 1958.

[Arc20] E. Archer. Random walks on decorated Galton-Watson trees. *arXiv preprint arXiv:2011.07266*, 2020.

[Arn13] L. Arnold. *Random dynamical systems*. Springer, 2013.

[AS96] L. Arnold and B. Schmalfuss. Fixed points and attractors for random dynamical systems. In *IUTAM Symposium on Advances in Nonlinear Stochastic Mechanics*, pages 19–28. Springer, 1996.

[AW13a] A. Adili and B. Wang. Random attractors for non-autonomous stochastic fitzhugh-nagumo systems with multiplicative noise. In *Conference Publications*, pages 1–10. AIMS, 2013.

[AW13b] A. Adili and B. Wang. Random attractors for stochastic FitzHugh-Nagumo systems driven by determinisits non-autonomous forcing. *Discrete Contin. Dyn. Syst. Ser. B*, 18(3):643–666, 2013.

[AW15] M. Aizenman and S. Warzel. *Random operators*. AMS, 2015.

[Bal77] J. M. Ball. Strongly continuous semigroups, weak solutions, and the variation of constants formula. *Proceedings of the American Mathematical Society*, 63(2):370–373, 1977.

[Bar91] D. Barkley. A model for fast computer simulation of waves in excitable media. *Physica D*, 49(1-2):61–70, 1991.

[BC95] L. Bertini and N. Cancrini. The stochastic heat equation: Feynman-Kac formula and intermittence. *J. Stat. Phys.*, 78(5-6):1377–1401, 1995.

[BG02] N. Berglund and B. Gentz. Pathwise description of dynamic pitchfork bifurcations with additive noise. *Probab. Theory Relat. Fields*, 122(3):341–388, 2002.

[BG03] N. Berglund and B. Gentz. Geometric singular perturbation theory for stochastic differential equations. *J. Differ. Equat.*, 191(1):1–54, 2003.

[BG06] N. Berglund and B. Gentz. *Noise-Induced Phenomena in Slow-Fast Dynamical Systems*. Springer, 2006.

[BH10] D. Blömker and Y. Han. Asymptotic compactness of stochastic complex Ginzburg–Landau equation on an unbounded domain. *Stoch. Dyn.*, 10(04):613–636, 2010.

[BJ13] D. Blömker and A. Jentzen. Galerkin approximations for the stochastic Burgers equation. *SIAM J. Numer. Anal.*, 51(1):694–715, 2013.

[BK01] M. Biskup and W. König. Long-time tails in the parabolic Anderson model with bounded potential. *Ann. Probab.*, 29(2):636–682, 2001.

[BK16] N. Berglund and C. Kuehn. Regularity structures and renormalisation of FitzHugh-Nagumo SPDEs in three space dimensions. *Electron. J. Probab.*, 21(18):1–48, 2016.

[Blö07] D. Blömker. *Amplitude equations for stochastic partial differential equations*. World Scientific, 2007.

[BLW09] P. W. Bates, K. Lu, and B. Wang. Random attractors for stochastic reaction–diffusion equations on unbounded domains. *J. Differ. Equat.*, 246(2):845–869, 2009.

[BM08] S. Bonaccorsi and E. Mastrogiacomo. Analysis of the stochastic FitzHugh-Nagumo system. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 11(03):427–446, 2008.

[BPSV82] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. Stochastic resonance in climatic change. *Tellus*, 34(11):10–16, 1982.

[BT72] A. Bensoussan and R. Temam. Équations aux dérivées partielles stochastiques non linéaires. *Israel J. Math.*, 11(1):95–129, 1972.

[CD17] R. Cherniha and V. Davydovych. *Nonlinear reaction-diffusion systems.* Springer, 2017.

[CDF97] H. Crauel, A. Debussche, and F. Flandoli. Random attractors. *J. Dyn. Differ. Equ.*, 9(2):307–341, 1997.

[CF94] H. Crauel and F. Flandoli. Attractors for random dynamical systems. *Probab. Theory Relat. Fields*, 100(3):365–393, 1994.

[CGALdlC17] T. Caraballo, M. J. Garrido-Atienza, and J. López-de-la Cruz. Dynamics of some stochastic chemostat models with multiplicative noise. *Commun. Pure Appl. Math.*, 16(5):1893–1914, 2017.

[CH17] T. Caraballo and X. Han. *Applied Nonautonomous and Random Dynamical Systems: Applied Dynamical Systems.* Springer, 2017.

[Cho14] P. L. Chow. *Stochastic Partial Differential Equations, Second Edition.* Advances in Applied Mathematics. Taylor & Francis, 2014.

[CK08] D. Croydon and T. Kumagai. Random walks on Galton-Watson trees with infinite variance offspring distribution conditioned to survive. *Electron. J. Probab.*, 13, 2008.

[CK19] C. Chong and P. Kevei. Intermittency for the stochastic heat equation with Lévy noise. *The Annals of Probability*, 47(4):1911–1948, 2019.

[CKNP20] L. Chen, D. Khoshnevisan, D. Nualart, and F. Pu. Spatial ergodicity and central limit theorems for parabolic Anderson model with delta initial condition. *arXiv preprint arXiv:2005.10417*, 2020.

[CKS04] T. Caraballo, P. E. Kloeden, and B. Schmalfuß. Exponentially stable stationary solutions for stochastic evolution equations and their perturbation. *Appl. Math. Optim.*, 50(3):183–207, 2004.

[CL08] T. Caraballo and K. Lu. Attractors for stochastic lattice dynamical systems with a multiplicative noise. *Front. Math. China*, 3(3):317–335, 2008.

[CLR00] T. Caraballo, J. A. Langa, and J. C. Robinson. Stability and random attractors for a reaction-diffusion equation with multiplicative noise. *Discrete Contin. Dyn. Syst.*, 6(4):875–892, 2000.

[CLR01] T. Caraballo, J. A. Langa, and J. C. Robinson. A stochastic pitchfork bifurcation in a reaction-diffusion equation. *Proc. R. Soc. Lond. Ser. A*, 457(2013):2041–2061, 2001.

[CM94] R. A. Carmona and S. A. Molchanov. *Parabolic Anderson problem and intermittency*, volume 518. AMS, 1994.

[CM95] R. A. Carmona and S. A. Molchanov. Stationary parabolic Anderson model and intermittency. *Probab. Theory Relat. Fields*, 102(4):433–453, 1995.

[Cn70] E. M. Cabaña. The vibrating string forced by white noise. *Probab. Theory Relat. Fields*, 15(2):111–130, 1970.

[CS10] I. Chueshov and B. Schmalfuss. Master-slave synchronization and invariant manifolds for coupled stochastic systems. *J. Math. Phys.*, 51(10):102702, 2010.

[CS17] T. Caraballo and S. Sonner. Random pullback exponential attractors: General existence results for random dynamical systems in Banach spaces. *Discrete Contin. Dyn. Syst. Ser. A*, 37(12):6383–6403., 2017.

[CV96] V. V. Chepyzhov and M. I. Vishik. Trajectory attractors for reaction-diffusion systems. *Topol. Methods Nonlinear Anal.*, 7(1):49–76, 1996.

[CV04] I. Chueshov and P.-A. Vuillermot. Non-random invariant sets for some systems of parabolic stochastic partial differential equations. *Stoch. Anal. Appl.*, 22(6):1421–1486, 2004.

[DB94] R. J. Deissler and H. R. Brand. Periodic, quasiperiodic, and chaotic localized solutions of the quintic complex Ginzburg-Landau equation. *Phys. Rev. Lett.*, 72(4):478–481, 1994.

[Deb97] A. Debussche. On the finite dimensionality of random attractors. *Stoch. Anal. Appl.*, 15(4):473–491, 1997.

[Deb98] A. Debussche. Hausdorff dimension of a random invariant set. *J. Math. Pures Appl.*, 77(10):967–988, 1998.

[dKd20] F. den Hollander, W. König, and R. dos Santos. The parabolic Anderson model on a Galton-Watson tree. *arXiv preprint arXiv:2001.05106*, 2020.

[DLS03] J. Duan, K. Lu, and B. Schmalfuss. Invariant manifolds for stochastic partial differential equations. *Ann. Probab.*, 31(4):2109–2135, 2003.

[Dos77] H. Doss. Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. Henri Poincare Probab. Stat.*, 13(2):99–125, 1977.

[DP12] G. Da Prato. *Kolmogorov equations for stochastic PDEs*. Birkhäuser, 2012.

[DPZ92] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge University Press, 1992.

[DQS11] R. C. Dalang and L. Quer-Sardanyons. Stochastic integrals for spde's: a comparison. *Expo. Math.*, 29(1):67–109, 2011.

[DW14] J. Duan and W. Wang. *Effective Dynamics of Stochastic Partial Differential Equations.* Elsevier, 2014.

[Eck11] W. Eckhaus. *Matched asymptotic expansions and singular perturbations.* Elsevier, 2011.

[EK20] M. Engel and C. Kuehn. Blow-up analysis of fast-slow PDEs with loss of hyperbolicity. *arXiv preprint arXiv:2007.09973*, 2020.

[EKN20] K. Eichinger, C. Kuehn, and A. Neamţu. Sample paths estimates for stochastic fast-slow systems driven by fractional Brownian motion. *J. Stat. Phys.*, 179:1222–1266, 2020.

[ET10] G. B. Ermentrout and D. H. Terman. *Mathematical foundations of neuroscience.* Springer, 2010.

[Fel68] W. Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[Fel71] W. Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

[Fen79] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *J. Differ. Equat.*, 31(1):53–98, 1979.

[Fit61] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.*, 1(6):445–466, 1961.

[FKN72] R. J. Field, E. Koros, and R. M. Noyes. Oscillations in chemical systems. II. Thorough analysis of temporal oscillation in the bromate-cerium-malonic acid system. *J. Am. Chem. Soc.*, 94(25):8649–8664, 1972.

[FM90] K. Fleischmann and S. A. Molchanov. Exact asymptotics in a mean field model with random potential. *Probab. Theory Relat. Fields*, 86(2):239–251, 1990.

[FM14] A. Fiodorov and S. Muirhead. Complete localisation and exponential shape of the parabolic Anderson model with Weibull potential field. *Electron. J. Probab.*, 19, 2014.

[FN74] R. J. Field and R. M. Noyes. Oscillations in chemical systems. IV. Limit cycle behavior in a model of a real chemical reaction. *J. Chem. Phys.*, 60(5):1877–1884, 1974.

[Fou22] J. B. J. Fourier. *Théorie analytique de la chaleur.* F. Didot, 1822.

[FS96] F. Flandoli and B. Schmalfuss. Random attractors for the 3D stochastic Navier-Stokes equation with multiplicative white noise. *Stochastics*, 59(1-2):21–45, 1996.

[GIP15] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics, Pi*, 3, 2015.

[GK98] J. Geiger and G. Kersting. The Galton-Watson tree conditioned on its height. *Probab. Theory and Math. Stat.*, pages 277–286, 1998.

[GK00] J. Gärtner and W. König. Moment asymptotics for the continuous parabolic Anderson model. *Ann. Appl. Probab.*, 10(1):192–217, 2000.

[GK05] J. Gärtner and W. König. The parabolic Anderson model. In *Interacting stochastic systems*, pages 153–179. Springer, 2005.

[GK15] K. Gowda and C. Kuehn. Warning signs for pattern-formation in SPDEs. *Comm. Nonl. Sci. & Numer. Simul.*, 22(1):55–69, 2015.

[GKM00] J. Gärtner, W. König, and S. A. Molchanov. Almost sure asymptotics for the continuous parabolic Anderson model. *Probab. Theory Relat. Fields*, 118(4):547–573, 2000.

[GKM07] J. Gärtner, W. König, and S. A. Molchanov. Geometric characterization of intermittency in the parabolic Anderson model. *Ann. Probab.*, 35(2):439–499, 2007.

[GKP19] M. V. Gnann, C. Kuehn, and A. Pein. Towards sample path estimates for fast–slow stochastic partial differential equations. *Eur. J. Appl. Math.*, 30(5):1004–1024, 2019.

[GLR11] B. Gess, W. Liu, and M. Röckner. Random attractors for a class of stochastic partial differential equations driven by general additive noise. *J. Differ. Equat.*, 251(4-5):1225–1253, 2011.

[GM90] J. Gärtner and S. A. Molchanov. Parabolic problems for the Anderson model. *Commun. Math. Phys.*, 132(3):613–655, 1990.

[GM98] J. Gärtner and S. A. Molchanov. Parabolic problems for the Anderson model II. Second-order asymptotics and structure of high peaks. *Probab. Theory Relat. Fields*, 111(1):17–55, 1998.

[GM10] L. Gawarecki and V. Mandrekar. *Stochastic differential equations in infinite dimensions: with applications to stochastic partial differential equations.* Springer, 2010.

[GOS12] J. García-Ojalvo and J. Sancho. *Noise in spatially extended systems.* Springer, 2012.

[GR11] M. Ghergu and V. Radulescu. *Nonlinear PDEs: Mathematical models in biology, chemistry and population genetics.* Springer, 2011.

[Gra53] I. S. Gradshtein. Application of AM Lyapunov's theory of stability to the theory of differential equations with small coefficients in the derivatives. *Matematicheskii Sbornik*, 74(2):263–286, 1953.

[GSB11] J. H. Goldwyn and E. Shea-Brown. The what and where of adding channel noise to the Hodgkin-Huxley equations. *PLoS Comput. Biol*, 7(11):e1002247, 2011.

[GW18] A. Gu and B. Wang. Asymptotic behavior of random FitzHugh-Nagumo systems driven by colored noise. *Discrete Contin. Dyn. Syst. Ser. B*, 23(4):1689–1720, 2018.

[Hai13] M. Hairer. Solving the KPZ equation. *Ann. Math.*, 178(2):559–664, 2013.

[Hai14] M. Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[Har63] T. E. Harris. *The theory of branching process.* Springer, 1963.

[Has76] K. Hasselmann. Stochastic climate models part I. Theory. *Tellus*, 28(6):473–485, 1976.

[HH52] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117(4):500–544, 1952.

[HHNT15] Y. Hu, J. Huang, D. Nualart, and S. Tindel. Stochastic heat equations with general multiplicative gaussian noises: Hölder continuity and intermittency. *Electron. J. Probab.*, 20, 2015.

[HK20] F. Hummel and C. Kuehn. Slow Manifolds for Infinite-Dimensional Evolution Equations. *arXiv preprint arXiv:2008.10700*, 2020.

[HSZS18] S. A. Hiremath, C. Surulescu, A. Zhigun, and S. Sonner. On a coupled SDE-PDE system modeling acid-mediated tumor invasion. *Discrete Contin. Dyn. Syst. Ser. B*, 23(6):2339–2369, 2018.

[IS01] P. Imkeller and B. Schmalfuss. The conjugacy of stochastic and random differential equations and the existence of global attractors. *J. Dyn. Differ. Equ.*, 13(2):215–249, 2001.

[Jan12] S. Janson. Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probab. Surv.*, 9:103–252, 2012.

[Jon99] B. J. T. Jones. The origin of scaling in the galaxy distribution. *Monthly Notices of the Royal Astronomical Society*, 307(2):376–386, 1999.

[Już14] M. Jużyniec. The existence of a weak solution of the semilinear first-order differential equation in a Banach space. *Czasopismo Techniczne*, 2014.

[KC12] J. K. Kevorkian and J. D. Cole. *Multiple scale and singular perturbation methods*. Springer, 2012.

[Kes86] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. Henri Poincare Probab. Stat.*, 22(4):425–487, 1986.

[Kho14] D. Khoshnevisan. *Analysis of stochastic partial differential equations*. AMS, 2014.

[Kin13] S. Kinoshita. *Pattern Formations and Oscillatory Phenomena*. Elsevier, 2013.

[KLMS09] W. König, H. Lacoin, P. Mörters, and N. Sidorova. A two cities theorem for the parabolic Anderson model. *Ann. Probab.*, 37(1):347–392, 2009.

[KMS06] W. König, P. Mörters, and N. Sidorova. Complete localisation in the parabolic Anderson model with Pareto-distributed potential. *arXiv preprint math/0608544*, 2006.

[KNP20] C. Kuehn, A. Neamţu, and A. Pein. Random attractors for stochastic partly dissipative systems. *Nonlinear Differ. Equ. Appl.*, 27(35), 2020.

[KNS20] C. Kuehn, A. Neamţu, and S. Sonner. Random attractors via pathwise mild solutions for stochastic parabolic evolution equations. *arXiv preprint arXiv:2006.02787*, 2020.

[Kön16] W. König. *The Parabolic Anderson Model: Random Walk in Random Potential*. Birkhäuser, 2016.

[Kön20] W. König. Branching random walks in random environment: a survey. *arXiv preprint arXiv:2010.06942*, 2020.

[Kor04] J. Korevaar. *Tauberian theory*. Springer, 2004.

[Kor12] I. Kortchemski. Invariance principles for Galton–Watson trees conditioned on the number of leaves. *Stoch. Process. Their Appl.*, 122(9):3126–3172, 2012.

[Kot92] P. Kotelenez. Comparison methods for a class of function valued stochastic partial differential equations. *Probab. Theory Relat. Fields*, 93(1):1–19, 1992.

[KPvZ20] W. König, N. Perkowski, and W. van Zuijlen. Longtime asymptotics of the two-dimensional parabolic Anderson model with white-noise potential. *arXiv preprint arXiv:2009.11611*, 2020.

[KS98] T. Kapitula and B. Sandstede. Instability mechanism for bright solitary-wave solutions to the cubic-quintic Ginzburg-Landau equation. *J. Opt. Soc. Am. B*, 15(11):2757–2762, 1998.

[KS99] H. Keller and B. Schmalfuß. Attractors for stochastic Sine Gordon equations via transformation into random equations. *Report 448, Universität Bremen*, 1999.

[Kue15a] C. Kuehn. Efficient gluing of numerical continuation and a multiple solution method for elliptic PDEs. *Appl. Math. Comput.*, 266:656–674, 2015.

[Kue15b] C. Kuehn. *Multiple time scale dynamics.* Springer, 2015.

[Kue15c] C. Kuehn. Numerical Continuation and SPDE Stability for the 2D Cubic-Quintic Allen–Cahn Equation. *SIAM/ASA J. Uncertain. Quantif.*, 3(1):762–789, 2015.

[Kun97] H. Kunita. *Stochastic flows and stochastic differential equations.* Cambridge University Press, 1997.

[LG16] J.-F. Le Gall. *Brownian motion, martingales, and stochastic calculus.* Springer, 2016.

[LPS14] G. J. Lord, C. E. Powell, and T. Shardlow. *An introduction to computational stochastic PDEs.* Cambridge University Press, 2014.

[LPS15] K. B. Liaskos, A. A. Pantelous, and I. G. Stratis. Linear stochastic degenerate Sobolev equations and applications. *Int. J. Control*, 88(12):2538–2553, 2015.

[LR15] W. Liu and M. Röckner. *Stochastic partial differential equations: an introduction.* Springer, 2015.

[LSG99] B. Lindner and L. Schimansky-Geier. Analytical approach to the stochastic FitzHugh-Nagumo system and coherence resonance. *Phys. Rev. E*, 60(6):7270–7276, 1999.

[LSP18] K. B. Liaskos, I. G. Stratis, and A. A. Pantelous. Stochastic degenerate Sobolev equations: well posedness and exact controllability. *Math. Methods Appl. Sci.*, 41(3):1025–1032, 2018.

[LY16] Y. Li and J. Yin. A modified proof of pullback attractors in a Sobolev space for stochastic FitzHugh-Nagumo equations. *Discrete Contin. Dyn. Syst. Ser. B*, 21(4):1203–1223, 2016.

[Mar87] M. Marion. Attractors for reaction-diffusion equations: existence and estimate of their dimension. *Applicable Analysis*, 25(1-2):101–147, 1987.

[Mar89] M. Marion. Finite-dimensional attractors associated with partly dissipative reaction-diffusion systems. *SIAM J. Math. Anal.*, 20(4):816–844, 1989.

[MCK06] A. Marciniak-Czochra and M. Kimmel. Dynamics of growth and signaling along linear and surface structures in very early tumors. *Comput. Math. Methods Med.*, 7(2-3):189–213, 2006.

[MCK08] A. Marciniak-Czochra and M. Kimmel. Reaction-difusion model of early carcinogenesis: The effects of influx of mutated cells. *Math. Model. Nat. Pheno.*, 3(7):90–114, 2008.

[MD14] D. Morgan and J.H.P. Dawes. The Swift-Hohenberg equation with a nonlocal nonlinearity. *Physica D*, 270:60–80, 2014.

[ML81] C. Morris and H. Lecar. Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.*, 35(1):193–213, 1981.

[Mol91] S. A. Molchanov. Ideas in the theory of random media. *Acta Applicandae Mathematica*, 22(2-3):139–282, 1991.

[Mör09] P. Mörters. The parabolic Anderson model with heavy-tailed potential. In *Surveys in Stochastic Processes, Proceedings of the 33rd SPA Conference in Berlin*, 2009.

[MOS11] P. Mörters, M. Ortgiese, and N. Sidorova. Ageing in the parabolic Anderson model. *Ann. Inst. Henri Poincare Probab. Stat.*, 47(4):969–1000, 2011.

[MP14] M. Métivier and J. Pellaumail. *Stochastic integration.* Academic Press, 2014.

[MP16] S. Muirhead and R. Pymar. Localisation in the Bouchaud–Anderson model. *Stoch. Process. Their Appl.*, 126(11):3402–3462, 2016.

[MTvdE01] A. J. Majda, I. Timofeyev, and E. van den Eijnden. A mathematical framework for stochastic climate models. *Commun. Pure Appl. Math.*, 54(8):891–974, 2001.

[Mue15] C. Mueller. Stochastic PDE from the point of view of particle systems and duality. In *Stochastic Analysis: A Series of Lectures*, pages 271–295. Springer, 2015.

[Mur07] J. D. Murray. *Mathematical biology: I. An introduction.* Springer, 2007.

[MZZ06] S.-E. A. Mohammed, T. Zhang, and H. Zhao. The stable manifold theorem for semilinear stochastic evolution equations and stochastic partial differential equations. *Memoirs of the AMS*, 98, 2006.

[Nag89] R. Nagel. Towards a "matrix theory" for unbounded operator matrices. *Math. Z.*, 201(1):57–68, 1989.

[Nar99] T. N. Narasimhan. Fourier's heat conduction equation: History, influence, and connections. *Rev. Geophys.*, 37(1):151–172, 1999.

[NAY62] J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, 1962.

[Nev86] J. Neveu. Arbres et processus de Galton-Watson. *Ann. Inst. Henri Poincare Probab. Stat.*, 22(2):199–207, 1986.

[NFK72] R. M. Noyes, R. J. Field, and E. Koros. Oscillations in chemical systems. I. Detailed mechanism in a system showing temporal oscillations. *J. Am. Chem. Soc.*, 94(4):1394–1395, 1972.

[Oks03] B. Oksendal. *Stochastic differential equations.* Springer, 2003.

[O'm91] R. E. O'malley. *Singular perturbation methods for ordinary differential equations.* Springer, 1991.

[OR16] M. Ortgiese and M. I. Roberts. Intermittency for branching random walk in Pareto environment. *Ann. Probab.*, 44(3):2198–2263, 2016.

[OR17] M. Ortgiese and M. I. Roberts. One-point localization for branching random walk in Pareto environment. *Electron. J. Probab.*, 22, 2017.

[OR18] M. Ortgiese and M. I. Roberts. Scaling limit and ageing for branching random walk in Pareto environment. *Ann. Inst. Henri Poincare Probab. Stat.*, 54(3):1291–1313, 2018.

[Paz12] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*. Springer, 2012.

[Per15] B. Perthame. *Parabolic Equations in Biology*. Springer, 2015.

[Pha20] C. Phan. Random attractor for stochastic Hindmarsh-Rose equations with multiplicative noise. *Discrete Contin. Dyn. Syst. Ser. B*, 25(8):3233–3256, 2020.

[PR07] C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*. Springer, 2007.

[PY19] C. Phan and Y. You. Random Attractor for Stochastic Hindmarsh–Rose Equations with Additive Noise. *J. Dyn. Differ. Equ.*, pages 1–22, 2019.

[PZ07] S. Peszat and J. Zabczyk. *Stochastic partial differential equations with Lévy noise: An evolution equation approach*. Cambridge University Press, 2007.

[RBW00] A. Rodriguez-Bernal and B. Wang. Attractors for partly dissipative reaction diffusion systems in $\mathbb{R}^n$. *J. Math. Anal. Appl*, 252(2):790–803, 2000.

[RR06] M. Renardy and R. C. Rogers. *An introduction to partial differential equations*. Springer, 2006.

[Sch92] B. Schmalfuss. Backward cocycles and attractors of stochastic differential equations. In *International Seminar on Applied Mathematics-Nonlinear Dynamics: Attractor Approximation and Global Behaviour*, pages 185–192, 1992.

[SK18] S. Sadhu and C. Kuehn. Stochastic mixed-mode oscillations in a three-species predator-prey model. *Chaos*, 28(3):033606, 2018.

[Sla68] R. S. Slack. A branching process with mean one and possibly infinite variance. *Probab. Theory Relat. Fields*, 9(2):139–145, 1968.

[Smo94] J. Smoller. *Shock waves and reaction-diffusion equations*. Springer, 1994.

[SMSG07] M. Sieber, H. Malchow, and L. Schimansky-Geier. Constructive effects of environmental noise in an excitable prey-predator plankton system with infected prey. *Ecol. Complex.*, 4(4):223–233, 2007.

[SRT04] J. Su, J. Rubin, and D. Terman. Effects of noise on elliptic bursters. *Nonlinearity*, 17(1):133–157, 2004.

[SS16] M. Sauer and W. Stannat. Analysis and approximation of stochastic nerve axon equations. *Math. Comput.*, 85(301):2457–2481, 2016.

[ST14] N. Sidorova and A. Twarowski. Localisation and ageing in the parabolic Anderson model with Weibull potential. *Ann. Probab.*, 42(4):1666–1698, 2014.

[Sus78] H. J. Sussmann. On the gap between deterministic and stochastic ordinary differential equations. *Ann. Probab.*, 6(1):19–41, 1978.

[SWLW19] L. Shi, R. Wang, K. Lu, and B. Wang. Asymptotic behavior of stochastic FitzHugh-Nagumo systems on unbounded thin domains. *J. Differ. Equat.*, 267(7):4373–4409, 2019.

[SY02] G. R. Sell and Y. You. *Dynamics of Evolutionary Equations*. Applied Mathematical Sciences. Springer, 2002.

[Szn93] A.-S. Sznitman. Brownian asymptotics in a Poissonian environment. *Probab. Theory Relat. Fields*, 95(2):155–174, 1993.

[Szn98] A.-S. Sznitman. *Brownian motion, obstacles and random media*. Springer, 1998.

[Tap13] S. Tappe. Foundations of the theory of semilinear stochastic partial differential equations. *Int. J. Stoch. Anal.*, 2013.

[Tem12] R. Temam. *Infinite-dimensional dynamical systems in mechanics and physics*. Springer, 2012.

[Tik52] A. N. Tikhonov. Systems of differential equations containing small parameters in the derivatives. *Matematicheskii sbornik*, 73(3):575–586, 1952.

[Tri78] H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. North-Holland Publishing Company, 1978.

[Tur90] A. M. Turing. The chemical basis of morphogenesis. *Bull. Math. Biol.*, 52(1-2):153–197, 1990.

[TY16] J. Tu and Y. You. Random attractor of stochastic Brusselator system with multiplicative noise. *Discrete Contin. Dyn. Syst. Ser. A*, 36(5):2757, 2016.

[vdHKM06] R. van der Hofstad, W. König, and P. Mörters. The universality classes in the parabolic Anderson model. *Commun. Math. Phys.*, 267(2):307–353, 2006.

[Ver05] F. Verhulst. *Methods and applications of singular perturbations: boundary layers and multiple timescale dynamics.* Springer, 2005.

[Ver10] M. C. Veraar. Non-autonomous stochastic evolution equations and applications to stochastic partial differential equations. *J. Evol. Equ.*, 10(1):85–127, 2010.

[VGVC07] M. Venkadesan, J. Guckenheimer, and F. J. Valero-Cuevas. Manipulating the edge of instability. *J. Biomech.*, 40(8):1653–1661, 2007.

[vN08] J. van Neerven. Stochastic evolution equations. *ISEM lecture notes*, 2008.

[Vol14] V. Volpert. *Elliptic Partial Differential Equations: Volume 2: Reaction-Diffusion Equations.* Springer, 2014.

[Wal81] J. B. Walsh. A stochastic model of neural response. *Advances in applied probability*, 13(2):231–281, 1981.

[Wal86] J. B. Walsh. An introduction to stochastic partial differential equations. In *École d'Été de Probabilités de Saint Flour XIV-1984*, pages 265–439. Springer, 1986.

[Wal00] P. Walters. *An introduction to ergodic theory.* Springer, 2000.

[Wan09a] B. Wang. Pullback attractors for the non-autonomous FitzHugh–Nagumo system on unbounded domains. *Nonlinear Anal. Theory Methods Appl.*, 70(11):3799–3815, 2009.

[Wan09b] B. Wang. Random attractors for the stochastic FitzHugh–Nagumo system on unbounded domains. *Nonlinear Anal. Theory Methods Appl.*, 71(7-8):2811–2828, 2009.

[WZ11] Z. Wang and S. Zhou. Random attractor for stochastic reaction–diffusion equation with multiplicative noise on unbounded domains. *J. Math. Anal. Appl.*, 384(1):160–172, 2011.

[Yag09] A. Yagi. *Abstract parabolic evolution equations and their applications.* Springer, 2009.

[You17] Y. You. Random dynamics of stochastic reaction–diffusion systems with additive noise. *J. Dyn. Differ. Equ.*, 29(1):83–112, 2017.

[Zam20] L. Zambotti. A brief and personal history of stochastic partial differential equations. *arXiv preprint arXiv:2004.09336*, 2020.

[Zei89] E. Zeidler. *Nonlinear Functional Analysis and Its Applications: Part 2 B: Nonlinear Monotone Operators*. Springer, 1989.

[ZG17] W. Zhao and A. Gu. Regularity of pullback attractors and random equilibrium for non-autonomous stochastic FitzHugh-Nagumo system on unbounded domains. *J. Appl. Anal. Comput.*, 7(4):1285–1311, 2017.

[ZMRS87] Y. B. Zel'Dovich, S. A. Molchanov, A. A. Ruzmaĭkin, and D. D. Sokolov. Intermittency in random media. *Soviet Physics Uspekhi*, 30(5):353, 1987.

[ZW16] S. Zhou and Z. Wang. Finite fractal dimensions of random attractors for stochastic FitzHugh–Nagumo system with multiplicative white noise. *J. Math. Anal. Appl.*, 441(2):648–667, 2016.