Small Noise Large Deviations for Infinite Dimensional Stochastic Dynamical Systems

Vasileios Maroulas

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research (Statistics).

Chapel Hill
2008

Approved by

Advisor: Amarjit Budhiraja

Reader: M. Ross Leadbetter

Reader: Chuanshu Ji

Reader: Vladas Pipiras

Reader: Gordon Simons

ABSTRACT

VASILEIOS MAROULAS: Small Noise Large Deviations for Infinite Dimensional Stochastic Dynamical Systems
(Under the direction of Amarjit Budhiraja)

Large deviations theory concerns with the study of precise asymptotics governing the decay rate of probabilities of rare events. A classical area of large deviations is the Freidlin–Wentzell (FW) theory that deals with path probability asymptotics for small noise Stochastic Dynamical Systems (SDS). For finite dimensional SDS, FW theory has been very well studied. The goal of the present work is to develop a systematic framework for the study of FW asymptotics for infinite dimensional SDS. Our first result is a general LDP for a broad family of functionals of an infinite dimensional small noise Brownian motion (BM). Depending on the application, the driving infinite dimensional BM may be given as a space–time white noise, a Hilbert space valued BM or a cylindrical BM. We provide sufficient conditions for LDP to hold for all such different model settings.

As a first application of these results we study FW LDP for a class of stochastic reaction diffusion equations. The model that we consider has been widely studied by several authors. Two main assumptions imposed in all previous studies are the boundedness of the diffusion coefficient and a certain geometric condition on the underlying domain. These restrictive conditions are needed in proofs of certain exponential probability estimates that form the basis of classical proofs of LDPs. Our proofs instead rely on some basic qualitative properties, eg. existence, uniqueness, tightness, of certain controlled analogues of the original systems. As a result, we are able to relax the two restrictive requirements described above.

As a second application, we study large deviation properties of certain stochastic diffeomorphic flows driven by an infinite sequence of i.i.d. standard real BMs. LDP for small noise finite dimensional flows has been studied by several authors. Typical space–time stochastic models with a realistic correlation structure in the spatial parameter naturally leads to infinite dimensional flows. We establish a LDP for such flows in the small noise limit. We also apply our result to a Bayesian formulation of an image analysis problem. An approximate maximum likelihood property is shown for the solution of an optimal image matching problem that involves the large deviation rate function.

To my parents, Ioannis and Sofia,

To my brothers, Nikolaos and Serafeim – Dionusios

ACKNOWLEDGEMENTS

First and foremost, I would like to express gratitude to my PhD thesis advisor, Professor Amarjit Budhiraja, for his endless encouragement in the PhD program of the Department of Statistics and Operations Research at The University of North Carolina at Chapel Hill. Professor Budhiraja was always available for both short and long periods of time (I will never forget the two and half hours meeting in early September of 2007). He wanted to be sure that at the end of our meeting a good amount of knowledge had been absorbed. Professor Budhiraja's passion for the mathematical sciences and his concrete way of explaining new concepts helped me to develop and expand my mathematical skills. Furthermore, I would like to thank Professor Budhiraja for his financial support through two of his grants received from the Army Research Office.

Next, I would like to thank my PhD Thesis Committee members: Professor M. Ross Leadbetter, for his teaching the foundations of Probability Theory, Professor Chuanshu Ji, for his infinite patience in explaining ideas and concepts of Mathematical Statistics, Professor Vladas Pipiras, for his support as a teacher and a friend, Professor Gordon Simons, for his thoughtful advice.

I could not close this section before I thank my family for its constant support and give a special thanks to my first teacher in all of my mathematical training, my father, Professor Ioannis (John) Maroulas.

CONTENTS

CHAPTER 1

Introduction

Theory of Large Deviations concerns with the study of probabilities of rare events. It is one of the most active research fields in probability, having many applications to areas such as statistical inference, queueing systems, communication networks, information theory, risk sensitive control, partial differential equations and statistical mechanics. We refer the reader to [13, 14, 15] for background, motivation, applications and fundamental results in the area.

Consider, for example, one of the classical subjects in Probability Theory–sum of i.i.d. random variables. Suppose that $X_1, X_2, \cdots$ is an i.i.d. sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with mean μ and variance σ^2. Let $S_n = \sum_{i=1}^n X_i,\ n \geq 1$ be the partial sums. Then by the strong law of large numbers (SLLN) $\frac{S_n}{n} \to \mu$ a.s., and by the central limit theorem (CLT) $\frac{S_n - n\mu}{\sigma\sqrt{n}} \to Z$ in distribution, where Z is a standard normal random variable. While the SLLN provides the convergence of the empirical average $\frac{S_n}{n}$ to μ as $n \to \infty$, the CLT gives the asymptotics of the probability of deviations of S_n from $n\mu$ by an amount of order $\sqrt{n}$. Deviations of this size are usually referred as "normal". On the other hand, deviations of order n; for example the event $A_n \doteq \{S_n \geq (\mu + a)n\},\ a > 0$, are referred as "large". Note that the probability of the above event tends to zero as $n \to \infty$. Theory of Large Deviations deals with the study of precise asymptotics governing the decay rate of such probabilities.

In particular, for the event A_n introduced above, the celebrated Cramér's theorem (cf. p.18 [13]) gives

$$-\inf_{x\in(\mu+\alpha,\infty)} I(x) \le \liminf_{n\to\infty} \frac{1}{n}\log \mathbb{P}(A_n) \le \limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(A_n) \le -\inf_{x\in[\mu+\alpha,\infty)} I(x)$$

where $I(x) \doteq \sup_{\lambda\in\mathbb{R}}\{\lambda x - \Lambda(\lambda)\}$ is the Legendre transformation of the log moment generating function of X_1, i.e. $\Lambda(\lambda) \doteq \log \mathbb{E}e^{\lambda X_1},\ \lambda \in \mathbb{R}$. The function I governs the rate of (exponential) decay of such probabilities of large deviations and is referred to as the rate function.

One of the classical areas of large deviations is the Freidlin–Wentzell theory that deals with path probability asymptotics for small noise stochastic dynamical systems. As a motivating example consider a k–dimensional stochastic differential equation of the form:

$$dX^\epsilon(t) = b\big(X^\epsilon(t)\big)dt + \sqrt{\epsilon}\, a\big(X^\epsilon(t)\big)dW(t),\ t \in [0,T], \tag{1.1}$$

with coefficients a, b satisfying suitable regularity properties and W a finite dimensional standard Brownian motion. As $\epsilon \to 0$, $X^\epsilon \xrightarrow{\mathbb{P}} X^0$ in $\mathcal{C}([0,T] : \mathbb{R}^k)$ (for a Polish space $\mathcal{E}$, $\mathcal{C}([0,T] : \mathcal{E})$ denotes the space of continuous functions from $[0,T]$ to $\mathcal{E}$), where X^0 solves the ODE $\dot{x} = b(x)$. Freidlin–Wentzell theory describes precise asymptotics (as $\epsilon \to 0$) of probabilities of events such as $\big\{\sup_{0\le t\le T} |X^\epsilon(t) - X^0(t)| > c\big\}$. For finite dimensional stochastic differential equations (SDEs) such a study is classical and we refer the reader to [20] for a comprehensive account.

The goal of the present work is to undertake a similar analysis for infinite dimensional stochastic dynamical systems. Although there are several works that have considered Freidlin–Wentzell asymptotics for infinite dimensional stochastic systems, typically proofs proceed through approximation and discretization arguments that are specific to the model under study. One of the main emphasis in the current work is to develop a unified framework for addressing such asymptotic questions for a broad family of infinite dimensional stochastic dynamical systems. Starting points of our work are the variational representation for a Hilbert space valued Brownian motion presented in Theorem 2.6.1 and a large deviation result for functionals of such Brownian motions, given in Theorem 3.3.1, that was established in [8].

The general large deviation result of Theorem 3.3.1 is our main tool in the study of small noise large deviations for stochastic differential equations driven by infinite dimensional Brownian motions. Depending on the application of interest the infinite dimensional nature of the driving noise may be given in a variety of forms. Some examples of such forms include–an infinite sequence of i.i.d standard (1–dim) Brownian motions, a cylindrical Brownian motion, a Hilbert space valued Brownian motion, a space–time Brownian sheet. We introduce in Chapter 2 all such different descriptions of an infinite dimensional Brownian motion and using relationships between them obtain, in Chapter 3, general large deviations results (analogous to Theorem 3.3.1) for functionals of Brownian motions given by such alternate descriptions. In fact we develop a somewhat strengthened form of Theorem 3.3.1 by establishing a uniform large deviation principle (LDP) with respect to some parameter (typically for SDEs, this parameter is the initial data). Uniform LDPs are needed in the study of exit time and invariant measure asymptotics for small noise Markov processes.

In Chapter 4, as a first application of the general large deviation results established in Chapter 3 (specifically Theorem 3.6.2), we will study Freidlin–Wentzell LDP for a class of reaction–diffusion stochastic partial differential equations (SPDE) [see (4.1)], for which well–posedness has been studied in [27] and a small noise LDP established in [26]. The class includes, as a special case, the reaction–diffusion SPDEs considered in [36] (see Remark 4.2.2).

Our proof of the LDP proceeds by verification of the general sufficient condition Assumption 3.3.1. The key ingredient in the verification of this assumption are the well–posedness and compactness for sequences of controlled versions of the original SPDE–see Theorems 4.2.3, 4.2.4 and 4.2.5. For comparison, the statements analogous to Theorems

4.2.3, 4.2.4 in the finite dimensional setting (1.1) would say that for any $\theta \in [0,1)$ and any L^2–bounded control u, (i.e. a predictable process satisfying $\int_0^T ||u(s)||^2 ds \leq M$, a.s. for some $M \in (0,\infty)$), and any initial condition $x \in \mathbb{R}^k$, the equation

$$dX_x^{\theta,u}(t) = b\big(X_x^{\theta,u}(t)\big)dt + \theta a\big(X_x^{\theta,u}(t)\big)dW(t) + a\big(X_x^{\theta,u}(t)\big)u(t)dt,\ X_x^{\theta,u}(0) = x \tag{1.2}$$

has a unique solution for $t \in [0,T]$. Also, the statement analogous to Theorem 4.2.5 in the finite dimensional setting would require that if $\theta(\varepsilon) \to \theta(0) = 0$, if a sequence of uniformly L^2–bounded controls u^ε satisfies $u^\epsilon \to u$ in distribution (with the weak topology on the bounded L^2-ball), and if $x^{\theta(\epsilon)} \to x$ (all as $\varepsilon \to 0$), then $X_{x^{\theta(\epsilon)}}^{\theta(\epsilon),u^\epsilon} \to X_x^{0,u}$ in distribution.

As one may expect, the techniques and estimates used to prove such properties for the original (uncontrolled) stochastic model can be applied here as well, and indeed proofs for the controlled SPDEs proceed in very much the same way as those of their uncontrolled counterparts. A side benefit of this pleasant situation is that one can often prove large deviation properties under mild conditions, and indeed conditions that differ little from those needed for a basic qualitative analysis of the original equation. In the present setting, we are able to relax two of the main technical conditions used in [26], which are the uniform boundedness of the diffusion coefficient [i.e., the function F in (4.1)] and the so called "cone condition" imposed on the underlying domain (cf. p.320 [25]). In place of these, we require only that the domain is a bounded open set and that the diffusion coefficients satisfy the standard linear growth condition. It is stated in Remark 3.2 of [26] that although unique solvability holds under the weaker linear growth condition, they are unable to derive the corresponding large deviation principle. The conditions imposed on F and $\mathcal{O}$ in [26] enter in an important way in their proofs of the large deviation principle which is based on obtaining suitable exponential tail probability estimates for certain stochastic convolutions in Hölder norms. This relies on the application of a generalization of Garsia's theorem [21], which requires the restrictive conditions alluded to above. An important point is that these conditions are not needed for unique solvability of the SPDE.

In contrast, the weak convergence proof presented here does not require any exponential probability estimates and hence these assumptions are no longer needed. Indeed, suitable exponential continuity (in probability) and exponential tightness estimates are perhaps the hardest and most technical parts of the usual proofs based on discretization and approximation arguments. This becomes particularly hard in infinite dimensional settings where these estimates are needed with metrics on exotic function spaces (e.g., Hölder spaces, spaces of diffeomorphisms, etc.).

Standard approaches to small noise LDP for infinite dimensional SDE build on the ideas of [2]. The key ingredients to the proof are as follows. One first considers an approximating Gaussian model which is obtained from the original SDE by freezing the coefficients of the right hand side according to a time discretization. Each such approximation is then further approximated by a finite dimensional system uniformly in the value of the frozen (state) variable. Next one establishes an LDP for the finite dimensional system and argues that the LDP continues to hold as one approaches the infinite dimensional model. Finally, one needs to obtain suitable exponential continuity estimates in order to obtain the LDP for the original non–Gaussian model from that for the frozen Gaussian model. Exponential continuity (in probability) and exponential tightness estimates that are used to justify these approximations are often obtained under additional conditions on the model than those needed for well posedness and compactness. In particular, as noted earlier, for the reaction diffusion systems considered here, these rely on exponential tail probability estimates in Hölder norms for certain stochastic convolutions which are only available for bounded integrands.

An alternative approach, based on nonlinear semigroup theory and infinite dimensional Hamilton–Jacobi (HJ) equations, has been developed in [18] (see also [19]). The method of proof involves showing that the value function of the limit control problem that is obtained by the law of large number analysis of certain controlled perturbations of the original stochastic model, uniquely solves an appropriate infinite dimensional HJ equation in a suitable viscosity sense. In addition, one needs to establish exponential tightness by verifying a suitable exponential compact containment estimate. Although both these steps have been verified for a variety of models (cf. [19]), the proofs are quite technical and rely on a uniqueness theory for infinite dimensional nonlinear PDEs. The uniqueness requirement on the limit HJ equation is an extraneous artifact of the approach, and different stochastic models seem to require different methods for this, in general very hard, uniqueness problem. In contrast to the weak convergence approach, it requires an analysis of the model that goes significantly beyond the unique solvability of the SPDE. In addition, as discussed previously the exponential tightness estimates are typically the most technical part of the large deviation analysis for infinite dimensional models, and are often only available under "sub–optimal" conditions when using standard techniques.

In Chapter 5 we will consider another application of the general large deviation principle established in Chapter 3. In this chapter we will prove a LDP for a wide family class of stochastic flows in the small noise limit. Stochastic flows of diffeomorphisms have been a subject of much research [4, 28, 17, 7]. In this work, we are interested in an important subclass of such flows, namely the Brownian flows of diffeomorphisms [28]. Our goal is to study small noise asymptotics, specifically, the large deviation principle (LDP) for such flows.

Elementary examples of Brownian flows are those constructed by solving finite dimensional Itô stochastic differential equations. More precisely, suppose b, f_i, $i = 1, \dots, m$ are functions from $\mathbb{R}^d \times [0,T]$ to $\mathbb{R}^d$ that are continuous in (x,t) and $(k+1)$–times continuously differentiable (with uniformly bounded derivatives) in x. Let $\beta_1, \dots \beta_m$ be independent standard real Brownian motions on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$. Then for each $s \in [0,T]$ and $x \in \mathbb{R}^d$, there is a unique continuous $\{\mathcal{F}_t\}-$adapted, $\mathbb{R}^d-$valued process $\phi_{s,t}(x)$, $s \le t \le T$, satisfying

$$\phi_{s,t}(x) = x + \int_s^t b(\phi_{s,r}(x), r)dr + \sum_{i=1}^m \int_s^t f_i(\phi_{s,r}(x), r)d\beta_i(r). \tag{1.3}$$

By choosing a suitable modification, $\{\phi_{s,t}, 0 \le s \le t \le T\}$ defines a Brownian flow of $\mathbb{C}^k-$diffeomorphisms (see Section 5.2). In particular, denoting by G^k the topological group of $\mathbb{C}^k$–diffeomorphisms (see Section 5.3 for precise definitions of the topology and the metric on G^k), one has that $\phi \equiv \{\phi_{0,t}, 0 \le t \le T\}$ is a random variable with values in the Polish space $\hat{W}_k = \mathcal{C}([0,T] : G^k)$. For $\varepsilon \in (0,\infty)$, when f_i is replaced by εf_i in (1.3), we write the corresponding flow as ϕ^ε. Large deviations for ϕ^ε in $\hat{W}_k$, as $\varepsilon \to 0$, have been studied for the case $k = 0$ in [32, 3] and for general k in [5].

As is well known [30, 4, 28], not all Brownian flows can be expressed as in (1.3) and in general one needs infinitely many Brownian motions to obtain a SDE representation for the flow. Indeed typical space–time stochastic models with a realistic correlation structure in the spatial parameter naturally lead to a formulation with infinitely many Brownian motions. One such example is given in Section 5.5. Thus, following Kunita's [28] notation for stochastic integration with respect to semi–martingales with a spatial parameter, the study of general Brownian flows of $\mathbb{C}^k$–diffeomorphisms leads to SDEs of the form

$$d\phi_{s,t}(x) = F(\phi_{s,t}(x), dt),\ \phi_{s,s}(x) = x,\ 0 \le s \le t \le T,\ x \in \mathbb{R}^d, \tag{1.4}$$

where $F(x,t)$ is a $\mathbb{C}^{k+1}$–Brownian motion (See Definition 5.2.2). Note that such an F can be regarded as a random variable with values in the Polish space $W_k = \mathcal{C}([0,T] : \mathbb{C}^{k+1}(\mathbb{R}^d))$, where $\mathbb{C}^{k+1}(\mathbb{R}^d)$ is the space of $(k+1)$–times continuously differentiable functions from $\mathbb{R}^d$ to $\mathbb{R}^d$. Representations of such Brownian motions in terms of infinitely many independent standard real Brownian motions is well known (see Exercise 3.2.10

[28]). Indeed, one can represent F as

$$F(x,t) \doteq \int_0^t b(x,r)dr + \sum_{i=1}^{\infty} \int_0^t f_i(x,r)d\beta_i(r),\ (x,t) \in \mathbb{R}^d \times [0,T], \tag{1.5}$$

where $\{\beta_i\}_{i=1}^{\infty}$ is an infinite sequence of i.i.d. real Brownian motions and b, f_i are suitable functions from $\mathbb{R}^d \times [0,T]$ to $\mathbb{R}^d$ (see below Definition 5.2.2 for details).

Letting $a(x,y,t) = \sum_{i=1}^{\infty} f_i(x,t)f_i'(y,t)$ for $x, y \in \mathbb{R}^d$, $t \in [0,T]$, the functions (a,b) are referred to as the *local characteristics* of the Brownian motion F. When equation (1.4) is driven by the Brownian motion F^ε with local characteristics $(\varepsilon a, b)$, we will denote the corresponding solution by ϕ^ε. In Chapter 5 we will establish a large deviation principle for $(\phi^\varepsilon, F^\varepsilon)$ in $\hat{W}_{k-1} \times W_{k-1}$. Note that the LDP is established in a larger space than the one in which $(\phi^\varepsilon, F^\varepsilon)$ take values (namely, $\hat{W}_k \times W_k$). This is consistent with results in [32, 3, 5], which consider stochastic flows driven by only finitely many real Brownian motions. The main technical difficulty in establishing the LDP in $\hat{W}_k \times W_k$ is the proof of a result analogous to Proposition 5.4.2, which establishes tightness of certain controlled processes, when $k-1$ is replaced by k.

The proof of our main result (Theorem 5.3.1) proceeds by verification of the general sufficient condition Theorem 3.4.1 obtained in Chapter 3. The verification of this condition essentially translates into establishing weak convergence of certain stochastic flows defined via controlled analogues of the original model (see Theorem 5.3.2). These weak convergence proofs proceed by first establishing convergence for $N-$point motions of the flow and then using Sobolev and Rellich–Kondrachov embedding theorems (see the proof of Proposition 5.4.2) to argue tightness and convergence as flows. The key point here is that the estimates needed in the proofs are precisely those that have been developed in [28] for general qualitative analysis (e.g. existence, uniqueness) of the uncontrolled versions of the flows. Unlike in [32, 3] and [5] (which consider only finite dimensional flows), the proof of the LDP does not require any exponential probability estimates or discretization/approximation of the original model.

In Section 5.5 we study an application of these results to a problem in image analysis. Stochastic diffeomorphic flows have been suggested for modeling *prior* statistical distributions on the space of possible images/targets of interest in the study of nonlinear inverse problems in this field (see [16] and references therein). Along with a data model, noise corrupted observations with such a prior distribution can then be used to compute a *posterior* distribution on this space, the "mode" of which yields an estimate of the true image underlying the observations. Motivated by such a Bayesian procedure a variational approach to this image matching problem has been suggested and analyzed in [16]. Our goal is to develop a rigorous asymptotic theory that relates standard stochastic Bayesian formulations of this problem, in the small noise limit, with the deterministic variational approach taken in [16]. This is established in Theorem 5.5.1 of Section 5.

Other possible applications of the results of Chapter 5 are as follows. Large deviations for stochastic flows were studied in [32] in order to obtain large deviation estimates for finite dimensional anticipative SDEs. The results of the current work are the first step towards the study of the analogous problem for infinite dimensional SDEs. The paper [5] used the LDP for stochastic diffeomorphic flows to study large deviation properties, as $\varepsilon \to 0$, of finite dimensional diffusions generated by $\varepsilon L_1 + L_2$, where L_1, L_2 are two second order differential operators. The analogous problem for infinite dimensional diffusions is currently open; a key ingredient is again the LDP for infinite dimensional flows obtained in the current work.

The dissertation is organized as follows. In Chapter 2 we give some background on infinite dimensional Brownian motions and recall the variational representation for a Hilbert space valued BM obtained in [8]. Chapter 3 reviews some background definitions and results in large deviations theory. We also strengthen the main large deviation result of [8] by proving a uniform LDP. Chapter 4 is devoted to the study of LDP for small noise stochastic reaction–diffusion SPDEs. Finally, in Chapter 5 we study large deviations properties of infinite dimensional stochastic flows. The two appendices cover some basic functional analysis notation, definitions and background material.

CHAPTER 2

Infinite dimensional Brownian motion

2.1 Introduction.

An infinite dimensional Brownian motion arises in a natural fashion in the study of stochastic processes with a spatial parameter. We refer the reader to [12, 25, 38], for numerous examples in physical sciences where an infinite dimensional Brownian motion is used to model the driving noise for some dynamical system. Depending on the application of interest the infinite dimensional nature of the driving noise may be expressed in a variety of forms. Some examples include–an infinite sequence of i.i.d. standard (1–dim) Brownian motions, a Hilbert space valued Brownian motion, a cylindrical Brownian motion, a space–time Brownian sheet. In this chapter, we will describe all of these models and explain how the various models are related to each other.

2.2 Infinite sequence of i.i.d Brownian motions.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with an increasing family of right continuous $\mathbb{P}$–complete sigma fields $\{\mathcal{F}_t\}_{t\geq 0}$. We will refer to $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ as a filtered probability space. Let $\{\beta_i\}_{i=1}^{\infty}$ be an infinite sequence of independent, standard, one dimensional, $\{\mathcal{F}_t\}$–Brownian motions given on this filtered probability space. We will frequently consider all our stochastic processes defined on a finite time interval $[0, T]$, where $T \in (0, \infty)$ is a fixed arbitrary terminal time. We denote by $\mathbb{R}^\infty$, the product space of countably infinite copies of the real line. Then $\beta = \{\beta_i\}_{i=1}^{\infty}$ is a random variable with values in the

Polish space $\mathcal{C}([0,T]:\mathbb{R}^\infty)$ and represents the simplest model for an infinite dimensional Brownian motion.

2.3 Hilbert space valued Brownian motion.

Frequently in applications it is convenient to express the driving noise, analogous to finite dimensional theory, as a Hilbert space valued stochastic processes. Let $(H, <\cdot,\cdot>)$ be a real separable Hilbert space. Let Q be a bounded, strictly positive, trace class operator on H. We refer the reader to Appendix A for some basic Functional Analytic definitions.

Definition 2.3.1. *An H–valued stochastic process $\{W(t), t \geq 0\}$, given on a filtered probability space defined as in Section 2.2 is called a Q–Wiener process with respect to $\{\mathcal{F}_t\}$ if for every non–zero $h \in H$,*

$$\left\{< Qh, h >^{-\frac{1}{2}} < W(t), h >,\ \{\mathcal{F}_t\}\right\}_{t\geq 0}$$

is a one–dimensional standard Wiener process.

It can be shown that if W is a Q–Wiener process as in Definition 2.3.1 then $\mathbb{P}[W \in \mathcal{C}([0,T]:H)] = 1$. Let $\{e_i\}_{i=1}^\infty$ be a complete orthonormal system (CONS) for the Hilbert space H such that $Qe_i = \lambda_i e_i$ where λ_i is the strictly positive ith eigenvalue of Q which corresponds to the eigenvector e_i. Since Q is a trace class operator, we have $\sum_{i=1}^\infty \lambda_i < \infty$. Define $\widetilde{\beta}_i(t) \doteq < W(t), e_i >,\ t \geq 0,\ i \in \mathbb{N}$. It is easy to check that $\{\widetilde{\beta}_i\}$ is a sequence of independent $\{\mathcal{F}_t\}$–Brownian motions with quadratic variation: $<< \widetilde{\beta}_i, \widetilde{\beta}_j >>_t = \lambda_i \delta_{ij} t$, where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. Setting $\beta_i = \frac{1}{\sqrt{\lambda_i}}\widetilde{\beta}_i$ we have that $\{\beta_i\}_{i=1}^\infty$ is a sequence of independent, standard, one dimensional, $\{\mathcal{F}_t\}$–Brownian motions. Thus starting from a Q–Wiener process one can produce an infinite collection of independent, standard Brownian motions in a straight forward manner. Conversely, given a collection of independent, standard Brownian motions $\{\beta_i\}_{i=1}^\infty$ and $(Q, \{e_i, \lambda_i\})$ as above one can obtain a Q–Wiener process W on setting:

$$W(t) \doteq \sum_{i=1}^{\infty} \sqrt{\lambda_i}\, \beta_i(t) e_i \tag{2.1}$$

The right hand side of (2.1) clearly converges for each fixed t in $L^2(\Omega)$. Furthermore, one can check that the series also converges in $\mathcal{C}([0,\infty) : H)$ almost surely, where $\mathcal{C}([0,\infty) : H)$ is the space of continuous functions from $[0,\infty)$ to the Hilbert space H (see Theorem 4.3, pp. 88–89, [12]). These observations lead to the following result. Throughout this work Polish spaces will always be endowed with their Borel $\sigma-$fields and explicit reference to this $\sigma-$field will be omitted in all measurability statements.

Proposition 2.3.1. *There exist measurable maps $f : \mathcal{C}([0,T] : \mathbb{R}^\infty) \longmapsto \mathcal{C}([0,T] : H)$ and $g : \mathcal{C}([0,T] : H) \longmapsto \mathcal{C}([0,T] : \mathbb{R}^\infty)$ such that $f(\beta) = W$ and $g(W) = \beta$ a.s.*

2.4 Cylindrical Brownian motion.

Equation (2.1) above can be interpreted as saying that the trace class operator Q injects a "coloring" to a white noise, namely an independent sequence of standard Brownian motions, in a manner such that the resulting process has better regularity. In some models of interest, such coloring is obtained indirectly in terms of (state dependent) diffusion coefficients. It is natural, in such situations to consider the driving noise as a "cylindrical Brownian motion" rather than a Hilbert space valued Brownian motion. Let $(H, <\cdot,\cdot>)$ be as in Section 2.3. We denote the norm on H by $||\cdot||$. Fix a filtered probability space as in Section 2.2.

Definition 2.4.1. *A family $\{B_t(h) \equiv B(t,h) : t \geq 0, h \in H\}$ of real random variables is said to be an $\{\mathcal{F}_t\}$–cylindrical Brownian motion if:*

1. *For every $h \in H$ with $||h|| = 1$, $\{B(t,h), \mathcal{F}_t\}_{t\geq 0}$ is a standard Wiener process.*

2. For every $t \geq 0$, $a_1, a_2 \in \mathbb{R}$ and $f_1, f_2 \in H$,

$$B(t, a_1 f_1 + a_2 f_2) = a_1 B(t, f_1) + a_2 B(t, f_2) \ \ a.s.$$

If $\{B_t(h) : t \geq 0, h \in H\}$ is a cylindrical Brownian motion as in Definition 2.4.1, and $\{e_i\}$ is a complete orthonormal system in H then setting $\beta_i(t) \doteq B(t, e_i)$, we see that $\{\beta_i\}$ is a sequence of independent, standard, real–valued Brownian motions. Conversely, given a sequence $\{\beta_i\}_{i=1}^{\infty}$ of independent, standard Brownian motions on a common filtered probability space,

$$B_t(h) \doteq \sum_{i=1}^{\infty} \beta_i(t) < e_i, h > \tag{2.2}$$

defines a cylindrical Brownian motion on H. The above series converges in $L^2(\Omega)$ for each fixed $t \geq 0, h \in H$ and also a.s. in $\mathcal{C}([0,\infty) : \mathbb{R})$ for each $h \in H$. Given a cylindrical Brownian motion B on H as above, one can construct a Hilbert space valued Wiener process W on an enlargement of H in a manner so that the two processes generate the same filtration. This construction proceeds as follows. Let $(H_1, < \cdot, \cdot >_1)$ be a Hilbert space such that $H_1 \supset H$, the identity map $i : H \to H_1$ is a Hilbert–Schmidt operator and $Q_1 = ii^*$ is a strictly positive trace class operator on H_1, where i^* is the adjoint operator of i. Obviously, H_1, Q_1 are not uniquely determined. For example, let $\{\lambda_i\}_{i=1}^{\infty}$ be an arbitrary sequence of strictly positive real numbers such that $\sum_{i=1}^{\infty} \lambda_i < \infty$. Define an inner product on H as

$$< h, k >_1 \doteq \sum_{i=1}^{\infty} \lambda_i < h, e_i >< e_i, k >, \ h, k \in H$$

and let H_1 be the closure of H in the above inner product. Then it is easy to check that the embedding map from $(H, < \cdot, \cdot >)$ into $(H_1, < \cdot, \cdot >_1)$ is Hilbert–Schmidt and the operator $Q_1 = ii^*$ on H_1 is a strictly positive trace class operator. The Hilbert–Schmidt embedding implies that if $\{e_i\}_{i=1}^{\infty}$ and $\{f_k\}_{k=1}^{\infty}$ are complete orthonormal system in H

and H_1, respectively, then

$$\sum_{i=1}^{\infty}\sum_{k=1}^{\infty} < e_i, f_k >_1^2 < \infty. \tag{2.3}$$

From (2.3) we infer that for each $t \geq 0$ the sequence $\{\sum_{j=1}^{n} e_j\beta_j(t)\}$ converges, in probability, in H_1 as $n \to \infty$, where $\{\beta_i\}$ is the sequence of real Brownian motions introduced in Definition 2.4.1, and

$$W^*(t) \doteq \sum_{j=1}^{\infty} e_j\beta_j(t) \tag{2.4}$$

is a Q_1–Wiener process on H_1. The series (2.4) above converges in $L^2(\Omega)$ for every t and in $\mathcal{C}([0,\infty) : H_1)$ for a.e. ω. The choice of the Hilbert space H_1 is immaterial in the following sense:

Proposition 2.4.1. *Let B be a cylindrical Brownian motion as in Definition 2.4.1 and let W^* be the Q_1–Wiener process as constructed above. Also, let $\beta = \{\beta_i\}_{i=1}^{\infty}$ be as introduced in Section 2.2. Then $\sigma\{W_s^* : 0 \leq s \leq t\} = \sigma\{B_s(h) : 0 \leq s \leq t, h \in H\} = \sigma\{\beta_i(s) : 0 \leq s \leq t, i \in \mathbb{N}\},\ t \geq 0$. In particular if X is a $\sigma\{B(s,h) : 0 \leq s \leq T, h \in H\}$ measurable random variable then there exist measurable maps $f : \mathcal{C}([0,T] : H_1) \longmapsto \mathbb{R}$ and $g : \mathcal{C}([0,T] : \mathbb{R}^{\infty}) \longmapsto \mathbb{R}$ such that $f(W^*) = g(\beta) = X$ a.s.*

2.5 Brownian sheet

In many physical dynamical systems with randomness, the driving noise is given as a space–time white noise process, also referred to as a Brownian sheet. In this section we introduce this stochastic process and describe its relationship with a cylindrical Brownian motion. We also briefly discuss stochastic integration with respect to a Brownian sheet. Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ be a filtered probability space and fix a bounded open subset $\mathcal{O} \subseteq \mathbb{R}^d$.

Definition 2.5.1. *A Gaussian family of real–valued random variables $\big\{B(t,x),\ (t,x) \in \mathbb{R}_+ \times \mathcal{O}\big\}$ on the above filtered probability space is called a Brownian sheet if*

1. $\mathbb{E}B(t,x) = 0,\ \forall (t,x) \in \mathbb{R}_+ \times \mathcal{O}$

2. $B(t,x) - B(s,x)$ is independent of $\{\mathcal{F}_s\}$, $\forall\, 0 \le s \le t$ and $x \in \mathcal{O}$

3. $Cov\big(B(t,x), B(s,y)\big) = \lambda(A_{t,x} \cap A_{s,y})$, where λ is the Lebesgue measure on $\mathbb{R}_+ \times \mathcal{O}$ and $A_{t,x} \doteq \big\{(s,y) \in \mathbb{R}_+ \times \mathcal{O}\big|\ 0 \le s \le t \text{ and } y_j \le x_j\ j = 1, \cdots, d\big\}$.

4. The map $(t,u) \mapsto B(t,u)$ from $[0,\infty) \times \mathcal{O}$ to $\mathbb{R}$ is continuous a.s.

To introduce stochastic integrals with respect to a Brownian sheet we begin with the following definitions.

Definition 2.5.2. *(Elementary and simple functions) A function $f : \mathcal{O} \times [0,T] \times \Omega \to \mathbb{R}$ is elementary if there exist $a, b \in [0,T]$, $a \le b$, a bounded $\{\mathcal{F}_a\}$–measurable random variable X and $A \in \mathcal{B}(\mathcal{O})$ such that*

$$f(x,s,\omega) = X(\omega)\mathbf{1}_{(a,b]}(s)\mathbf{1}_A(x)$$

A finite sum of elementary functions is referred to as a simple function. We denote by $\overline{\mathcal{S}}$ the class of all simple functions.

Definition 2.5.3. *(Predictable σ-field) The predictable σ–field $\mathcal{P}$ on $\Omega \times \mathbb{R}_+ \times \mathcal{O}$ is the σ–field generated by $\overline{\mathcal{S}}$. A function $f : \Omega \times \mathbb{R}_+ \times \mathcal{O} \to \mathbb{R}$ is called a predictable process if it is $\mathcal{P}$–measurable.*

For fixed $T \ge 0$ let $\mathcal{P}_2(T)$ be the class of all squared integrable and predictable processes f such that $\int_{[0,T]\times\mathcal{O}} f^2(s,x)dsdx < \infty$, a.s. Also, let $\mathcal{L}_2(T)$ be the subset of those processes that satisfy $\int_{[0,T]\times\mathcal{O}} \mathbb{E}f^2(s,x)dsdx < \infty$. We will suppress T in the notations $\mathcal{P}_2(T)$, $\mathcal{L}_2(T)$ unless needed. For all $f \in \mathcal{P}_2$ the stochastic integral $M_t(f) \doteq \int_{[0,t]\times\mathcal{O}} f(s,u)B(dsdu)$, $t \in [0,T]$ is well defined as in Chapter 2 of [38]. Furthermore for all $f \in \mathcal{P}_2$, $\{M_t(f)\}_{0 \le t \le T}$ is a continuous $\{\mathcal{F}_t\}$–local martingale which is in fact a square integrable martingale if $f \in \mathcal{L}_2$. The quadratic variation of this local martingale is given as $<< M >>_t \doteq \int_{[0,t]\times\mathcal{O}} f^2(s,x)dsdx$. Additional properties of the stochastic integral can be found in [38].

Let $\{\phi_i\}_{i=1}^{\infty}$ be a complete orthonormal system in $L^2(\mathcal{O})$ (space of real, square integrable functions on $\mathcal{O}$). Then it is easy to verify that $\beta \equiv \{\beta_i\}_{i=1}^{\infty}$ defined as $\beta_i(t) \doteq \int_{[0,t]\times\mathcal{O}} \phi_i(x)B(dsdx),\ i \geq 1,\ t \in [0,T]$ is a sequence of independent standard, real Brownian motions. Also for $(t,u) \in [0,\infty) \times \mathcal{O}$

$$B(t,x) = \sum_{i=1}^{\infty} \beta_i(t) \int \phi_i(y) \mathbf{1}_{(-\infty,x]}(y) dy, \tag{2.5}$$

where $(-\infty, x] = \{y \in \mathcal{O} \mid y_i \leq x_i, \quad \forall i = 1, \cdots, d\}$ and the above series converges in $L^2(\Omega)$ for each (t,x). From these considerations it follows that

$$\sigma\{B(t,x);\ t \in [0,T],\ x \in \mathcal{O}\} = \sigma\{\beta_i(t),\ i \geq 1,\ t \in [0,T]\}. \tag{2.6}$$

As a consequence of (2.6) we have the following result.

Proposition 2.5.1. *Let $f : \mathcal{C}([0,T] \times \mathcal{O} : \mathbb{R}) \to \mathbb{R}$ be a measurable map. Then there exists a measurable map $g : \mathcal{C}([0,T] \times \mathcal{O} : \mathbb{R}^{\infty}) \to \mathbb{R}$ such that $f(B) = g(\beta)$ a.s., where β is as defined above (2.5).*

2.6 Variational representations for infinite dimensional Brownian motions

The large deviation results established in this work critically use certain variational representations for infinite dimensional Brownian motions that we now present.

Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ be a filtered probability space and let W be an H–valued Q–Wiener process, where Q is a bounded, strictly positive, trace class operator on the Hilbert space H. Let $H_0 = Q^{1/2}H$, then H_0 is a Hilbert space with the inner product $< h, k >_0 \doteq < Q^{-1/2}h, Q^{-1/2}k >,$ where $h, k \in H_0$. Also the embedding map $i : H_0 \mapsto H$ is a Hilbert–Schmidt operator and $ii^* = Q$. Analogously to Definitions 2.5.2 and 2.5.3, a function $f : H_0 \times [0,T] \times \Omega \to \mathbb{R}$ is elementary if there exist $a, b \in [0,T],\ a \leq b$, a bounded $\{\mathcal{F}_a\}$–measurable random variable X and $A \in \mathcal{B}(H_0)$ such that $f(x,s,\omega) = X(\omega)\mathbf{1}_{(a,b]}(s)\mathbf{1}_A(x)$. Here, and throughout this work, for a Polish space $\mathcal{E}$, $\mathcal{B}(\mathcal{E})$ denotes the Borel $\sigma-$field on $\mathcal{E}$. A finite sum of elementary functions is referred to as a simple function. We denote by $\overline{\mathcal{S}}$ the class of all simple functions. The predictable σ–field $\mathcal{P}$ on $\Omega \times \mathbb{R}_+ \times H_0$ is the σ–field generated by $\overline{\mathcal{S}}$. A function $f : \Omega \times \mathbb{R}_+ \times H_0 \to \mathbb{R}$ is called a predictable process if it is $\mathcal{P}$–measurable. Define for a fixed $T \in (0,\infty)$

$$\mathcal{A} = \Big\{\phi \mid \phi \text{ is } H_0\text{–valued, } \{\mathcal{F}_t\}\text{–predictable process and } \mathbb{P}\big[\int_0^T ||\phi(s)||_0^2 ds < \infty\big] = 1\Big\}. \tag{2.7}$$

Theorem 2.6.1. *(Budhiraja and Dupuis [8]) Let f be a bounded, Borel measurable function mapping $\mathcal{C}([0,T] : H)$ into $\mathbb{R}$. Then,*

$$-\log \mathbb{E}(\exp\{-f(W)\}) = \inf_{u \in \mathcal{A}} \mathbb{E}\big(\frac{1}{2}\int_0^T ||u(s)||_0^2 ds + f(W + \int_0^{\cdot} u(s)ds)\big).$$

As immediate corollary to Theorem 2.6.1, Proposition 2.4.1 and 2.5.1, we have the following representation theorems for a cylindrical Brownian motion and a Brownian sheet respectively. Define $\mathcal{A}^*$ to be the class of H–valued $\{\mathcal{F}_t\}$–predictable processes ϕ, satisfying $\mathbb{P}\{\int_0^T ||\phi(s)||^2 ds < \infty\} = 1$.

Corollary 2.6.1. *Let $\{e_i\}$, as before, be a complete orthonormal system in H and let $\{B_t(h) : 0 \le t \le T, h \in H\}$ be an $\{\mathcal{F}_t\}$–cylindrical Brownian motion. Let X be a bounded random variable which is measurable with respect to $\sigma\{B_s(h) : 0 \le s \le T, h \in H\}$. Then,*

$$\begin{aligned} -\log \mathbb{E}(\exp\{-X\}) &= \inf_{u \in \mathcal{A}^*} \mathbb{E}\big(\frac{1}{2}\int_0^T \parallel u(s) \parallel^2 ds + f(W^*(\cdot) + \int_0^{\cdot} u(s)ds)\big) \\ &= \inf_{u \in \mathcal{A}^*} \mathbb{E}\Big(\frac{1}{2}\int_0^T \parallel u(s) \parallel^2 ds + g\big(\beta^u\big)\Big) \end{aligned} \tag{2.8}$$

where $\beta^u \doteq \{B_i^u = B(e_i) + \int_0^{\cdot} u_i(s)ds\}_{i=1}^{\infty}$, $u_i(s) =< u(s), e_i >$ and f, g, W^* are related to B and X as in Lemma 2.4.1.

Corollary 2.6.2. *Let $f : \mathcal{C}([0,T] \times \mathcal{O} : \mathbb{R}) \to \mathbb{R}$ be a bounded measurable map. Let B be a Brownian sheet as in Definition 2.5.1. Then,*

$$-\log \mathbb{E}(\exp\{-f(B)\}) = \inf_{u \in \mathcal{P}_2} \mathbb{E}\big(\frac{1}{2}\int_0^T \int_{\mathcal{O}\cap(-\infty,x]} u^2(s,r)drds + f(B^u)\big) \tag{2.9}$$

where $B^u(t,x) = B(t,x) + \int_0^t \int_{[0,x]\cap\mathcal{O}} u(s,y)dyds$ and $[0,x] = \{y \in \mathcal{O} \mid 0 \le y_i \le x_i, \ \forall i = 1, \cdots, d\}$.

CHAPTER 3

Large deviations

3.1 Introduction

In this chapter we present the general large deviation principle for functionals of a Hilbert space valued Brownian motion that was established in [8]. Next, using results of Chapter 2 we translate this result in Sections 3.4 and 3.5 to settings where the underlying infinite dimensional noise is given as a cylindrical Brownian motion or as a space–time white noise. Finally in Section 3.6 we establish a strengthened version of the result in [8] (and the results in Sections 3.4 and 3.5) by proving a uniform (in initial condition) large deviation principle. Such uniform large deviation estimates are critical in the study of exit time and invariant measures asymptotics for the corresponding stochastic processes. We begin with some basic definitions and background results in Large Deviations Theory.

3.2 Large deviation principle and the Laplace principle

Let $\{X^\epsilon, \epsilon > 0\}$ be a collection of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in a Polish space (i.e., complete separable metric space) $\mathcal{E}$. Denote the metric on $\mathcal{E}$ as $d(\cdot,\cdot)$ and expectation with respect to $\mathbb{P}$ by $\mathbb{E}$. The theory of large deviations concerns with events A for which the probabilities $\mathbb{P}(X^\epsilon \in A)$ converge to zero exponentially fast as $\epsilon \to 0$. The exponential decay rate of such probabilities is typically expressed in terms of a " rate function" I mapping $\mathcal{E}$ into $[0, \infty]$.

Definition 3.2.1. *(Rate function) A function $I : \mathcal{E} \to [0, \infty]$ is called a rate function on $\mathcal{E}$ if for each $M < \infty$ the level set $\{x \in \mathcal{E} : I(x) \leq M\}$ is a compact subset of $\mathcal{E}$.*

Remark 3.2.1. 1. *The compactness of level sets implies that a rate function is a lower semi–continuous function. In many works, a rate function is defined with the compact level set requirement replaced by the statement that I is lower semi–continuous; while a function that in addition has compact level sets, is referred to as a good rate function. In our work all rate functions will be good and so the adjective "good" will be omitted.*

2. *With an abuse of notation for a subset A of $\mathcal{E}$ we write $I(A) \doteq \inf_{x \in A} I(x)$.*

Definition 3.2.2. *(Large deviation principle) Let I be a rate function on $\mathcal{E}$. The sequence $\{X^\epsilon\}$ is said to satisfy the large deviation principle on $\mathcal{E}$, as $\epsilon \to 0$, with rate function I if the following two conditions hold.*

1. ***Large deviation upper bound.*** *For each closed subset F of $\mathcal{E}$*

$$\limsup_{\epsilon \to 0} \epsilon \log \mathbb{P}(X^\epsilon \in F) \leq -I(F). \tag{3.1}$$

2. ***Large deviation lower bound.*** *For each open subset G of $\mathcal{E}$*

$$\liminf_{\epsilon \to 0} \epsilon \log \mathbb{P}(X^\epsilon \in G) \geq -I(G). \tag{3.2}$$

Remark 3.2.2. *If a sequence of random variables satisfies the large deviation principle with some rate function, then the rate function is unique (see Theorem 1.3.1 pp. 17–18 [15]).*

In many problems one is interested in obtaining exponential estimates on functions which are more general than indicator functions of closed or open sets. This leads to the study of the, so called, Laplace principle.

Definition 3.2.3. *(Laplace principle) Let I be rate function on $\mathcal{E}$. The family $\{X^\epsilon\}$ is said to satisfy the Laplace principle on $\mathcal{E}$, as $\epsilon \to 0$, with rate function I, if for all bounded continuous functions $h : \mathcal{E} \to \mathbb{R}$, the following two conditions hold.*

1. ***Laplace principle upper bound.***

$$\limsup_{\epsilon \to 0} \epsilon \log \mathbb{E}\{exp[-\frac{1}{\epsilon} h(X^\epsilon)]\} \leq -\inf_{x \in \mathcal{E}}\{h(x) + I(x)\}. \tag{3.3}$$

2. ***Laplace principle lower bound.***

$$\liminf_{\epsilon \to 0} \epsilon \log \mathbb{E}\{exp[-\frac{1}{\epsilon} h(X^\epsilon)]\} \geq -\inf_{x \in \mathcal{E}}\{h(x) + I(x)\}. \tag{3.4}$$

One of the main results of the theory is the equivalence between the Laplace principle and the large deviation principle. This equivalence can be regarded as an analogue of the Portmanteau theorem in the theory of weak convergence of probability measures $\big($ see, e.g. Theorem 2.1, p.16 [6]$\big)$. The precise result describing this equivalence is as follows. For a proof we refer the reader to Section 1.2 of [15].

Theorem 3.2.1. *The family $\{X^\epsilon\}_{\epsilon>0}$ satisfies the Laplace principle upper (respectively lower) bound for all bounded and continuous functions h with a rate function I on $\mathcal{E}$ if and only if $\{X^\epsilon\}_{\epsilon>0}$ satisfies the large deviation upper (respectively lower) bound for all closed sets (respectively open sets) with the rate function I.*

3.3 A general large deviation principle

In this section we present a large deviation principle established in [8], for functionals of a Hilbert space valued Wiener process. Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$, $(H, < \cdot, \cdot >)$, Q be as in

Section 2.3. Let W be an $H-$valued Wiener process with trace class covariance Q given on the above filtered probability space. Let $\mathcal{E}$ be a Polish space and for each $\epsilon > 0$ let $\mathcal{G}^\epsilon : \mathcal{C}([0,T] : H) \to \mathcal{E}$ be a measurable map. We now present a set of sufficient conditions for large deviation principle to hold for the family $\{X^\epsilon \doteq \mathcal{G}^\epsilon(\sqrt{\epsilon}W), \epsilon > 0\}$ as $\epsilon \to 0$. For a Hilbert space H, let $L^2([0,T] : H)$ denote the Hilbert space of measurable maps f from $[0,T]$ to H satisfying $\int_0^T |f(s)|^2 ds < \infty$. Let H_0 be as in Section 2.6 and define for $N \in \mathbb{N}$

$$S_N \doteq \Big\{u \in L^2([0,T] : H_0) : \int_0^T ||u(s)||_0^2 ds \leq N\Big\}. \tag{3.5}$$

$$\mathcal{A}_N \doteq \{u \in \mathcal{A} : u(\omega) \in S_N, \mathbb{P} - a.s.\}. \tag{3.6}$$

It is easy to check that S_N is a compact metric space with the metric $d_1(x,y) = \sum_{i=1}^\infty \frac{1}{2^i} |\int_0^T < x(s) - y(s), e_i(s) >_0 ds|$. Henceforth, wherever we refer to S_N, we will consider it endowed with the topology obtained from the metric d_1 and refer to this topology as weak topology on S_N. We will abbreviate the statement "X_n converges to X in distribution" as $X_n \xrightarrow{d} X$.

Assumption 3.3.1. *There exists a measurable map $\mathcal{G}^0 : \mathcal{C}([0,T] : H) \to \mathcal{E}$ such that the following hold:*

1. *For every $M < \infty$ the set*

$$\Gamma_M \doteq \{\mathcal{G}^0(\int_0^\cdot u(s)ds) : u \in S_M\} \tag{3.7}$$

   is a compact subset of $\mathcal{E}$.

2. *Consider $M < \infty$ and a family $\{u^\epsilon\} \subset \mathcal{A}_M$, such that u^ϵ converges in distribution (as S_M−valued random elements) to u. Then*

$$\mathcal{G}^\epsilon(\sqrt{\epsilon}W(\cdot) + \int_0^\cdot u^\epsilon(s)ds) \xrightarrow{d} \mathcal{G}^0(\int_0^\cdot u(s)ds). \tag{3.8}$$

The following result was presented (without proof) in [8]. We provide a proof below for the sake of completeness.

Proposition 3.3.1. *For $f \in \mathcal{E}$ define*

$$I(f) = \inf_{\{u \in L^2([0,T]:H_0): f = \mathcal{G}^0(\int_0^\cdot u(s)ds)\}} \{\frac{1}{2}\int_0^T ||u(s)||_0^2 ds\}, \tag{3.9}$$

where $\mathcal{G}^0 : \mathcal{C}([0,T] : H) \to \mathcal{E}$ is a measurable map. If $\mathcal{G}^0$ satisfies the first condition in Assumption 3.3.1, then I is a rate function on $\mathcal{E}$.

Proof. We need to show that the set $E_M \doteq \{f \in \mathcal{E} : I(f) \leq M\}$ is compact $\forall M \in (0, \infty)$. It suffices to show that $E_M = \bigcap_{n \geq 1} \Gamma_{2M+\frac{1}{n}}$ since, by assumption, for each $k \in \mathbb{R}_+$, Γ_k is compact. Let $f \in E_M$. Then for every $n \geq 1$, there exists $u \in L^2([0,T] : H_0)$ such that $\frac{1}{2}\int_0^T ||u(s)||_0^2 ds \leq M + \frac{1}{2n}$ and $f = \mathcal{G}^0(\int_0^\cdot u(s)ds)$. In particular, $u \in S_{2M+\frac{1}{n}}$ and thus $f \in \Gamma_{2M+\frac{1}{n}}$. Since $n \geq 1$ is arbitrary we get $f \in \bigcap_{n\geq 1}\Gamma_{2M+\frac{1}{n}}$. Thus $E_M \subseteq \bigcap_{n\geq 1}\Gamma_{2M+\frac{1}{n}}$. Conversely, let $f \in \Gamma_{2M+\frac{1}{n}}$ for all $n \geq 1$ and let $u_n \in S_{2M+\frac{1}{n}}$ be such that $f = \mathcal{G}^0(\int_0^\cdot u_n(s)ds)$. Then $I(f) \leq \frac{1}{2}\int_0^T ||u_n(s)||^2 ds \leq \frac{1}{2}(2M + \frac{1}{n})$. Sending $n \to \infty$ we get that $I(f) \leq M$. Thus $f \in E_M$ and the inclusion $\bigcap_{n\geq 1}\Gamma_{2m+\frac{1}{n}} \subseteq E_M$ follows. Combining the two set inclusions established above we have the result. □

The following is the main result of [8].

Theorem 3.3.1. *(Budhiraja and Dupuis [8]) Let $\mathcal{G}^\epsilon$ and X^ϵ be as introduced above. Suppose that Assumption 3.3.1 holds. Then the family $\{X^\epsilon, \epsilon > 0\}$ satisfies the large deviation principle on $\mathcal{E}$, as $\epsilon \to 0$, with rate function I defined in (5.7).*

3.4 Large deviation principle for functionals of a cylindrical Brownian motion

In the previous section a large deviation principle for certain functionals of a Hilbert space valued Brownian motion was established. In this section, using Proposition 2.4.1, which relates a cylindrical Brownian motion with a Hilbert space valued Brownian motion, we establish a large deviation principle for a family of random elements measurable with respect to a cylindrical Brownian motion. Let $(H, < \cdot, \cdot >)$ be a separable Hilbert space and $\{B(t,h) \equiv B_t(h) : h \in H\}$ be a cylindrical Brownian motion defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$. Analogous to sets defined in, (2.7), (3.5) and (3.6), we define:

$$\mathcal{A}[l_2] = \Big\{\phi \equiv \{\phi\}_{i=1}^{\infty} |\ \phi_i : [0,T] \to \mathbb{R}\ \text{, is } \{\mathcal{F}_t\}\text{–predictable for all } i \\ \text{and } \mathbb{P}\{\int_0^T ||\phi(s)||_{l_2}^2 ds < \infty\} = 1\Big\} \tag{3.10}$$

where $\int_0^T ||\phi(s)||_{l_2}^2 ds \doteq \sum_{i=1}^{\infty} \int_0^T |\phi_i(s)|^2 ds$. Note that $\mathcal{A}[l_2]$ is same as $\mathcal{A}$ defined in (2.7) with H_0 there replaced by the Hilbert space $l_2 = \{\underline{x} = (x_1, x_2, \cdots) : x_i \in \mathbb{R}, i \geq 1 \text{ and } \sum_{i=1}^{\infty} x_i^2 < \infty\}$ (the inner product on l_2 is defined as $< x, y >_{l_2} \doteq \sum_{i=1}^{\infty} x_i y_i,\ x, y \in l_2$). Define,

$$S_N[l_2] = \Big\{\phi \equiv \{\phi_i\}_{i=1}^{\infty} \text{ in } L^2([0,T] : l^2) \text{ and } \int_0^T ||\phi(s)||_0^2 ds \leq N\Big\}. \tag{3.11}$$

Recall that $S_N[l_2]$ is a compact Polish space endowed with the weak topology. Finally define,

$$\mathcal{A}_N[l_2] = \Big\{u \in \mathcal{A}[l_2] : u(\omega) \in S_N,\ \mathbb{P}\text{–a.s.}\Big\}. \tag{3.12}$$

In the rest of this section we will write $\mathcal{A}[l_2], S_N[l_2], \mathcal{A}_N[l_2]$ as $\mathcal{A}, S_N, \mathcal{A}_N$ respectively. Let $\{e_i\}$ be a complete orthonormal system in H and let $\beta_i(t) \doteq B_t(e_i),\ 0 \leq t \leq T,\ i \geq 1$. Noting that $\beta \equiv \{\beta_i\}$ is a sequence of independent, standard, real Brownian motions, we have that β is a $\Big(\mathcal{C}([0,T] : \mathbb{R}^{\infty}), \mathcal{B}\big(\mathcal{C}([0,T] : \mathbb{R}^{\infty})\big)\Big) \equiv (S, \mathcal{S})$ valued random variable, where $\mathbb{R}^{\infty}$ denotes the product space of countably many copies of the real line. Note that S is a Polish space and $\mathcal{S}$ is its corresponding Borel σ–field. We now introduce the analog of Assumption 3.3.1 for the case of a cylindrical Brownian motion. Let $\mathcal{E}$

be a Polish space and for each $\epsilon > 0$, $\mathcal{G}^\epsilon : S \to \mathcal{E}$, $\epsilon > 0$ be a measurable map. The following is the main assumption for the large deviation principle to hold for the family $\Big\{X^\epsilon \doteq \mathcal{G}^\epsilon\big(\sqrt{\epsilon}\beta\big),\ \epsilon > 0\Big\}$.

Assumption 3.4.1. *There exists a measurable map $\mathcal{G}^0 : S \to \mathcal{E}$ such that the following hold:*

1. *For every $M < \infty$ the set*

$$\Gamma_M \doteq \{\mathcal{G}^0(\int_0^{\cdot} u(s)ds) : u \in S_M\} \tag{3.13}$$

   is a compact subset of $\mathcal{E}$.

2. *Consider $M < \infty$ and a family $\{u^\epsilon\} \subset \mathcal{A}_M$, such that u^ϵ converges in distribution (as S_M–valued random elements) to u, then*

$$\mathcal{G}^\epsilon(\sqrt{\epsilon}\beta + \int_0^{\cdot} u^\epsilon(s)ds) \xrightarrow{d} \mathcal{G}^0(\int_0^{\cdot} u(s)ds). \tag{3.14}$$

Define for each $f \in \mathcal{E}$

$$I(f) = \inf_{\{u \in L^2([0,T]:l_2) : f = \mathcal{G}^0(\int_0^{\cdot} u(s)ds)\}} \{\frac{1}{2}\int_0^T ||u(s)||_{l_2}^2 ds\}, \tag{3.15}$$

where $\mathcal{G}^0 : S \to \mathcal{E}$ is a measurable map.

Theorem 3.4.1. *Let $\mathcal{G}^0 : S \to \mathcal{E}$ be a measurable map satisfying the first condition in Assumption 3.4.1. Then $I : \mathcal{E} \to [0, \infty]$, defined by (3.15) is a rate function on $\mathcal{E}$. If, furthermore $\{\mathcal{G}^\epsilon\}_{\epsilon>0}$ and $\mathcal{G}^0$ satisfy the second condition of Assumption 3.4.1 then the family $\{X^\epsilon, \epsilon > 0\}$ satisfies the large deviation principle, as $\epsilon \to 0$, with rate function I.*

Proof. Let $\{\lambda_i\}_{i=1}^\infty$ be a sequence of positive real numbers such that $\sum_{i=1}^\infty \lambda_i < \infty$. Define $l_2^* \doteq \Big\{x \equiv (x_1, x_2, ...) |\ x_i \in \mathbb{R} \text{ and } \sum \lambda_i x_i^2 < \infty\Big\}$. Note that l_2^* is a Hilbert space with inner product $< x, y >_{l_2^*} \doteq \sum \lambda_i x_i y_i$. Also note that $l_2 \subseteq l_2^*$ and the embedding map is a Hilbert–Schmidt operator. It can be easily checked that $\mathcal{C}([0,T] : l_2^*)$ is a measurable subset of S–in fact the embedding of $\mathcal{C}([0,T] : l_2^*)$ into $\mathcal{C}([0,T] : \mathbb{R}^\infty)$ is continuous. Note that $\widehat{\mathcal{G}}^0 \doteq \mathcal{G}^0\big|_{\mathcal{C}([0,T]:l_2^*)}$ is a measurable map from $\mathcal{C}([0,T] : l_2^*)$ to $\mathcal{E}$ satisfying the first part of Assumption 3.3.1 with H there replaced by l_2^* and H_0 by l_2. Thus the first part of the theorem follows from Proposition 3.3.1. Next note that $\mathbb{P}\big[\beta \in \mathcal{C}([0,T] : l_2^*)\big] = 1$ and is a l_2^*–valued Wiener process with covariance operator Q defined as $(Qx)_i = \lambda_i x_i,\ i = 1, 2, \cdots$. Let $\widehat{\mathcal{G}}^\epsilon \doteq \mathcal{G}^\epsilon\big|_{\mathcal{C}([0,T]:l_2^*)}$. Then $(\widehat{\mathcal{G}}^\epsilon, \widehat{\mathcal{G}}^0)$ satisfy Assumption 3.3.1 with W replaced by β and H_0, H there replaced by l_2 and l_2^* respectively. Finally, observing that $X^\epsilon = \widehat{\mathcal{G}}^\epsilon(\sqrt{\epsilon}\beta)$ the result follows from Theorem 3.3.1. □

3.5 Large deviation principle for functionals of a Brownian sheet

In this section, using Lemma 2.5.1 and Theorem 3.4.1, we establish a large deviation principle for a family of random elements measurable with respect to a Brownian sheet. Let $\mathcal{O}$ be a bounded open subset of $\mathbb{R}^d$ and $\{B(t,x) : (t,x) \in \mathbb{R}_+ \times \mathcal{O}\}$ be a Brownian sheet defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ (see Section 2.5). We denote by $L^2([0,T] \times \mathcal{O})$ the space of real (Lebesgue) square integrable functions on $[0,T] \times \mathcal{O}$. Analogous to classes defined in, (2.7), (3.5) and (3.6), we define $\mathcal{A}^{\text{BS}} \doteq \mathcal{P}_2(T)$, where $\mathcal{P}_2(T)$ is as introduced in Section 2.5, and,

$$S_N^{\text{BS}} \doteq \Big\{\phi \in L^2([0,T] \times \mathcal{O}) : \int_{[0,T]\times\mathcal{O}} \phi^2(s,r)drds \le N\Big\} \tag{3.16}$$

$$\mathcal{A}_N^{\text{BS}} \doteq \big\{u \in \mathcal{P}_2(T) : u(\omega) \in S_N,\ \mathbb{P}\text{–a.s.}\big\} \tag{3.17}$$

We now introduce the analog of Assumption 3.3.1 (or equivalently Assumption 3.4.1) for the case of a Brownian sheet. Let $\mathcal{E}$ be a Polish space and $\mathcal{G}^\epsilon : \mathcal{C}([0,T] \times \mathcal{O} : \mathbb{R}) \to \mathcal{E},\ \epsilon > 0$ be a family of measurable maps. The following is the main assumption for the large deviation principle to hold for the family $\big\{X^\epsilon \doteq \mathcal{G}^\epsilon(\sqrt{\epsilon}B),\ \ \epsilon > 0)\big\}$. For $u \in L^2([0,T]\times\mathcal{O})$, define $\mathcal{I}(u) \in \mathcal{C}([0,T]\times\mathcal{O}:\mathbb{R})$ as

$$\mathcal{I}(u)(t,x) \doteq \int_{[0,t]\times(\mathcal{O}\cap(-\infty,x])} u(s,y)dsdy, \tag{3.18}$$

where as before $(-\infty,x] = \{y \in \mathcal{O} \mid y_i \le x_i,\ \forall i = 1,\cdots,d\}$.

Assumption 3.5.1. *There exists a measurable map $\mathcal{G}^0 : \mathcal{C}([0,T]\times\mathcal{O}:\mathbb{R}) \to \mathcal{E}$ such that the following hold:*

1. *For every $M < \infty$ the set*

$$\Gamma_M \doteq \Big\{\mathcal{G}^0\big(\mathcal{I}(u)\big) : u \in S_M\Big\} \tag{3.19}$$

   is a compact subset of $\mathcal{E}$.

2. *Consider $M < \infty$ and a family $\{u^\epsilon\} \subset \mathcal{A}_M$, such that u^ϵ converges in distribution (as S_M–valued random elements) to u. Then*

$$\mathcal{G}^\epsilon\big(\sqrt{\epsilon}B + \mathcal{I}(u^\epsilon)\big) \xrightarrow{d} \mathcal{G}^0\big(\mathcal{I}(u)\big) \tag{3.20}$$

Define for $f \in \mathcal{E}$

$$I(f) = \inf_{\{u\in L^2([0,T]\times\mathcal{O}):f=\mathcal{G}^0(\mathcal{I}(u))\}} \Big\{\frac{1}{2}\int_{[0,T]\times\mathcal{O}} u^2(s,r)drds\Big\}. \tag{3.21}$$

Theorem 3.5.1. *Let $\mathcal{G}^0 : \mathcal{C}([0,T]\times\mathcal{O}:\mathbb{R}) \to \mathcal{E}$ be a measurable map satisfying the first condition in Assumption 3.5.1. Then $I : \mathcal{E} \to [0,\infty]$, defined by (3.21), is a rate function on $\mathcal{E}$. Furthermore, if $\{\mathcal{G}^\epsilon\}_{\epsilon>0}$ and $\mathcal{G}^0$ satisfy in addition the second part of Assumption 3.5.1 then the family $\{X^\epsilon\}_{\epsilon>0}$ satisfies the large deviation principle, as $\epsilon \to 0$, with rate function I.*

Proof. Let $\{\phi_i\}$ be a CONS in $L^2(\mathcal{O})$. Define $\beta_i(t) \doteq \int_0^t \int_{\mathcal{O}} \phi_i(x)B(dsdx),\ t \in [0,T]$. Then $\beta \equiv \{\beta\}_{i=1}^{\infty}$ is a l_2^*–valued Brownian motion, where l_2^* is as in the proof of Theorem 3.4.1. From Lemma 2.5.1 we have that there is a measurable map $\Psi : \mathcal{C}([0,T] : l_2^*) \to \mathcal{C}([0,T] \times \mathcal{O} : \mathbb{R})$ such that $\Psi(\beta) = B$ a.s. An application of Girsanov's theorem shows that for every $u \in L^2([0,T] \times \mathcal{O} : \mathbb{R})$, $\Psi(\beta^u) = B + \mathcal{I}(u)$ a.s., where $\beta^u \equiv \{\beta_i^u\}_{i=1}^{\infty}$; $\beta_i^u(t) = \beta_i(t) + \int_0^t \int_{\mathcal{O}} \phi_i(y)u(s,y)dsdy$. Also note that if $g \equiv \{g_i\}_{i=1}^{\infty} \in L^2([0,T] : l_2)$, then defining

$$u_g(t,y) \doteq \sum_{i=1}^{\infty} \phi_i(y)g_i(t), \quad (t,y) \in [0,T] \times \mathcal{O}, \tag{3.22}$$

we have that $u_g \in L^2([0,T] \times \mathcal{O} : \mathbb{R})$ and $\beta^{u_g} = \beta + \int_0^{\cdot} g(s)ds$ a.s. Define the maps $\hat{\mathcal{G}}^{\epsilon}, \hat{\mathcal{G}}^0$ from $\mathcal{C}([0,T] : l_2^*)$ to $\mathcal{E}$ as $\hat{\mathcal{G}}^{\epsilon}(\sqrt{\epsilon}b) \doteq \mathcal{G}^{\epsilon}(\sqrt{\epsilon}\Psi(b))$, $b \in \mathcal{C}([0,T] : l_2^*)$ and $\hat{\mathcal{G}}^0(b) \doteq \mathcal{G}^0\big(\mathcal{I}(u_g)\big)$ if $b = \int_0^{\cdot} g_s ds$ for some $g \in L^2([0,T] : l^2)$, and $\hat{\mathcal{G}}^0(b) \doteq 0$ otherwise. Here u_g is given by (3.22). It is easy to check that Assumption 3.3.1 is satisfied with $(\mathcal{G}^0, \mathcal{G}^{\epsilon})$ there replaced by $(\hat{\mathcal{G}}^0, \hat{\mathcal{G}}^{\epsilon})$, H replaced by l_2^* and W by β. Thus from Theorem 3.3.1, $\mathcal{G}^{\epsilon}(\sqrt{\epsilon}B) = \hat{\mathcal{G}}^{\epsilon}(\sqrt{\epsilon}\beta)$ satisfies the large deviation principle on $\mathcal{E}$ with rate function

$$\hat{I}(f) = \inf_{\{g \in L^2([0,T]:l_2) : f = \hat{\mathcal{G}}^0(\int_0^{\cdot} g)\}} \Big\{ \frac{1}{2} \int_0^T ||g(t)||_{l_2}^2 dt \Big\} = I(f).$$

This proves the result. □

3.6 Uniform large deviation principle

Let $\mathcal{E}_0$ and $\mathcal{E}$ be Polish spaces and for each $\epsilon > 0$ $\mathcal{G}^{\epsilon} : \mathcal{E}_0 \times S \to \mathcal{E}$ be a measurable map, where S is as in Section 3.4 . For $x \in \mathcal{E}_0$ let

$$X^{\epsilon,x} \doteq \mathcal{G}^{\epsilon}(x, \sqrt{\epsilon}\beta), \tag{3.23}$$

where β is as in Section 2.2. In this section we will give sufficient conditions for the Laplace principle for $\{X^{\epsilon,x}\}$ to hold uniformly in x for compact subsets of $\mathcal{E}_0$. We begin with the following definitions.

Definition 3.6.1. *A family of rate functions I_x on $\mathcal{E}$ parameterized by $x \in \mathcal{E}_0$ is said to have compact level sets on compacts if for all compact subsets K of $\mathcal{E}_0$ and each $M < \infty$ $\Lambda_{M,K} \doteq \cup_{x \in K}\{f \in \mathcal{E} : I_x(f) \leq M\}$ is a compact subset of $\mathcal{E}$.*

Definition 3.6.2. *Let I_x be a family of rate functions on $\mathcal{E}$ parameterized by $x \in \mathcal{E}_0$ and assume that this family has compact level sets on compacts. The family $\{X^{\epsilon,x}\}$ is said to satisfy the Laplace principle on $\mathcal{E}$, as $\epsilon \to 0$, with rate function I_x, uniformly for x in compacts, if for all compact subsets K on $\mathcal{E}_0$ and all bounded continuous functions h mapping $\mathcal{E}$ into $\mathbb{R}$:*

$$\lim_{\epsilon \to 0} \sup_{x \in K} \Big| \epsilon \log \mathbb{E}\big\{\exp[-\frac{1}{\epsilon} h(X^{\epsilon,x})]\big\} - F(x,h) \Big| = 0,$$

where $F(x,h) \doteq -\inf_{f \in \mathcal{E}}\{h(f) + I_x(f)\}$.

We now formulate a sufficient condition for the validity of uniform LDP for the family $\{X^{\epsilon,x}\}$.

Assumption 3.6.1. *There exists a measurable map $\mathcal{G}^0 : \mathcal{E}_0 \times S \to \mathcal{E}$ such that the following hold:*

1. *For every $M < \infty$ and compact set $K \subseteq \mathcal{E}_0$ the set*

$$\Gamma_{M,K} \doteq \Big\{\mathcal{G}^0\big(x, \int_0^{\cdot} u(s)ds\big) : u \in S_M[l_2], x \in K\Big\} \tag{3.24}$$

   is a compact subset of $\mathcal{E}$.

2. *Consider $M < \infty$ and a family $\{u^\epsilon\} \subset \mathcal{A}_M[l_2]$, such that u^ϵ converges in distribution (as $S_M[l_2]$–valued random elements) to u. Let $x_\epsilon \in \mathcal{E}_0$ be such that $x_\epsilon \to x$ as $\epsilon \to 0$. Then as $\epsilon \to 0$,*

$$\mathcal{G}^\epsilon\Big(x_\epsilon, \sqrt{\epsilon}\beta + \int_0^{\cdot} u^\epsilon(s)ds\Big) \xrightarrow{d} \mathcal{G}^0\Big(x, \int_0^{\cdot} u(s)ds\Big). \tag{3.25}$$

Theorem 3.6.1. *Let $X^{\epsilon,x}$ be as in (3.23) and suppose that Assumption 3.6.1 holds. Let for $x \in \mathcal{E}_0$ and $f \in \mathcal{E}$, $I_x(f)$ be defined as:*

$$I_x(f) = \inf_{\{u \in L^2([0,T]:l_2): f = \mathcal{G}^0(x, \int_0^\cdot u(s)ds)\}} \{\frac{1}{2}\int_0^T \sum_{i=1}^{\infty} u_i^2(s)ds\}. \tag{3.26}$$

Suppose that for all $f \in \mathcal{E}$, $x \mapsto I_x(f)$ is a lower semi–continuous map from $\mathcal{E}_0$ to $[0,\infty]$. Then the family $\{I_x(\cdot),\ x \in \mathcal{E}_0\}$ of rate functions has compact level sets on compacts. Furthermore, the family $\{X^{\epsilon,x}\}_{\epsilon>0}$ satisfies the large deviation principle on $\mathcal{E}$, with rate function I_x, uniformly on compact subsets of $\mathcal{E}_0$.

Proof. For the first part of the theorem we will proceed as in the proof of Proposition 3.3.1, i.e., we will show that $\Lambda_{M,K}$ equals to $\bigcap_{n\geq 1}\Gamma_{2M+\frac{1}{n},K}$. In view of Assumption 3.6.1, the compactness of $\Lambda_{M,K}$ will then follow. Let $f \in \Lambda_{M,K}$. There exists $x \in K$ such that $I_x(f) \leq M$. We can now find for each $n \geq 1$, $u_n \in L^2([0,T]:l_2)$ such that $f = \mathcal{G}^0\big(x, \int_0^\cdot u_n(s)ds\big)$ and $\frac{1}{2}\sum_{i=1}^{\infty}\int_0^T u_{n,i}^2(s)ds \leq M + \frac{1}{2n}$. In particular $u_n \in S_{2M+\frac{1}{n}}$ and so $f \in \Gamma_{2M+\frac{1}{n},K}$. Since $n \geq 1$ is arbitrary, we have $\Lambda_{M,K} \subseteq \bigcap_{n\geq 1}\Gamma_{2M+\frac{1}{n},K}$. Conversely, suppose $f \in \Gamma_{2M+\frac{1}{n},K}$, for all $n \geq 1$. Then for all $n \geq 1$ there exists $x_n \in K, u_n \in S_{2M+\frac{1}{n}}$ such that $f = \mathcal{G}^0\big(x_n, \int_0^\cdot u_n(s)ds\big)$. In particular, we have $\inf_{x\in K} I_x(f) \leq I_{x_n}(f) \leq M + \frac{1}{2n}$. Sending $n \to \infty$ we see that $\inf_{x\in K} I_x(f) \leq M$. Therefore $f \in \Lambda_{M,K}$ and the inclusion $\bigcap_{n\geq 1}\Gamma_{2M+\frac{1}{n},K} \subseteq \Lambda_{M,K}$ follows. This proves the first part of the theorem.

For the second part let fix $x \in \mathcal{E}_0$. Let $\{x^\epsilon > 0\} \subseteq \mathcal{E}_0$ be such that $x^\epsilon \to x$ as $\epsilon \to 0$. For notational convenience we will write $\mathcal{A}[l_2]$, $\mathcal{A}_M[l_2]$. $S_M[l_2]$ simply as $\mathcal{A}$, $\mathcal{A}_M$, S_M respectively.

- Proof of the lower bound. Note that

$$-\epsilon \log \mathbb{E}\Big[\exp\big(-\frac{1}{\epsilon}h(X^{\epsilon,x})\big)\Big] = \inf_{u\in\mathcal{A}} \mathbb{E}\Big[\frac{\epsilon}{2}\int_0^T ||u(s)||_{l_2}^2 ds + h \circ \mathcal{G}^\epsilon\big(x_\epsilon, \beta + \int_0^\cdot u(s)ds\big)\Big] \tag{3.27}$$

Fix $\delta \in (0,1)$. Then for every $\epsilon > 0$ there exists $u^\epsilon \in \mathcal{A}$ such that the right hand side of (3.27) is bounded below by

$$\mathbb{E}\Big[\frac{1}{2}\int_0^T ||u^\epsilon(s)||_{l_2}^2 ds + h\circ\mathcal{G}^\epsilon\big(x_\epsilon, \sqrt{\epsilon}\beta + \int_0^\cdot u^\epsilon(s)ds\big)\Big] - \delta. \tag{3.28}$$

Using the fact that h is bounded we can assume without loss of generality (we refer the reader to Theorem 4.4 of [8] where a similar argument is used) that

$$\sup_{\epsilon>0}\int_0^T \sum_{i=1}^\infty (u_i^\epsilon)^2(s)ds \leq N, \text{ a.s.}. \tag{3.29}$$

In order to prove the lower bound it suffices to show that

$$\liminf_{\epsilon\to 0}\mathbb{E}\Big[\frac{1}{2}\int_0^T ||u^\epsilon(s)||_{l_2}^2 ds + h\circ\mathcal{G}^\epsilon\big(x_\epsilon, \sqrt{\epsilon}\beta + \int_0^\cdot u^\epsilon(s)ds\big)\Big] \geq \inf_{f\in\mathcal{E}}\{I_x(f)+h(f)\}. \tag{3.30}$$

Pick a subsequence (relabeled by ϵ) along which u^ϵ converges in distribution to some $u\in\mathcal{A}_N$ as S_N–valued random elements. We now infer from the Assumption (3.25) that:

$$\begin{aligned}
&\liminf_{\epsilon\to 0}\mathbb{E}\Big[\frac{1}{2}\int_0^T ||u^\epsilon(s)||_{l_2}^2 ds + h\circ\mathcal{G}^\epsilon\big(x_\epsilon, \sqrt{\epsilon}\beta + \int_0^\cdot u^\epsilon(s)ds\big)\Big]\\
&\quad\geq \mathbb{E}\Big[\frac{1}{2}\int_0^T ||u(s)||_{l_2}^2 ds + h\circ\mathcal{G}^0\big(x_\epsilon, \beta + \int_0^\cdot u(s)ds\big)\Big]\\
&\quad\geq \inf_{\{(f,u)\in\mathcal{E}\times L^2([0,T]:l_2): f=\mathcal{G}^0(x,\int_0^\cdot u(s)ds)\}}\Big\{\frac{1}{2}\int_0^T ||u(s)||_{l_2}^2 ds + h(f)\Big\}\\
&\quad\geq \inf_{f\in\mathcal{E}}\{I_x(f)+h(f)\}.
\end{aligned}$$

- Proof of the upper bound. We need to show that

$$\limsup_{\epsilon\to 0} -\epsilon\log\mathbb{E}(-\frac{1}{\epsilon}h[X^{\epsilon,x_\epsilon}]) \leq \inf_{f\in\mathcal{E}}\{I_x(f)+h(f)\}.$$

Without loss of generality we can assume that $\inf_{f\in\mathcal{E}}\{I_x(f)+h(f)\}<\infty$. Let $\delta>0$

be arbitrary, and let $f_0 \in \mathcal{E}$ be such that

$$I_x(f_0) + h(f_0) \leq \inf_{f \in \mathcal{E}} \{I_x(f) + h(f)\} + \frac{\delta}{2}. \tag{3.31}$$

Choose $\widetilde{u} \in L^2([0,T] : l_2)$ such that:

$$\frac{1}{2}\int_0^T ||\tilde{u}(s)||_{l_2}^2 ds \leq I_x(f_0) + \frac{\delta}{2} \text{ and } f_0 = \mathcal{G}^0\big(x, \int_0^{\cdot} \tilde{u}(s)ds\big). \tag{3.32}$$

Then,

$$\begin{aligned}
&\limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}\big[-\frac{1}{\epsilon}h(X^{\epsilon,x_\epsilon})\big] = \\
&= \limsup_{\epsilon\to 0} \inf_{u\in\mathcal{A}} \mathbb{E}\Big[\frac{1}{2}\int_0^T ||u(s)||_{l_2}^2 ds + h\circ\mathcal{G}^\epsilon\big(x_\epsilon, \sqrt{\epsilon}\beta + \int_0^{\cdot} u(s)ds\big)\Big] \\
&\leq \limsup_{\epsilon\to 0} \mathbb{E}\Big[\frac{1}{2}\int_0^T ||\tilde{u}(s)||_{l_2}^2 ds + h\circ\mathcal{G}^\epsilon\big(x_\epsilon, \sqrt{\epsilon}\beta + \int_0^{\cdot} \tilde{u}(s)ds\big)\Big] \\
&= \frac{1}{2}\int_0^T ||\tilde{u}(s)||_{l_2}^2 ds + \limsup_{\epsilon\to 0} \mathbb{E}\Big[h\circ\mathcal{G}^\epsilon\big(x_\epsilon, \sqrt{\epsilon}\beta + \int_0^{\cdot} \tilde{u}(s)ds\big)\Big]
\end{aligned} \tag{3.33}$$

By Assumption (3.25) $\lim_{\epsilon\to 0} \mathbb{E}\Big[h\circ\mathcal{G}^\epsilon\big(x_\epsilon, \sqrt{\epsilon}\beta + \int_0^{\cdot} \tilde{u}(s)ds\big)\Big] = h\big(\mathcal{G}^0(x, \int_0^{\cdot} \tilde{u}(s)ds)\big) = h(f_0)$. Thus in view of (3.31) and (3.32) the expression (3.33) can be at most $\inf_{f\in\mathcal{E}}\{I(f) + h(f)\} + \delta$. Since δ is arbitrary, the proof is complete.

□

We now present the analog of Theorem 3.6.1 for the case of a Brownian sheet.

Assumption 3.6.2. *There exists a measurable map $\mathcal{G}^0 : \mathcal{E}_0 \times \mathcal{C}([0,T]\times\mathcal{O} : \mathbb{R}) \to \mathcal{E}$ such that the following hold:*

1. *For every $M < \infty$ and compact set $K \subseteq \mathcal{E}_0$, the set,*

$$\Gamma_{M,K} \doteq \Big\{\mathcal{G}^0(z, \mathcal{I}(u)) : u \in S_M,\ z \in K\Big\} \tag{3.34}$$

is a compact subset of $\mathcal{E}$, where $\mathcal{I}(u)$ is as defined above Assumption 3.5.1.

2. *Consider $M < \infty$ and a family $\{u^\epsilon, \ \epsilon > 0\} \subset \mathcal{A}_M$, such that u^ϵ converges in distribution (as S_M–valued random elements) to $u \in \mathcal{A}$. Let $\{z^\epsilon, \epsilon > 0\} \subseteq \mathcal{E}_0$ be such that $z_\epsilon \to z$ as $\epsilon \to 0$. Then as $\epsilon \to 0$,*

$$\mathcal{G}^\epsilon\big(z^\epsilon, \sqrt{\epsilon}B + \mathcal{I}(u^\epsilon)\big) \xrightarrow{d} \mathcal{G}^0\big(z, \mathcal{I}(u)\big) \tag{3.35}$$

Define for $f \in \mathcal{E}$ and $z \in \mathcal{E}_0$,

$$I_z(f) = \inf_{\{u \in L^2([0,T]\times\mathcal{O}): f = \mathcal{G}^0(z,\mathcal{I}(u))\}} \Big\{ \frac{1}{2} \int_{[0,T]\times\mathcal{O}} u^2(s,r)drds \Big\}. \tag{3.36}$$

Theorem 3.6.2. *Let $\mathcal{G}^0 : \mathcal{E}_0 \times \mathcal{C}([0,T] \times \mathcal{O} : \mathbb{R}) \to \mathcal{E}$ be a measurable map satisfying the first part of Assumption 3.6.2. Then $I_z : \mathcal{E} \to [0,\infty]$, defined by (3.36), is a rate function on $\mathcal{E}$ and the family $\{I_z(\cdot), z \in \mathcal{E}_0\}$ of rate functions has compact level sets on compacts. Furthermore, if the second part of Assumption 3.6.2 holds then the family $\{X^{z,\epsilon} \doteq \mathcal{G}^\epsilon(z, \sqrt{\epsilon}B), \epsilon > 0)\}$ satisfies the large deviation principle on $\mathcal{E}$, with rate function I_z, uniformly for z in compact subsets of $\mathcal{E}_0$.*

CHAPTER 4

Large Deviations for Stochastic Reaction–Diffusion equations

4.1 Introduction

In this chapter we will use results from Chapter 3, in particular Theorem 3.6.2, to study the small noise uniform large deviations principle for solutions of a class of stochastic partial differential equations (SPDE) that have been considered in [27]. The class includes, as a special case, the reaction–diffusion SPDEs considered in [36]. Results of this chapter are contained in [9]. The main result of the chapter is Theorem 4.2.2 which establishes the uniform Freidlin–Wentzell LDP for such SPDEs. Our proofs do not require any discretization, approximation or exponential probability estimates that are commonly used in standard approaches to the problem. As a result, we are able to relax two main conditions which have been used in prior works (see e.g. [26]) on large deviations analysis for this class of problems. These two restrictive conditions imposed in [26] are the uniform boundedness of the diffusion coefficients and a certain geometric condition, the so called "cone condition" (cf. p.320 [25]), on the underlying domain. In the current work, our only requirement on the domain is that it be a bounded open set and for the diffusion coefficient we require that it satisfy a standard linear growth condition (instead of uniform boundedness).

4.2 Setting

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with an increasing family of right–continuous, $\mathbb{P}$–complete σ–fields $\{\mathcal{F}_t\}_{0\le t\le T}$. Let $\mathcal{O} \subseteq \mathbb{R}^d$ be a bounded open set and $\{B(t,x) : (t,x) \in \mathbb{R}_+ \times \mathcal{O}\}$ be a Brownian sheet given on this filtered probability space. Consider the SPDE

$$\begin{aligned} &dX(t,r) = (L(t)X(t,r) + R\,(t,r,X(t,r)))\,drdt + \sqrt{\epsilon}F\,(t,r,X(t,r))\,B(drdt) \qquad (4.1)\\ &X(0,r) = \xi(r), \end{aligned}$$

where $X(0,r)$ is the initial condition. Here F and R are measurable maps from $[0,T] \times \mathcal{O} \times \mathbb{R}$ to $\mathbb{R}$ and $\epsilon \in (0,\infty)$. Also, $\{L(t) : t \ge 0\}$ is a family of linear, closed, densely defined operators on $\mathcal{C}(\mathcal{O})$ (space of real continuous functions on $\mathcal{O}$) that generates a two parameter strongly continuous semigroup $\{U(t,s) : 0 \le s \le t\}$ on $\mathcal{C}(\mathcal{O})$, with kernel function $G(t,s,r,q), 0 \le s < t,\ r,q \in \mathcal{O}$. We refer the reader to Appendix A for these functional analytic notions. In particular, we have for $f \in \mathcal{C}(\mathcal{O}),$ and $t \in [0,T]$

$$(U(t,s)f)(r) = \int_{\mathcal{O}} G(t,s,r,q)f(q)dq,\ r \in \mathcal{O},\ 0 \le s < t \le T.$$

For notational convenience we write $f(r) = \int_{\mathcal{O}} G(0,0,r,q)f(q)dq$ for $f \in \mathcal{C}(\mathcal{O})$. By a solution of the SPDE (4.1), we mean the following:

Definition 4.2.1. *A random field $X \equiv \{X(t,r) : t \in [0,T], r \in \mathcal{O}\}$ is called a mild solution of the stochastic partial differential equation (4.1) with initial condition ξ if $(t,r) \mapsto X(t,r)$ is continuous a.s., $X(t,r)$ is $\{\mathcal{F}_t\}$–measurable for any $t \in [0,T],\ r \in \mathcal{O}$, and if*

$$\begin{aligned} X(t,r) \;=\; & \int_{\mathcal{O}} G(t,0,r,q)\xi(q)dq + \int_0^t \int_{\mathcal{O}} G(t,s,r,q)R\,(s,q,X(s,q))\,dqds \\ & + \sqrt{\epsilon}\int_0^t \int_{\mathcal{O}} G(t,s,r,q)F\,(s,q,X(s,q))\,B(dqds)\ a.s. \qquad (4.2) \end{aligned}$$

Implicit in Definition 4.2.1 is the requirement that the integrals in (4.2) are well defined. We will shortly introduce conditions on G, F and R that ensure that for a continuous adapted random field X, all the integrals in (4.2) are meaningful. As a convention, we take $G(t,s,r,q)$ to be zero when $0 \le t \le s \le T,\ r, q \in \mathcal{O}$.

For $u \in \mathcal{A}_N^{BS}$ [which was defined in (3.17)] the controlled analogue of (4.2) is

$$\begin{aligned} Y(t,r) \;=\; & \int_{\mathcal{O}} G(t,0,r,q)\xi(q)dq + \int_0^t \int_{\mathcal{O}} G(t,s,r,q)R(s,q,Y(s,q))dqds \\ & + \sqrt{\epsilon}\int_0^t \int_{\mathcal{O}} G(t,s,r,q)F(s,q,Y(s,q))B(dqds) \\ & + \int_0^t \int_{\mathcal{O}} G(t,s,r,q)F(s,q,Y(s,q))u(s,q)dqds. \end{aligned} \tag{4.3}$$

As discussed previously, the main work in proving an LDP for (4.2) will be to prove qualitative properties (existence and uniqueness, tightness properties, and stability under perturbations) for solutions to (4.3). We begin by discussing known qualitative theory for (4.2).

For $\alpha > 0$, let $\mathbb{B}_\alpha = \{\psi \in \mathcal{C}(\mathcal{O}) : ||\psi||_\alpha < \infty\}$ be the Banach space with norm

$$||\psi||_\alpha = ||\psi||_0 + \sup_{r,q\in\mathcal{O}} \frac{|\psi(r) - \psi(q)|}{|r-q|^\alpha},$$

where $||\psi||_0 = \sup_{r\in\mathcal{O}} |\psi(r)|$. The Banach space $\mathbb{B}_\alpha([0,T]\times\mathcal{O})$ is defined similarly and for notational convenience we denote this space by $\mathbb{B}_\alpha^T$. For $\alpha = 0$ the space $\mathbb{B}_0^T$ is the space of all continuous maps from $[0,T]\times\mathcal{O}$ to $\mathbb{R}$ endowed with the sup–norm. The following will be a standing assumption for this section. In the assumption, $\bar{\alpha}$ is a fixed constant, and the large deviation principle will be proved in the topology of $\mathcal{C}([0,T] : \mathbb{B}_\alpha)$, for any fixed $\alpha \in (0, \bar{\alpha})$.

Assumption 4.2.1. *The following two conditions hold.*

1. *There exist constants $K(T) < \infty$ and $\gamma \in (d, \infty)$ such that*

(a) for all $t, s \in [0,T],\ r \in \mathcal{O}$*,*

$$\int_{\mathcal{O}} |G(t,s,r,q)|dq \leq K(T), \tag{4.4}$$

(b) for all $0 \leq s < t \leq T$ *and* $r, q \in \mathcal{O}$*,*

$$|G(t,s,r,q)| \leq K(T)(t-s)^{-\frac{d}{\gamma}}, \tag{4.5}$$

(c) if $\bar{\alpha} = \frac{\gamma-d}{2\gamma}$*, then for any* $\alpha \in (0, \bar{\alpha})$ *and for all* $r_1, r_2, q \in \mathcal{O},\ 0 \leq s < t_1 \leq t_2 \leq T$

$$\begin{aligned} & |G(t_1,s,r_1,q) - G(t_2,s,r_2,q)| \\ \leq\ & K(T)\left[(t_2-t_1)^{1-\frac{d}{\gamma}}(t_1-s)^{-1} + |r_1-r_2|^{2\alpha}(t_1-s)^{-\frac{d+2\alpha}{\gamma}}\right], \end{aligned} \tag{4.6}$$

(d) for all $x, y \in \mathbb{R}$*,* $r \in \mathcal{O}$ *and* $0 \leq t \leq T$*,*

$$|R(t,r,x) - R(t,r,y)| + |F(t,r,x) - F(t,r,y)| \leq K(T)|x-y| \tag{4.7}$$

and

$$|R(t,r,x)| + |F(t,r,x)| \leq K(T)(1+|x|). \tag{4.8}$$

2. *For any* $\alpha \in (0,\bar{\alpha})$ *and* $\xi \in \mathbb{B}_\alpha$*, the trajectory* $t \mapsto \int_{\mathcal{O}} G(t,0,\cdot,q)\xi(q)dq$ *belongs to* $\mathcal{C}([0,T] : \mathbb{B}_\alpha)$ *and the map*

$$\mathbb{B}_\alpha \ni \xi \longmapsto \left\{ t \mapsto \int_{\mathcal{O}} G(t,0,\cdot,q)\xi(q)dq \right\} \in \mathcal{C}([0,T] : \mathbb{B}_\alpha)$$

is a continuous map.

For future reference we recall that $\bar{\alpha} = \frac{\gamma-d}{2\gamma}$ and note that $\bar{\alpha} \in (0, 1/2)$.

Remark 4.2.1. *1. We refer the reader to [26] for examples of families $\{L(t)\}_{t\geq 0}$ that satisfy this assumption.*

2. Using (4.4) and (4.5) it follows that for any $0 \leq s < t \leq T$ and $r \in \mathcal{O}$

$$\int_{\mathcal{O}} |G(t,s,r,q)|^2 dq \leq K^2(T)(t-s)^{-\frac{d}{\gamma}}. \tag{4.9}$$

This in particular ensures that the stochastic integral in (4.2) is well defined.

3. Lemma 4.1(ii) of [26] shows that under Assumption 4.2.1, for any $\alpha < \bar{\alpha}$ there exists a constant $\tilde{K}(\alpha)$ such that for all $0 \leq t_1 \leq t_2 \leq T$ and all $r_1, r_2 \in \mathcal{O}$

$$\int_0^T \int_{\mathcal{O}} |G(t_1,s,r_1,q) - G(t_2,s,r_2,q)|^2 dq ds \leq \tilde{K}(\alpha)\rho\left((t_1,r_1),(t_2,r_2)\right)^{2\alpha}, \tag{4.10}$$

where ρ is the Euclidean distance in $[0,T] \times \mathcal{O} \subset \mathbb{R}^{d+1}$. This estimate will be used in the proof of Lemma 4.3.2.

The following theorem is due to Kotelenez (see Theorem 2.1 and Theorem 3.4 in [27]; see also Theorem 3.1 in [26]).

Theorem 4.2.1. *Fix $\alpha \in (0, \bar{\alpha})$. There exists a measurable function*

$$\mathcal{G}^\epsilon : \mathbb{B}_\alpha \times \mathbb{B}_0^T \to \mathcal{C}([0,T] : \mathbb{B}_\alpha)$$

such that for any filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ with a Brownian sheet B as above and $x \in \mathbb{B}_\alpha$, $X^{\epsilon,x} \doteq \mathcal{G}^\epsilon(x, \sqrt{\varepsilon}B)$ is the unique mild solution of (4.1) (with initial condition x), and satisfies $\sup_{0\leq t\leq T} \mathbb{E}||X^{\epsilon,x}(t)||_0^p < \infty$ for all $p \geq 0$.

For the rest of the section we will only consider $\alpha \in (0, \bar{\alpha})$. For $f \in \mathcal{C}([0,T] : \mathbb{B}_\alpha)$ define

$$I_x(f) \doteq \inf_u \int_{[0,T]\times\mathcal{O}} u^2(s,q) dq ds, \tag{4.11}$$

where the infimum is taken over all $u \in L^2([0,T] \times \mathcal{O})$ such that

$$\begin{aligned} f(t,r) &= \int_{\mathcal{O}} G(t,0,r,q)x(q)dq + \int_{[0,t]\times\mathcal{O}} G(t,s,r,q)R(s,q,f(s,q))dqds \\ &\quad + \int_{[0,t]\times\mathcal{O}} G(t,s,r,q)F(s,q,f(s,q))u(s,q)dqds. \end{aligned} \tag{4.12}$$

The following is the main result of this section.

Theorem 4.2.2. *Let $X^{\epsilon,x}$ be as in Theorem 4.2.1. Then I_x defined by* (4.11) *is a rate function on $\mathcal{C}([0,T]:\mathbb{B}_\alpha)$ and the family $\{I_x, x \in \mathbb{B}_\alpha\}$ of rate functions has compact level sets on compacts. Furthermore, $\{X^{\epsilon,x}\}$ satisfies the Laplace principle on $\mathcal{C}([0,T]:\mathbb{B}_\alpha)$ with the rate function I_x, uniformly for x in compact subsets of $\mathbb{B}_\alpha$.*

Remark 4.2.2.

1. *If Assumption 4.2.1 (2) is weakened to merely the requirement that for every $\xi \in \mathbb{B}_\alpha$, $t \mapsto \int_{\mathcal{O}} G(t,0,\cdot,q)\xi(q)dq$ is in $\mathcal{C}([0,T]:\mathbb{B}_\alpha)$, then the proof of Theorem 4.2.2 shows that for all $x \in \mathbb{B}_\alpha$, the large deviation principle for $\{X^{\epsilon,x}\}$ on $\mathcal{C}([0,T]:\mathbb{B}_\alpha)$ holds (but not necessarily uniformly).*

2. *The small noise LDP for a class of reaction–diffusion SPDEs, with $\mathcal{O} = [0,1]$ and a bounded diffusion coefficient, has been studied in [36]. A difference in the conditions on the kernel G in [36] is that instead of* (4.6)*, G satisfies the L^2 estimate in Remark 4.2.1 (3) with $\bar{\alpha} = 1/4$. One finds that the proof of Lemma 4.3.2, which is at the heart of the proof of Theorem 4.2.2, only uses the L^2 estimate rather than the condition* (4.6)*. Using this observation one can, in a straightforward manner, extend results of [36] to the case where the diffusion coefficient, instead of being bounded, satisfies the linear growth condition* (4.8)*.*

Since the proof of Theorem 4.2.2 relies on properties of the controlled process (4.3), the first step is to prove existence and uniqueness of solutions. This follows from a standard application of Girsanov's Theorem.

Theorem 4.2.3. *Let $\mathcal{G}^\epsilon$ be as in Theorem 4.2.1 and let $u \in \mathcal{A}_N^{BS}$ for some $N \in \mathbb{N}_0$ where $\mathcal{A}_N^{BS}$ is as defined in (3.17). For $\epsilon > 0$ and $x \in \mathbb{B}_\alpha$ define*

$$X_x^{\epsilon,u} \doteq \mathcal{G}^\epsilon\left(x, \sqrt{\varepsilon}B + \mathcal{I}(u)\right)$$

where $\mathcal{I}$ is defined in (3.18). Then $X_x^{\epsilon,u}$ is the unique solution of (4.3).

Proof. Fix $u \in \mathcal{A}_N^{BS}$. Since

$$\mathbb{E}\left(\exp\left\{-\frac{1}{\sqrt{\epsilon}}\int_{[0,T]\times\mathcal{O}} u(s,q)B(dqds) - \frac{1}{2\epsilon}\int_{[0,T]\times\mathcal{O}} u^2(s,q)dqds\right\}\right) = 1,$$

the measure $\gamma^{u,\epsilon}$ defined by

$$d\gamma^{u,\epsilon} = \exp\left\{-\frac{1}{\sqrt{\epsilon}}\int_{[0,T]\times\mathcal{O}} u(s,q)B(dqds) - \frac{1}{2\epsilon}\int_{[0,T]\times\mathcal{O}} u^2(s,q)dqds\right\} d\mathbb{P}$$

is a probability measure on $(\Omega, \mathcal{F}, \mathbb{P})$. Furthermore, $\gamma^{u,\epsilon}$ is mutually absolutely continuous with respect to $\mathbb{P}$ and by Girsanov's theorem (see Theorem 10.14 [12]) the process $\widetilde{B} = B + \epsilon^{-1/2}\mathcal{I}(u)$ on $(\Omega, \mathcal{F}, \gamma^{u,\epsilon}, \{\mathcal{F}_t\})$ is a Brownian sheet. Thus, by Theorem 4.2.1 $X_x^{\epsilon,u} = \mathcal{G}^\epsilon(x, \sqrt{\varepsilon}B + \mathcal{I}(u))$ is the unique solution of (4.2), with B there replaced by $\widetilde{B}$, on $(\Omega, \mathcal{F}, \gamma^{u,\epsilon}, \{\mathcal{F}_t\})$. However equation (4.2) with $\widetilde{B}$ is precisely same as equation (4.3), and since $\gamma^{u,\epsilon}$ and $\mathbb{P}$ are mutually absolutely continuous, we get that $X_x^{\epsilon,u}$ is the unique solution of (4.3) on $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$. This completes the proof. □

In the next subsection we will study, under the standing assumption of this section, the following two basic qualitative results regarding the processes $X_x^{\epsilon,u}$. The first is simply the controlled, zero–noise version of the theorem just stated and its proof, being very similar to the proof of Theorem 4.2.1, is omitted. The next is a standard convergence result whose proof is given in Section 4.3.

Theorem 4.2.4. *Fix $x \in \mathbb{B}_\alpha$ and $u \in L^2([0,T] \times \mathcal{O})$. Then there is a unique function f in $\mathcal{C}([0,T] : \mathbb{B}_\alpha)$ which satisfies equation (4.12).*

In analogy with the notation $X_x^{\varepsilon,u}$ for the solution of (4.3), we will denote the unique solution f given by Theorem 4.2.4 as $X_x^{0,u}$. Let $\theta : [0,1) \to [0,1)$ be a measurable map such that $\theta(r) \to \theta(0) = 0$ as $r \to 0$.

Theorem 4.2.5. *Let $M < \infty$, and suppose that $x^\epsilon \to x$ and $u^\epsilon \to u$ in distribution as $\epsilon \to 0$ with $\{u^\epsilon\} \subset \mathcal{A}_M^{BS}$. Then $X_{x^\epsilon}^{\theta(\epsilon),u^\epsilon} \to X_x^{0,u}$ in distribution.*

PROOF OF THEOREM 4.2.2. Define the map $\mathcal{G}^0 : \mathbb{B}_\alpha \times \mathbb{B}_0^T \to \mathcal{C}([0,T] : \mathbb{B}_\alpha)$ as follows. For $x \in \mathbb{B}_\alpha$ and $\phi \in \mathbb{B}_0^T$ of the form $\phi(t,x) \doteq \mathcal{I}(u)(t,x)$ for some $u \in L^2([0,T] \times \mathcal{O})$, we define $\mathcal{G}^0(x,\phi) = X_x^{0,u}$. Set $\mathcal{G}^0(x,\phi) = 0$ for all other $\phi \in \mathbb{B}_0^T$. In view of Theorem 3.6.2, it suffices to show that $(\mathcal{G}^\epsilon, \mathcal{G}^0)$ satisfy Assumption 3.6.2 with $\mathcal{E}_0$ and $\mathcal{E}$ there replaced by $\mathbb{B}_\alpha$ and $\mathcal{C}([0,T]; \mathbb{B}_\alpha)$ respectively; and for all $f \in \mathcal{E}$, the map $x \mapsto I_x(f)$ is l.s.c. The latter property and the first part of Assumption 3.6.2 is immediate on applying Theorem 4.2.4 and Theorem 4.2.5 with $\theta = 0$. The second part of Assumption 3.6.2 follows on applying Theorem 4.2.5 with $\theta(r) = r$, $r \in [0,1)$. $\square$

4.3 Qualitative Properties of Controlled Reaction–Diffusion SDEs

This section is devoted to the proof of Theorem 4.2.5. Our first result shows that L^p bounds hold for controlled SDEs, uniformly when the initial condition and controls lie in compact sets and $\varepsilon \in [0,1)$. Note in particular that $\varepsilon = 0$ is allowed.

Lemma 4.3.1. *If K is any compact subset of $\mathbb{B}_\alpha$ and $M < \infty$, then for all $p \in [1,\infty)$*

$$\sup_{u \in \mathcal{P}_2^M} \sup_{x \in K} \sup_{\epsilon \in [0,1)} \sup_{(t,r) \in [0,T] \times \mathcal{O}} \mathbb{E}|X_x^{\epsilon,u}(t,r)|^p < \infty.$$

Proof. By Doob's inequality there exists a suitable constant c_1 such that

$$\begin{aligned}\mathbb{E}|X_x^{\epsilon,u}(t,r)|^p \le c_1 \left|\int_{\mathcal{O}} G(t,0,r,q)x(q)dq\right|^p + c_1\mathbb{E}\left|\int_0^t\int_{\mathcal{O}} G(t,s,r,q)R\left(s,q,X_x^{\epsilon,u}(s,q)\right)dqds\right|^p \\ + c_1\mathbb{E}\left[\int_0^t\int_{\mathcal{O}} |G(t,s,r,q)|^2\left|F\left(s,q,X_x^{\epsilon,u}(s,q)\right)\right|^2 dqds\right]^{\frac{p}{2}} \\ + c_1\mathbb{E}\left[\int_0^t\int_{\mathcal{O}} |G(t,s,r,q)|\left|F\left(s,q,X_x^{\epsilon,u}(s,q)\right)\right||u(s,q)|dqds\right]^p.\end{aligned}$$

Using (4.8) and the Cauchy–Schwarz inequality the right hand side above can be bounded by

$$c_2\left[1+\mathbb{E}\left(\int_0^t\int_{\mathcal{O}}|G(t,s,r,q)|^2|X_x^{\epsilon,u}(s,q)|^2dqds\right)^{\frac{p}{2}}\right].$$

Hölder's inequality yields for $p>2$ that

$$\Lambda_p(t)\le c_2\left[1+\left(\int_0^T\int_{\mathcal{O}}|G(t,s,r,q)|^{2\tilde{p}}dqds\right)^{\frac{p-2}{2}}\int_0^t\Lambda_p(s)ds\right],$$

where $\Lambda_p(t)=\sup_{u\in P_2^M}\sup_{x\in K}\sup_{\epsilon\in[0,1)}\sup_{r\in\mathcal{O}}\mathbb{E}|X_x^{\epsilon,u}(t,r)|^p$ and $\tilde{p}=\frac{p}{p-2}$. Choose p_0 large enough that $(\frac{2p_0}{p_0-2}-1)(1-2\bar{\alpha})<1$. Using (4.4) and (4.5), we have for all $p\ge p_0$ that

$$\left[\int_0^T\int_{\mathcal{O}}|G(t,s,r,q)|^{2\tilde{p}}dqds\right]^{\frac{p-2}{2}}\le c_3T^{(1-(2\tilde{p}-1)(1-2\bar{\alpha}))\frac{p-2}{2}}.$$

Thus for every $p\ge p_0$ there exists a constant c_4 such that $\Lambda_p(t)\le c_4\left[1+\int_0^t\Lambda_p(s)ds\right]$. The result now follows from Gronwall's lemma. □

The following lemma will be instrumental in proving tightness and weak convergence in Banach spaces such as $\mathbb{B}_\alpha$ and $\mathbb{B}_\alpha^T$.

Lemma 4.3.2. *Let $\mathcal{A}\subseteq\mathcal{A}_N^{BS}$ be a family such that for all $p\ge 2$*

$$\sup_{f\in\mathcal{A}}\sup_{(t,r)\in[0,T]\times\mathcal{O}}\mathbb{E}|f(t,r)|^p<\infty. \tag{4.13}$$

Also, let $\mathcal{B} \subseteq \mathcal{A}_M^{BS}$ for some $M < \infty$. For $f \in \mathcal{A}$ and $u \in \mathcal{B}$ define

$$\begin{aligned}
\Psi_1(t,r) &\doteq \int_0^t \int_{\mathcal{O}} G(t,s,r,q) f(s,q) B(dqds), \\
\Psi_2(t,r) &\doteq \int_0^t \int_{\mathcal{O}} G(t,s,r,q) f(s,q) u(s,q) dqds,
\end{aligned}$$

where the dependence on f and u is not made explicit in the notation. Then for any $\alpha < \bar{\alpha}$ and $i = 1, 2$,

$$\sup_{f \in \mathcal{A}, u \in \mathcal{B}} \mathbb{E} \left\{ \sup_{\rho((t,r),(s,q))<1} \frac{|\Psi_i(t,r) - \Psi_i(s,q)|}{\rho\left((t,r),(s,q)\right)^{\alpha}} \right\} < \infty.$$

Proof. We will prove the result for $i = 1$; the proof for $i = 2$ is identical (except an additional application of the Cauchy–Schwarz inequality) and thus it is omitted. Henceforth we write, for simplicity, Ψ_1 as Ψ. We will apply Theorem 6 of [23], according to which it suffices to show that for all $0 \le t_1 < t_2 \le T$, $r_1, r_2 \in \mathcal{O}$,

$$\sup_{f \in \mathcal{A}, u \in \mathcal{B}} \mathbb{E} \left| \Psi(t_2, r_2) - \Psi(t_1, r_1) \right|^p \le c_p \left(\hat{\omega} \left(\rho \left((t_1, r_1), (t_2, r_2) \right) \right) \right)^p, \tag{4.14}$$

for a suitable constant c_p; a $p > 2$; and a function $\hat{\omega} : [0, \infty) \to [0, \infty)$ satisfying

$$\int_0^1 \frac{\hat{\omega}(u)}{u^{1+\alpha+(d+1)/p}} du < \infty.$$

We will show that (4.14) holds with $\hat{\omega}(u) = u^{\alpha_0}$ for some $\alpha_0 \in (\alpha, \bar{\alpha})$ and all p sufficiently large. This will establish the result.

Fix $\alpha_0, \tilde{\alpha}$ such that $\alpha < \alpha_0 < \tilde{\alpha} < \bar{\alpha}$ and let $t_1 < t_2$, $r_1, r_2 \in \mathcal{O}$ and $p > 2$. We will need p to be sufficiently large and the choice of p will be fixed in the course of the proof. By Doob's inequality there exists a constant c_1 such that:

$$\mathbb{E}\left|\Psi(t_2,r_2)-\Psi(t_1,r_1)\right|^p \leq c_1\mathbb{E}\left[\int_0^T\int_{\mathcal{O}}\left|G(t_2,s,r_2,q)-G(t_1,s,r_1,q)\right|^2|f(s,q)|^2 dqds\right]^{\frac{p}{2}}. \tag{4.15}$$

Let $\tilde{p}=p/(p-2)$ and $\delta=4/p$. Note that $(2-\delta)\tilde{p}=\delta p/2=2$. Hölder's inequality, (4.9) and (4.13) give that the right hand side of (4.15) is bounded by

$$\begin{aligned} & c_1\left[\int_0^T\int_{\mathcal{O}}\left|G(t_2,s,r_2,q)-G(t_1,s,r_1,q)\right|^{(2-\delta)\tilde{p}} dqds\right]^{\frac{p-2}{2}} \\ & \quad\times\left[\int_0^T\int_{\mathcal{O}}\left|G(t_2,s,r_2,q)-G(t_1,s,r_1,q)\right|^{\delta p/2}\mathbb{E}|f(s,q)|^p dqds\right] \\ & \quad\leq c_2\left[\int_0^T\int_{\mathcal{O}}\left|G(t_2,s,r_2,q)-G(t_1,s,r_1,q)\right|^2 dqds\right]^{\frac{p-2}{2}} \end{aligned} \tag{4.16}$$

for a suitable constant c_2 that is independent of f. From Remark 4.2.1(3), the expression in (4.16) can be bounded (for p large enough) by

$$c_3\rho\left((t_1,r_1),(t_2,r_2)\right)^{\tilde{\alpha}(p-2)} \leq c_4\rho\left((t_1,r_1),(t_2,r_2)\right)^{\alpha_0 p}.$$

The result follows. □

The next result will be used to prove the stochastic integral in (4.3) converges to 0 in $\mathcal{C}([0,T]\times\mathcal{O})\equiv\mathcal{C}([0,T]\times\mathcal{O}:\mathbb{R})$, which will be strengthened shortly.

Lemma 4.3.3. *Let $\mathcal{A}$ and Ψ_1 be as in Lemma 4.3.2 and let $Z^{\epsilon}_f \doteq \sqrt{\epsilon}\,\Psi_1$. Then for every sequence $\{f^{\varepsilon}\}\subset\mathcal{A}$, $Z^{\epsilon}_{f^{\varepsilon}}\xrightarrow{\mathbb{P}}0$ in $\mathcal{C}([0,T]\times\mathcal{O})$, as $\varepsilon\to 0$.*

Proof. Arguments similar to those lead to (4.16) along with (4.4), (4.5) and (4.13) yield that $\sup_{f\in\mathcal{A}}\mathbb{E}|\Psi_1(t,r)|^2<\infty$. This shows that for each $(t,r)\in[0,T]\times\mathcal{O}$, $Z^{\epsilon}_{f^{\varepsilon}}(t,r)\xrightarrow{\mathbb{P}}0$ (in fact in L^2). Defining

$$\omega(x,\delta)\doteq\sup\left\{|x(t,r)-x(t',r')| : \rho\left((t,r),(t',r')\right)\leq\delta\right\}$$

for $x \in \mathcal{C}([0,T] \times \mathcal{O})$ and $\delta \in (0,1)$, we see that $\omega(Z^{\epsilon}_{f^{\varepsilon}}, \delta) = \sqrt{\epsilon}\delta^{\alpha} M^{\epsilon}_{f^{\varepsilon}}$ where $M^{\epsilon}_{f} \doteq \sup_{\rho((t,r),(s,q))<1} \frac{|\Psi_1(t,r)-\Psi_1(s,q)|}{\rho((t,r),(s,q))^{\alpha}}$. Therefore from Lemma 4.3.2

$$\lim_{\delta\to 0}\lim_{\epsilon\to 0} \mathbb{E}\omega(Z^{\epsilon}_{f^{\varepsilon}}, \delta) = 0.$$

The result now follows from Theorem 14.5 of [24]. □

We now establish the main convergence result.

PROOF OF THEOREM 4.2.5. Given $x \in K, u \in \mathcal{A}^{BS}_{M}, \epsilon \in [0,1)$, define

$$\begin{aligned}
Z^{\epsilon,u}_{1,x}(t,r) &= \int_{\mathcal{O}} G(t,0,r,q)x(q)dq \\
Z^{\epsilon,u}_{2,x}(t,r) &= \int_0^t \int_{\mathcal{O}} G(t,s,r,q)R(s,q,X^{\theta(\epsilon),u}_x(s,q))dqds \\
Z^{\epsilon,u}_{3,x}(t,r) &= \sqrt{\theta(\epsilon)} \int_0^t \int_{\mathcal{O}} G(t,s,r,q)F(s,q,X^{\theta(\epsilon),u}_x(s,q))B(dqds) \\
Z^{\epsilon,u}_{4,x}(t,r) &= \int_0^t \int_{\mathcal{O}} G(t,s,r,q)F(s,q,X^{\theta(\epsilon),u}_x(s,q))u(s,q)dqds.
\end{aligned}$$

We first show that each $Z^{\epsilon,u^{\epsilon}}_{i,x^{\epsilon}}$ is tight in $\mathcal{C}([0,T] : \mathbb{B}_{\alpha})$, for $i = 1,2,3,4$. For $i = 1$ this follows from part 2 of Assumption 4.2.1. Recalling that $\mathbb{B}^T_{\alpha^*}$ is compactly embedded in $\mathbb{B}^T_{\alpha}$ for $\bar{\alpha} > \alpha^* > \alpha$, it suffices to show that for some $\alpha^* \in (\alpha, \bar{\alpha})$

$$\sup_{\epsilon\in(0,1)} \mathbb{P}\left[||Z^{\epsilon,u^{\epsilon}}_{i,x^{\epsilon}}||_{\mathbb{B}^T_{\alpha^*}} > K\right] \to 0 \text{ as } K \to \infty \text{ for } i = 2,3,4. \tag{4.17}$$

For $i = 2,4$, (4.17) is an immediate consequence of

$$\sup_{\epsilon\in(0,1)} \mathbb{E}||Z^{\epsilon,u^{\epsilon}}_{i,x^{\epsilon}}||_{\mathbb{B}^T_{\alpha^*}} < \infty,$$

as follows from Lemma 4.3.2, the linear growth condition (4.8) and Lemma 4.3.1. For

$i = 3$, in view of Lemma 4.3.3, it suffices to establish

$$\sup_{\epsilon \in (0,1)} \mathbb{E}[Z^{\epsilon,u^\epsilon}_{3,x^\epsilon}]_{\mathbb{B}^T_{\alpha^*}} < \infty,$$

where for $z \in \mathbb{B}^T_\alpha$, $[z]_{\mathbb{B}^T_\alpha} = ||z||_{\mathbb{B}^T_\alpha} - ||z||_0$. Once more, this follows as an immediate consequence of Lemma 4.3.2, the linear growth condition (4.8) and Lemma 4.3.1.

Having shown tightness of $Z^{\epsilon,u^\epsilon}_{i,x^\epsilon}$ for $i = 1, 2, 3, 4$, we can extract a subsequence along which each of these processes and $X^{\epsilon,u^\epsilon}_{x^\epsilon}$ converges in distribution in $\mathcal{C}([0,T] : \mathbb{B}_\alpha)$. Let $Z^{0,u}_{i,x}$ and $X^{0,u}_x$ denote the respective limits. We will show that

$$\begin{aligned}
Z^{0,u}_{1,x}(t,r) &= \int_{\mathcal{O}} G(t,0,r,q)x(q)dq \\
Z^{0,u}_{2,x}(t,r) &= \int_0^t \int_{\mathcal{O}} G(t,s,r,q)R(s,q,X^{0,u}_x(s,q))dqds \\
Z^{0,u}_{3,x}(t,r) &= 0 \\
Z^{0,u}_{4,x}(t,r) &= \int_0^t \int_{\mathcal{O}} G(t,s,r,q)F(s,q,X^{0,u}_x(s,q))u(s,q)dqds.
\end{aligned} \tag{4.18}$$

The uniqueness result Theorem 4.2.4 will then complete the proof. Convergence for $i = 1$ follows from part 2 of Assumption 4.2.1. The case $i = 3$ follows from Lemma 4.3.3, Lemma 4.3.1 and the linear growth condition. To deal with the cases $i = 2, 4$ we invoke the Skorokhod Representation Theorem [29], which allows us to assume with probability one convergence for the purposes of identifying the limits. We give the proof of convergence only for the harder case $i = 4$. Denote the right side of (4.18) by $\hat{Z}^{0,u}_{4,x}(t,r)$. Then

$$\begin{aligned}
&\left| Z^{\epsilon,u^\epsilon}_{4,x^\epsilon}(t,r) - \hat{Z}^{0,u}_{4,x}(t,r) \right| \\
&\quad \leq \int_0^t \int_{\mathcal{O}} |G(t,s,r,q)| \left| F(s,q,X^{\epsilon,u^\epsilon}_{x^\epsilon}(s,q)) - F(s,q,X^{0,u}_x(s,q)) \right| |u^\epsilon(s,q)| \, dqds \\
&\qquad + \left| \int_0^t \int_{\mathcal{O}} G(t,s,r,q)F(s,q,X^{0,u}_x(s,q)) \left(u^\epsilon(s,q) - u(s,q)\right) dqds \right|.
\end{aligned} \tag{4.19}$$

By the Cauchy–Schwarz inequality, equation (4.9) and the uniform Lipschitz property of F we see that, for a suitable constant $c \in (0, \infty)$, the first term on the right side of (4.19) can be bounded above by

$$\sqrt{M}\left[\int_0^t \int_{\mathcal{O}} |G(t,s,r,q)|^2 \left|F(s,q,X_{x^\epsilon}^{\epsilon,u^\epsilon}(s,q)) - F(s,q,X_x^{0,u}(s,q))\right|^2 dqds\right]^{1/2}$$
$$\leq \; c\left(\sup_{(s,q)\in[0,T]\times\mathcal{O}} \left|X_{x^\epsilon}^{\epsilon,u^\epsilon}(s,q) - X_x^{0,u}(s,q)\right|\right),$$

and thus converges to 0 as $\varepsilon \to 0$. The second term in (4.19) converges to 0 as well, since $u^\varepsilon \to u$ and

$$\int_0^t \int_{\mathcal{O}} \left(G(t,s,r,q)F(s,q,X_x^{0,u}(s,q))\right)^2 dqds < \infty.$$

By uniqueness of limits and noting that $\hat{Z}_{4,x}^{0,u}$ is a continuous random field, we see that $Z_{4,x}^{0,u} = \hat{Z}_{4,x}^{0,u}$ and the proof is complete. $\square$

4.4 Other Infinite Dimensional Models

The key ingredients in the proof of the LDP for the solution of the infinite dimensional SDE are the qualitative properties in Theorems 4.2.4 and 4.2.5 of the controlled SDE (4.3). Once these properties are verified, the LDP follows as an immediate consequence of Theorem 3.6.2. Furthermore, one finds that the estimates needed for the proof of Theorems 4.2.4 and 4.2.5 are essentially the same as those needed for establishing unique solvability of (4.1). This is a common theme that appears in all proofs of LDPs, for small noise stochastic dynamical systems, that are based on variational representations such as in Section 2.6. Indeed, one can argue that the variational representation approach makes the small noise large deviation analysis a transparent and a largely straightforward exercise, once one has the estimates for the unique solvability of the stochastic equation. This statement has been affirmed by several recent works on Freidlin–Wentzell large deviations for infinite dimensional SDEs that are based on the variational representation approach (specifically Theorem 2.6.1), and carry out the verification of statements analogous to

Theorems 4.2.4 and 4.2.5. Some of these works are summarized below.

4.4.1 SDEs driven by infinitely many Brownian motions

Ren and Zhang [34] consider a SDE driven by infinitely many Brownian motions with non–Lipschitz diffusion coefficients. Prior results on strong existence and uniqueness of the solutions to the SDE yield continuous (in time and initial condition) random field solutions. The authors prove a small noise LDP in the space $\mathcal{C}([0,T] \times \mathbb{R}^d)$. The proof relies on the representation formula for an infinite sequence of real Brownian motions $\{\beta_i\}$ given in Corollary 2.6.1 and the general Laplace principle of the form in Theorem 3.4.1. Non–Lipschitz coefficients make the standard discretization and approximation approach intractable for this example. The authors verify the analogues of Theorems 4.2.4 and 4.2.5 in Theorems 3.1, Lemma 3.4 and Lemma 3.11 of the cited paper. In the final section of the paper, Schilder's theorem for Brownian motion on the group of homeomorphisms of the circle is obtained. The proof here is also by verification of steps analogous to Theorems 4.2.4 and 4.2.5 regarding solvability and convergence in the space of homeomorphisms. Once more, exponential probability estimates with the natural metric on the space of homeomorphisms, needed in the standard proofs of the LDP, do not appear to be straightforward. Using similar ideas based on representations for infinite dimensional Brownian motions, a LDP for flows of homeomorphisms, extending results of the final section of [34] to multi–dimensional SDEs with non–Lipschitz coefficients, has been studied in [33].

4.4.2 Stochastic PDE with varying boundary conditions

Wang and Duan [39] study stochastic parabolic PDEs with rapidly varying random dynamical boundary conditions. The formulation of the SPDE as an abstract stochastic evolution equation in an appropriate Hilbert space leads to a non–Lipschitz nonlinearity with polynomial growth. Deviations of the solution from the limiting effective system (as the parameter governing the rapid component approaches its limit) are studied by establishing a large deviation principle. The proof of the LDP uses the variational representation for functionals of a Hilbert space valued Wiener process as in Theorem 2.6.1 and the general Laplace principle given in Theorem 3.3.1. Once more, the hardest part in the analysis is establishing the wellposedness (i.e., existence, uniqueness) of the stochastic evolution equation. Once estimates for existence/uniqueness are available, the proof of the LDP becomes a straightforward verification of Assumption 3.3.1.

4.4.3 Stochastic Navier–Stokes equation

Sritharan and Sundar [37] study small noise large deviations for a two dimensional Navier–Stokes equation in an (possibly) unbounded domain and with multiplicative noise. The equation can be written as an abstract stochastic evolution equation in an appropriate function space. The solution lies in the Polish space $\mathcal{C}([0,T]:H)\bigcap L^2([0,T]:V)$ for some Hilbert spaces H and V and can be expressed as $\mathcal{G}^\varepsilon(\sqrt{\varepsilon}W)$ for a H valued Wiener process W. Authors prove existence and uniqueness of solutions and then apply Theorem 3.3.1 by verifying Assumption 3.3.1 for their model.

CHAPTER 5

Large deviations for stochastic flows of diffeomorphisms

5.1 Introduction

In this chapter we consider a second application of the general large deviation results from Chapter 3. Using Theorem 3.4.1, we will establish a large deviation principle for a general class of stochastic flows of diffeomorphisms, driven by an infinite dimensional Brownian motion, in the small noise limit. This result is then applied in Section 5.5 to a Bayesian formulation of an image matching problem, and an approximate maximum likelihood property is shown for the solution of an optimization problem involving the large deviations rate function.

5.2 Preliminaries

We refer list below some standard notation that will be used in this chapter.

- Let $\circ$ denote the composition of maps and let id denote the identity map on $\mathbb{R}^d$.

- Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_d)$ be a multi index of non–negative integers and $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_d$. For a $|\alpha|$–times differentiable function $f : \mathbb{R}^d \to \mathbb{R}$, set $\partial^\alpha f \doteq \partial_x^\alpha f = \frac{\partial^{|\alpha|} f}{(\partial x_1)^{\alpha_1} \cdots (\partial x_d)^{\alpha_d}}$. For such an f, we write $\frac{\partial f(x)}{\partial x_i}$ as $\partial_i f$. If $f \equiv (f_1, f_2, \cdots, f_d)'$ is a $|\alpha|$–times differentiable function from $\mathbb{R}^d$ to $\mathbb{R}^d$, we write $\partial^\alpha f \doteq (\partial^\alpha f_1, \partial^\alpha f_2, \cdots, \partial^\alpha f_d)'$. $K \subset\subset \mathbb{R}^d$ will denote the statement that K is a compact subset of $\mathbb{R}^d$.

- For $m \geq 0$ denote by $\mathbb{C}^m$ the space of m–times continuously differentiable functions from $\mathbb{R}^d$ to $\mathbb{R}$, which endowed with seminorms $||f||_{m;K} = \sum_{0\leq|\alpha|\leq m} \sup_{x\in K} |\partial^\alpha f(x)|$, $K \subset\subset \mathbb{R}^d$, is a Fréchet space, where $\partial^0 f = f$. Also, for $0 < \delta \leq 1$, let

$$\mathbb{C}^{m,\delta} \doteq \big\{f \in \mathbb{C}^m : ||f||_{m,\delta;K} < \infty \text{ for any } K \subset\subset \mathbb{R}^d\big\},$$

  where

$$||f||_{m,\delta;K} = ||f||_{m;K} + \sum_{|\alpha|=m} \sup_{x,y\in K; x\neq y} \frac{|\partial^\alpha f(x) - \partial^\alpha f(y)|}{|x-y|^\delta}.$$

  The seminorms $\{||\cdot||_{m,\delta;K}; K \subset\subset \mathbb{R}^d\}$ make $\mathbb{C}^{m,\delta}$ a Fréchet space.

- For $m \geq 0$ denote by $\widetilde{\mathbb{C}}^m$ the space of functions $g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ such that $g(x,y)$, $x, y \in \mathbb{R}^d$ is m–times continuously differentiable with respect to both x and y. Endowed with the seminorms

$$||g||^\sim_{m;K} = \sum_{0\leq|\alpha|\leq m} \sup_{x,y\in K} |\partial^\alpha_x \partial^\alpha_y g(x,y)|,$$

  where $K \subset\subset \mathbb{R}^d$, $\widetilde{\mathbb{C}}^m$ is a Fréchet space. Also, for $0 < \delta \leq 1$, let $\widetilde{\mathbb{C}}^{m,\delta} \doteq \big\{g \in \widetilde{\mathbb{C}}^m; ||g||^\sim_{m,\delta;K} < \infty,\ K \subset\subset \mathbb{R}^d\big\}$, where

$$||g||^\sim_{m,\delta;K} = ||g||^\sim_{m;K} + \sum_{|\alpha|=m} \sup_{\substack{x\neq x', y\neq y' \\ x,y,x',y'\in K}} \frac{|\Delta_{x,x'} g(y) - \Delta_{x,x'} g(y')|}{|x-x'|^\delta |y-y'|^\delta},$$

  where $\Delta_{x,x'} g(y) \doteq \hat{\partial}^\alpha_{x,y} g(x,y) - \hat{\partial}^\alpha_{x',y} g(x',y)$, $\hat{\partial}^\alpha_{x,y} g(x,y) \doteq \partial^\alpha_x \partial^\alpha_y g(x,y)$. The seminorms $\{||\cdot||^\sim_{m,\delta;K}, K \subset\subset \mathbb{R}^d\}$ make $\widetilde{\mathbb{C}}^{m,\delta}$ a Fréchet space.

- We will write $||f||_{m;\mathbb{R}^d}$ as $||f||_m$. The norms $||\cdot||_{m,\delta}, ||\cdot||^\sim_m, ||\cdot||^\sim_{m,\delta}$ are to be interpreted in a similar manner.

- Let $\mathbb{C}^m(\mathbb{R}^d) \doteq \big\{f = (f_1, f_2, \cdots, f_d)' : f_i \in \mathbb{C}^m, i = 1, 2, \cdots, d\big\}$ and $||f||_m = \sum_{i=1}^d ||f_i||_m$. The spaces $\mathbb{C}^{m,\delta}(\mathbb{R}^d), \widetilde{\mathbb{C}}^m(\mathbb{R}^{d\times d})$, $\widetilde{\mathbb{C}}^{m,\delta}(\mathbb{R}^{d\times d})$ and their corresponding norms are defined similarly. In particular, note that $h \in \widetilde{\mathbb{C}}^{m,\delta}(\mathbb{R}^{d\times d})$ is map from $\mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^{d\times d}$.

- Let $\mathbb{C}_T^{m,\delta}(\mathbb{R}^d)$ and $\widetilde{\mathbb{C}}_T^{m,\delta}(\mathbb{R}^{d\times d})$ be classes of measurable functions $b : [0,T] \to \mathbb{C}^{m,\delta}(\mathbb{R}^d)$ and $a : [0,T] \to \widetilde{\mathbb{C}}^{m,\delta}(\mathbb{R}^{d\times d})$ respectively such that

$$||b||_{T,m,\delta} \doteq \sup_{0\le t\le T} ||b(t)||_{m,\delta} < \infty \text{ and } ||a||^{\sim}_{T,m,\delta} \doteq \sup_{0\le t\le T} ||a(t)||^{\sim}_{m,\delta} < \infty.$$

Definition 5.2.1. (*Stochastic flows of homeomorphisms/diffeomorphisms*) *A collection* $\{\phi_{s,t}(x) : 0 \le s \le t \le T, x \in \mathbb{R}^d\}$ *of* $\mathbb{R}^d$*–valued random variables on some filtered probability space* $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ *is called a forward stochastic flow of homeomorphisms, if there exists* $N \in \mathcal{F}$*, with* $\mathbb{P}(N) = 0$*, such that for any* $\omega \in N^c$*:*

1. $(s,t,x) \mapsto \phi_{s,t}(x,\omega)$ *is a continuous map,*
2. $\phi_{s,u}(\omega) = \phi_{t,u}(\omega) \circ \phi_{s,t}(\omega)$ *holds for all* s,t,u*,* $0 \le s \le t \le u \le T$*,*
3. $\phi_{s,s}(\omega) = id$ *for all* s*,* $0 \le s \le T$*,*
4. *the map* $\phi_{s,t}(\omega) : \mathbb{R}^d \to \mathbb{R}^d$ *is an onto homeomorphism for all* $s,t, 0 \le s \le t \le T$*.*

If in addition $\phi_{s,t}(x,\omega)$ *is* k*–times differentiable with respect to* x *for all* $s \le t$ *and the derivatives are continuous in* (s,t,x)*, it is called a stochastic flow of* $\mathbb{C}^k-$*diffeomorphisms.*

We now introduce a Brownian motion with a spatial parameter, with local characteristics (a,b). Throughout the dissertation we will assume that $(a,b) \in \widetilde{\mathbb{C}}_T^{k,\delta}(\mathbb{R}^{d\times d}) \times \mathbb{C}_T^{k,\delta}(\mathbb{R}^d)$, for some $k \in \mathbb{N}$ and $\delta \in (0,1]$. Fix ν such that $0 < \nu < \delta$.

Definition 5.2.2. ($\mathbb{C}^{k,\nu}-$*Brownian motion*) *A continuous stochastic process* $\{F(t)\}_{t\ge 0}$ *on some filtered probability space* $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ *with values in* $\mathbb{C}^{k,\nu}(\mathbb{R}^d)$ *is said to be a* $\mathbb{C}^{k,\nu}-$*Brownian motion with local characteristics* (a,b)*, if* $F(0)$*,* $F(t_{i+1}) - F(t_i)$*,* $i = 0,1,\ldots,n-1$*, are independent* $\mathbb{C}^{k,\nu}(\mathbb{R}^d)-$*valued random variables whenever* $0 \le t_0 < t_1 <$

$\cdots < t_n \leq T$, and if for each $x \in \mathbb{R}^d$, $M(x,t) \doteq F(x,t) - \int_0^t b(x,r)dr$ is a continuous (d–dimensional) martingale such that $<< M(x,\cdot), M(y,\cdot) >>_t = \int_0^t a(x,y,r)dr$ for all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$.

The existence of a $\mathbb{C}^{k,\nu}-$Brownian motion with local characteristics (a,b) follows from [28] (see, e.g., Theorem 3.1.2 and Exercise 3.2.10). Indeed, for any $\gamma < \delta$ one can represent F as in (1.5), where $f_i : \mathbb{R}^d \times [0,T] \to \mathbb{R}^d$ are such that for each $t \in [0,T]$, $f_i(\cdot,t) \in \mathbb{C}^{k,\gamma}(\mathbb{R}^d)$,

$$a(x,y,t) = \sum_{i=1}^{\infty} f_i(x,t)f_i'(y,t), \text{ a.e. } t,$$

and

$$\int_0^T \sum_{i=1}^{\infty} |f_i(x,r)|^2 dr \leq T||a||_{T,k,\delta}^{\sim} < \infty.$$

In particular, note that if F is a $\mathbb{C}^{k,\nu}-$valued Brownian motion, its finite dimensional restriction $(F(x_1,\cdot), F(x_2,\cdot), \ldots, F(x_n,\cdot))'$ is an $nd-$dimensional Brownian motion (with suitable mean and covariance) for any $(x_1,\ldots,x_n) \in \mathbb{R}^{nd}$. If F is as defined by (1.5) and $\{\phi_t\}_{0\leq t\leq 1}$ is a continuous $\mathbb{R}^d-$valued $\{\mathcal{F}_t\}-$adapted stochastic process, the stochastic integral $\int_0^t F(\phi_r, dr)$ is a well–defined $d-$dimensional continuous $\{\mathcal{F}_t\}-$adapted stochastic process (see Chapter 3, Section 2, pp. 71–86 of [28]).

Definition 5.2.3. *Let F be as in Definition 5.2.2. Then for each $s \in [0,T]$ and $x \in \mathbb{R}^d$, there is a unique continuous $\{\mathcal{F}_t\}-$adapted, $\mathbb{R}^d-$valued process $\phi_{s,t}(x)$, $s \leq t \leq T$ satisfying $\phi_{s,t}(x) = x + \int_s^t F(\phi_{s,r}(x), dr)$, $t \in [s,T]$. This stochastic process is called the solution of Itô's stochastic differential equation based on the Brownian motion F.*

From Theorem 4.6.5 [28] it follows that $\{\phi_{s,t}\}_{0\leq s\leq t\leq T}$ as introduced in Definition 5.2.3 has a modification that is a forward stochastic flow of $\mathbb{C}^k$–diffeomorphisms.

5.3 Large deviation principle

Given $\varepsilon > 0$, let F^ε be a $\mathbb{C}^{k,\nu}-$ Brownian motion on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$, with local characteristics $(\varepsilon a, b)$, where (k,ν) and (a,b) are as in Section

5.2. Without loss of generality we assume that F^ε is represented as

$$F^\varepsilon(x,t) \doteq \int_0^t b(x,r)dr + \sqrt{\varepsilon}\sum_{l=1}^\infty \int_0^t f_l(x,r)d\beta_l(r),\ (x,t) \in \mathbb{R}^d \times [0,T], \tag{5.1}$$

where $(\beta_l, f_l)_{l\geq 1}$ are as in Section 5.2. Note in particular that

$$F^\varepsilon(x,t) - \int_0^t b(x,r)dr = \sqrt{\varepsilon}M(x,t).$$

With an abuse of notation, when $\varepsilon = \varepsilon_n$ we write F^ε as F^n. Observe that for all $t \in [0,T]$, $<< M(x,\cdot), \beta_l(\cdot) >>_t = \int_0^t f_l(x,r)dr$, a.s. Let $\phi^\varepsilon \equiv \{\phi^\varepsilon_{s,t}(x), x \in \mathbb{R}^d, 0 \leq s \leq t \leq T\}$ be the forward stochastic flow of $\mathbb{C}^k-$diffeomorphisms based on F^ε. With another abuse of notation,we write $\phi^\varepsilon_{0,t}$ as ϕ^ε_t and $\phi^\varepsilon = \{\phi^\varepsilon_t(x), 0 \leq t \leq T, x \in \mathbb{R}^d\}$.

The goal of this work is to show that the family $(\phi^\varepsilon, F^\varepsilon)_{\varepsilon>0}$ satisfies a LDP on a suitable function space, as $\varepsilon \to 0$. For $m \in \mathbb{N}$, let G^m be the group of $\mathbb{C}^m-$diffeomorphisms on $\mathbb{R}^d$. G^m is endowed with the metric

$$d_m(\phi,\psi) = \lambda_m(\phi,\psi) + \lambda_m(\phi^{-1},\psi^{-1}), \tag{5.2}$$

where

$$\lambda_m(\phi,\psi) = \sum_{|\alpha|\leq m} \rho(\partial^\alpha\phi, \partial^\alpha\psi), \tag{5.3}$$
$$\rho(\phi,\psi) = \sum_{N=1}^\infty \frac{1}{2^N}\frac{\sup_{|x|\leq N}|\phi(x)-\psi(x)|}{1+\sup_{|x|\leq N}|\phi(x)-\psi(x)|}.$$

Under this metric G^m is a Polish space. Let $\hat{W}_m \doteq \mathcal{C}\big([0,T] : G^m\big)$ be the set of all continuous maps from $[0,T]$ to G^m and $W_m \doteq \mathcal{C}\big([0,T] : \mathbb{C}^m(\mathbb{R}^d)\big)$ be the set of all continuous maps from $[0,T]$ to $\mathbb{C}^m(\mathbb{R}^d)$. The space $\hat{W}_m$ endowed with the metric $\hat{d}_m(\phi,\psi) = \sup_{0\leq t\leq T} d_m(\phi(t),\psi(t))$ and the space W_m with the metric $\bar{d}_m(\phi,\psi) = \sup_{0\leq t\leq T}\lambda_m(\phi(t),\psi(t))$ are Polish spaces. Note that $(\phi^\varepsilon, F^\varepsilon)$ belongs to $\hat{W}_k \times W_k \subseteq$

$\hat{W}_{k-1} \times W_{k-1} \subseteq W_{k-1} \times W_{k-1}$. We will show that the pair $(\phi^{\varepsilon}, F^{\varepsilon})_{\varepsilon>0}$ satisfies LDPs in both of the spaces $\hat{W}_{k-1} \times W_{k-1}$ and $W_{k-1} \times W_{k-1}$, with a rate function I that is introduced below.

Let $u \equiv \{u_l\}_{l=1}^{\infty} \in \bigcup_{N\geq 1} \mathcal{A}_N[l_2]$. Given any such control, we want to construct a corresponding controlled flow in the form of a perturbed analogue of (5.1). Observe that $Z_t \doteq \sum_{l=1}^{\infty} \int_0^t u_l(s) d\beta_l(s)$ is a continuous square integrable martingale. For any $\gamma < \delta$ one can find $b_u : \mathbb{R}^d \times [0,T] \times \Omega \to \mathbb{R}^d$ such that $b_u(t,\omega) \in \mathbb{C}^{k,\gamma}(\mathbb{R}^d)$ for a.e. (t,ω), such that for each $x \in \mathbb{R}^d$, $b_u(x,\cdot)$ is predictable, and such that $\int_0^t b_u(x,s)ds =<< Z, M(x,\cdot) >>_t$ for each $(x,t) \in \mathbb{R}^d \times [0,T]$. In particular, for each $x \in \mathbb{R}^d$, $b_u(x,t) \doteq \sum_{l=1}^{\infty} u_l(t) f_l(x,t)$ a.e. (t,ω). Furthermore, for some $c \in (0,\infty)$,

$$||b_u(t)||_{k,\gamma}^2 \leq c||a||_{T,k,\delta}^{\sim} \sum_{l=1}^{\infty} |u_l(t)|^2, \; [dt \otimes \mathbb{P}] - \text{a.e. in } (t,\omega). \tag{5.4}$$

The proofs of these statements follow along the lines of Exercise 3.2.10 and Lemma 3.2.3 of [28]. Next, define

$$F^{0,u}(x,t) \doteq \int_0^t b_u(x,s)ds + \int_0^t b(x,s)ds. \tag{5.5}$$

It follows that $F^{0,u}(\cdot,t)$ is a $\mathbb{C}^{k,\gamma}(\mathbb{R}^d)$–valued continuous adapted stochastic process. Let $\hat{b}_u \doteq b_u + b$ and for $(t_0,x) \in [0,T] \times \mathbb{R}^d$ let $\{\phi_{t_0,t}^{0,u}(x)\}_{t_0 \leq t \leq T}$ be the unique solution of the equation

$$\phi_{t_0,t}^{0,u}(x) \doteq x + \int_{t_0}^t \hat{b}_u\big(\phi_{t_0,r}^{0,u}(x), r\big)dr, \; t \in [t_0,T]. \tag{5.6}$$

From Theorem 4.6.5 of [28] it follows that $\{\phi_{s,t}^{0,u}, 0 \leq s \leq t \leq T\}$ is a forward flow of $\mathbb{C}^k-$diffeomorphisms.

For $(\phi^0, F^0) \in \hat{W}_k \times W_k$ define

$$I(\phi^0, F^0) \doteq \inf_{u \in \mathcal{L}(\phi_0, F_0)} \frac{1}{2} \int_0^T ||u(s)||_{l_2}^2 ds, \tag{5.7}$$

where $\mathcal{L}(\phi^0, F^0) = \{u \in L^2([0,T] : l_2) \big| (\phi^0, F^0) = (\phi^{0,u}, F^{0,u})\}$. Note in particular that u in (5.7) is deterministic. If $(\phi^0, F^0) \in (W_{k-1} \times W_{k-1}) \setminus (\hat{W}_k \times W_k)$ then we set $I(\phi^0, F^0) = \infty$. We denote the restriction of I to $\hat{W}_{k-1} \times W_{k-1}$ by the same symbol. The following is the main result of the section.

Theorem 5.3.1. (*Large deviation principle*) *The family $(\phi^\varepsilon, F^\varepsilon)_{\varepsilon>0}$ satisfies a LDP in the spaces $\hat{W}_{k-1} \times W_{k-1}$ and $W_{k-1} \times W_{k-1}$ with rate function I.*

Let $\{u_n\}_{n=1}^\infty$ $\big(u_n \equiv \{u_l^n\}_{l=1}^\infty\big)$ be a sequence in $\mathcal{A}_N[l_2]$ for some fixed $N < \infty$. Let $\{\varepsilon_n\}_{n\geq 0}$ be a sequence such that $\varepsilon_n \geq 0$ for each n and $\varepsilon_n \to 0$ as $n \to \infty$. Note that we allow $\varepsilon_n = 0$ for all n. Recall that $M(x,t) = \sum_{i=1}^\infty \int_0^t f_i(x,r) d\beta_i(r), (x,t) \in \mathbb{R}^d \times [0,T]$. Define

$$\hat{F}^n(x,t) \doteq \int_0^t \hat{b}_{u_n}(x,r)dr + \sqrt{\varepsilon_n}\, M(x,t), \tag{5.8}$$

and let ϕ^n be the solution to

$$\phi_t^n(x) = x + \int_0^t \hat{b}_{u_n}\big(\phi_r^n(x), r\big)dr + \sqrt{\varepsilon_n} \int_0^t M\big(\phi_r^n(x), dr\big). \tag{5.9}$$

Clearly $\hat{F}^n \in W_k$, and from Theorem 4.6.5 of [28], equation (5.9) has a unique solution $\phi^n \in \hat{W}_k$ a.s. We next introduce some basic weak convergence definitions.

Definition 5.3.1. *Let $u \in \mathcal{A}_N[l_2]$ and $\{\phi^n\}$ be as above. Let $\hat{\mathbb{P}}_{k-1}^n, \hat{\mathbb{P}}_{k-1}^0$ be the measures induced by $(\phi^n, \hat{F}^n), (\phi^{0,u}, F^{0,u})$ respectively, on $\hat{W}_{k-1} \times W_{k-1}$. Thus for $A \in \mathcal{B}(\hat{W}_{k-1} \times W_{k-1})$,*

$$\hat{\mathbb{P}}_{k-1}^n(A) = \mathbb{P}\big((\phi^n, \hat{F}^n) \in A\big), \quad \hat{\mathbb{P}}_{k-1}^0(A) = \mathbb{P}\big((\phi^{0,u}, F^{0,u}) \in A\big).$$

The sequence $\big\{(\phi^n, \hat{F}^n)\big\}_{n\geq 1}$ is said to converge weakly as $G^{k-1}-$flows to $(\phi^{0,u}, F^{0,u})$ as $n \to \infty$ if $\hat{\mathbb{P}}_{k-1}^n$ converges weakly to $\hat{\mathbb{P}}_{k-1}^0$ as $n \to \infty$.

Definition 5.3.2. *Let $\mathbb{P}_{k-1}^n, \mathbb{P}_{k-1}^0$ be the measures induced by $(\phi^n, \hat{F}^n), (\phi^{0,u}, F^{0,u})$ respectively on $W_{k-1} \times W_{k-1}$. The sequence $\big\{(\phi^n, \hat{F}^n)\big\}_{n\geq 1}$ is said to converge weakly as $\mathbb{C}^{k-1}-$flows to $(\phi^{0,u}, F^{0,u})$ as $n \to \infty$ if $\mathbb{P}_{k-1}^n$ converges weakly to $\mathbb{P}_{k-1}^0$ as $n \to \infty$.*

As noted in the introduction, proofs of large deviations properties based on the general framework developed in Chapter 3 essentially reduce to weak convergence analysis for controlled analogues of the original process. For our problem the following theorem gives the needed result. The proof is given in the next subsection.

Theorem 5.3.2. *Let $\{u_n\}$ converge to u in distribution as an $S_N[l_2]-$valued sequence of random variables. Then the sequence $\{(\phi^n, \hat{F}^n)\}_{n\geq 1}$ converges weakly as $\mathbb{C}^{k-1}-$flows and $G^{k-1}-$flows to the pair $(\phi^{0,u}, F^{0,u})$ as $n \to \infty$.*

We will use Theorem 3.4.1 from Chapter 3. Recall the spaces $\mathcal{A}[l_2], S_N[l_2], \mathcal{A}_N[l_2]$, $S \equiv \mathcal{C}([0,T] : \mathbb{R}^\infty)$ (with the usual topology is a Polish space), introduced in Section 3.4, and $\beta \equiv \{\beta_i\}_{i=1}^\infty$ is a $S-$valued random variable.

PROOF OF THEOREM 5.3.1. We will only show that the sequence $(\phi^\varepsilon, F^\varepsilon)$ satisfies a LDP in $\hat{W}_{k-1} \times W_{k-1}$ with rate function I defined as in (5.7). The LDP in $W_{k-1} \times W_{k-1}$ follows similarly. Let $\mathcal{G}^\varepsilon : S \to \hat{W}_{k-1} \times W_{k-1}$ be a measurable map such that $\mathcal{G}^\varepsilon(\sqrt{\varepsilon}\beta) = (\phi^\varepsilon, F^\varepsilon)$ a.s., where F^ε is given by (5.1) and ϕ^ε is the associated flow based on F^ε. Define $\mathcal{G}^0 : S \to \hat{W}_{k-1} \times W_{k-1}$ by $\mathcal{G}^0\big(\int_0^\cdot u(s)ds\big) = (\phi^0, F^0)$ if $u \in L^2([0,T] : l_2)$ and with ϕ^0, F^0 as defined in (5.6) and (5.5), respectively. We set $\mathcal{G}^0(f) = 0$ for all other $f \in S$.

Fix $N < \infty$ and consider $\Gamma_N = \big\{\mathcal{G}^0(\int_0^\cdot u(s)ds), u \in S_N[l_2]\big\}$. We first show that Γ_N is a compact subset of $\hat{W}_{k-1} \times W_{k-1}$. For that it suffices to show that if $u_n, u \in S_N[l_2]$ are such that $u_n \to u$, then $\mathcal{G}^0(\int_0^\cdot u_n(s)ds) \to \mathcal{G}^0(\int_0^\cdot u(s)ds)$ in $\hat{W}_{k-1} \times W_{k-1}$. This is immediate from Theorem 5.3.2 on noting that $\mathcal{G}^0(\int_0^\cdot u_n(s)ds) = (\phi^n, \hat{F}^n)$, where $\phi^n, \hat{F}^n$ are as in (5.9) and (5.8) respectively with $\varepsilon_n = 0$; and $\mathcal{G}^0(\int_0^\cdot u(s)ds) = (\phi^{0,u}, F^{0,u})$, where $\phi^{0,u}, F^{0,u}$ are as in (5.6) and (5.5) respectively.

Next let $\{u_n\} \subset \mathcal{A}_N[l_2]$ and $\varepsilon_n \in (0,\infty)$ be such that $\varepsilon_n \to 0$ and u_n converges in distribution to some u as $n \to \infty$. In order to complete the proof, it is enough, in view of Theorem 3.4.1 and the definition of I in (5.7), to show that $\mathcal{G}^{\varepsilon_n}(\sqrt{\varepsilon_n}\beta + \int_0^\cdot u_n(s)ds) \to \mathcal{G}^0(\int_0^\cdot u(s)ds)$ in $\hat{W}_{k-1} \times W_{k-1}$, as $n \to \infty$. An application of Girsanov's theorem shows that $\mathcal{G}^{\varepsilon_n}(\sqrt{\varepsilon_n}\beta + \int_0^\cdot u_n(s)ds) = (\phi^n, \hat{F}^n)$, where $\phi^n, \hat{F}^n$ are defined as in (5.9) and (5.8) respectively. Also $\mathcal{G}^0(\int_0^{\cdot} u(s)ds) = (\phi^{0,u}, F^{0,u})$, where $\phi^{0,u}, F^{0,u}$ are the same as in (5.6), (5.5) respectively. The result now follows from Theorem 5.3.2. $\square$

5.4 Proof of Theorem 5.3.2

This section will present the proof of Theorem 5.3.2. It is worth recalling assumptions that will be in effect for this section, which are that $\{u_n\}$ is converging to u in distribution as an $S_N[l_2]-$valued sequence of random variables, and that $(a, b) \in \widetilde{\mathbb{C}}_T^{k,\delta}(\mathbb{R}^{d\times d}) \times \mathbb{C}_T^{k,\delta}(\mathbb{R}^d)$, for some $k \in \mathbb{N}$ and $\delta \in (0, 1]$.

We begin by introducing the $(m+p)$–point motion of the flow and the related notion of "convergence as diffusions". Let $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_p)$ be arbitrary fixed points in $\mathbb{R}^{d\times m}$ and $\mathbb{R}^{d\times p}$, respectively. Set

$$\phi_t^n(\mathbf{x}) = \big(\phi_t^n(x_1), \phi_t^n(x_2), \ldots, \phi_t^n(x_m)\big)$$

and

$$\hat{F}^n(\mathbf{y}, t) = \big(\hat{F}^n(y_1, t), \hat{F}^n(y_2, t), \ldots, \hat{F}^n(y_p, t)\big).$$

Then the pair $\{\phi_t^n(\mathbf{x}), \hat{F}^n(\mathbf{y}, t)\}$ is a $\mathbb{R}^{d\times m} \times \mathbb{R}^{d\times p}-$valued continuous stochastic process and is called an $(m + p)$–point motion of the flow. Let $V_m \doteq \mathcal{C}\big([0, T] : \mathbb{R}^{d\times m}\big)$ be the Fréchet space of all continuous maps from $[0, T]$ to $\mathbb{R}^{d\times m}$, with the usual semi–norms, and let $V_{m,p} = V_m \times V_p$ be the product space.

Definition 5.4.1. *Let $\mathbb{P}^n_{(\mathbf{x},\mathbf{y})}$ and $\mathbb{P}^0_{(\mathbf{x},\mathbf{y})}$ be the measures induced by $(\phi^n(\mathbf{x}), \hat{F}^n(\mathbf{y}))$ and $(\phi^{0,u}(\mathbf{x}), F^{0,u}(\mathbf{y}))$, respectively, on $V_{m,p}$. Thus for $A \in \mathcal{B}(V_{m,p})$*

$$\mathbb{P}^n_{(\mathbf{x},\mathbf{y})} = \mathbb{P}\big((\phi^n(\mathbf{x}), \hat{F}^n(\mathbf{y})) \in A\big), \quad \mathbb{P}^0_{(\mathbf{x},\mathbf{y})} = \mathbb{P}\big((\phi^{0,u}(\mathbf{x}), F^{0,u}(\mathbf{y})) \in A\big).$$

The sequence $\big\{(\phi^n, \hat{F}^n)\big\}_{n\geq 1}$ is said to converge weakly as diffusions to $(\phi^{0,u}, F^{0,u})$ as $n \to \infty$ if $\mathbb{P}^n_{(\mathbf{x},\mathbf{y})}$ converges weakly to $\mathbb{P}^0_{(\mathbf{x},\mathbf{y})}$ as $n \to \infty$ for each $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{d\times m} \times \mathbb{R}^{d\times p}$, *and* $m, p = 1, 2, \ldots$.

The following well known result (c.f., Theorem 5.1.1[28]) is a key ingredient to the proof of Theorem 5.3.2.

Theorem 5.4.1. *The family of probability measures* $\hat{\mathbb{P}}^n_{k-1}$*(respectively,* $\mathbb{P}^n_{k-1}$*) converges weakly to probability measures* $\hat{\mathbb{P}}^0_{k-1}$*(respectively,* $\mathbb{P}^0_{k-1}$*), as* $n \to \infty$ *if and only if the following two conditions are satisfied:*

1. *the sequence* $\left\{(\phi^n, \hat{F}^n)\right\}_{n\geq 1}$ *converges weakly as diffusions to* $(\phi^{0,u}, F^{0,u})$ *as* $n \to \infty$*,*
2. *the sequence* $\left\{\hat{\mathbb{P}}^n_{k-1}\right\}$ *(respectively,* $\left\{\mathbb{P}^n_{k-1}\right\}$*) is tight.*

We will show first that under the condition of Theorem 5.3.2 the sequence $\left\{(\phi^n, \hat{F}^n)\right\}_{n\geq 1}$ converges weakly as diffusions to $(\phi^{0,u}, F^{0,u})$ as $n \to \infty$. We begin with the following lemma.

Lemma 5.4.1. *For each* $x \in \mathbb{R}^d$

$$\mathbb{E} \sup_{0\leq t\leq T} \left| \sum_{k=1}^{\infty} \int_0^t f_k(x,s) d\beta_k(s) \right|^2 < \infty, \tag{5.10}$$

$$\sup_n \mathbb{E} \sup_{0\leq t\leq T} \left| \sum_{k=1}^{\infty} \int_0^t f_k(\phi^n_s(x),s) d\beta_k(s) \right|^2 < \infty. \tag{5.11}$$

Proof. We will only prove (5.11). The proof of (5.10) follows in a similar manner. From the Bürkholder–Davis–Gundy inequality the left hand side of (5.11) is bounded by

$$c_1 \mathbb{E} \left| \sum_{l=1}^{\infty} \int_0^T \mathrm{Tr}\big(f_l f_l'\big)\big(\phi_r(x), r\big) dr \right| = c_1 \mathbb{E} \left| \int_0^T \mathrm{Tr}\Big(a(\phi_r(x), \phi_r(x), r)\Big) dr \right| \leq c_2 ||a||^{\sim}_{T,k,\delta}.$$

The last expression is finite since a belongs to $\widetilde{\mathbb{C}}^{k,\delta}_T(\mathbb{R}^{d\times d})$. □

An immediate consequence of Lemma 5.4.1 is the following corollary (c.f. (5.8), (5.9)).

Corollary 5.4.1. *For each $x \in \mathbb{R}^d$ and $t \in [0,T]$,*

$$\hat{F}^n(x,t) = \int_0^t \hat{b}_{u_n}(x,r)dr + S_n(x,t)$$

and

$$\phi_t^n(x) = x + \int_0^t \hat{b}_{u_n}(\phi_r^n(x),r)dr + T_n(x,t),$$

where $S_n(x,\cdot)$ and $T_n(x,\cdot)$ are continuous stochastic processes with values in $\mathbb{R}^d$, satisfying $\sup_{0\le t\le T}\big\{|S_n(x,t)| + |T_n(x,t)|\big\} \to 0$ in probability as $n \to \infty$.

The following lemma, showing the tightness of $\mathbb{P}^n_{(\mathbf{x},\mathbf{y})}$, plays an important role in the proof of the weak convergence as diffusions.

Lemma 5.4.2. *For each $x \in \mathbb{R}^d$ the sequence $\{\big(\phi^n(x), \hat{F}^n(x)\big)\}_{n\ge1}$ is tight in $\mathcal{C}\big([0,T] : \mathbb{R}^d \times \mathbb{R}^d\big)$.*

Proof. We will only argue the tightness of $\{\phi^n(x)\}$. Tightness of $\{\hat{F}^n(x)\}$ is proved similarly. Corollary 5.4.1 yields that $T_n(x,\cdot)$ is tight in $\mathcal{C}\big([0,T] : \mathbb{R}^d\big)$. Thus it suffices to show the tightness of $\big\{\int_0^\cdot \hat{b}_{u_n}\big(\phi_r^n(x),r\big)dr\big\}$. Fix $p > 0$. From the Cauchy–Schwarz inequality, (5.4), and recalling that $u_n \in \mathcal{A}_N[l_2]$, $\mathbb{E}\big|\int_s^t \hat{b}_{u_n}(\phi_r^n(x),r)dr\big|^p$ can be bounded by

$$\mathbb{E}\left[\int_s^t \big|\hat{b}_{u_n}\big(\phi_r^n(x),r\big)\big|^2 dr\right]^{p/2}(t-s)^{p/2} \le c_1\{||a||^{\sim}_{T,k,\delta} + ||b||^2_{T,k,\delta}\}^{p/2}(t-s)^{p/2} \le c_2(t-s)^{p/2}.$$

The result follows. □

Proposition 5.4.1. *Let $u_n \to u$ in distribution as $S_N[l_2]$–valued random variables. Then the sequence $\big\{(\phi^n, \hat{F}^n)\big\}_{n\ge1}$ converges weakly as diffusions to $(\phi^{0,u}, F^{0,u})$ as $n \to \infty$.*

Proof. In view of the tightness established in Lemma 5.4.2 and Corollary 5.4.1, it suffices to show that for each $t \in [0,T]$, the map $(\xi, v) \mapsto \int_0^t \hat{b}_v(\xi_s, s)ds$, from $\mathcal{C}([0,T] : \mathbb{R}^d) \times S_N[l_2]$

to $\mathbb{R}^d$, is continuous. Let $(\xi_n, v_n) \to (\xi, v)$ in $\mathcal{C}([0,T]:\mathbb{R}^d) \times S_N[l_2]$. Then,

$$\begin{aligned}\left|\int_0^t \big(\hat{b}_{v_n}(\xi_s^n, s) - \hat{b}_v(\xi_s, s)\big)ds\right| &\le \left|\int_0^t \big(\hat{b}_{v_n}(\xi_s^n, s) - \hat{b}_{v_n}(\xi_s, s)\big)ds\right| \\ &+ \left|\int_0^t \big(\hat{b}_{v_n}(\xi_s, s) - \hat{b}_v(\xi_s, s)\big)ds\right| \equiv L_1 + L_2. \end{aligned} \tag{5.12}$$

For each $x \in \mathbb{R}^d$ we have that

$$\left|\int_0^t \big(\hat{b}_{v_n}(x, s) - \hat{b}_v(x, s)\big)ds\right| = \left|\sum_{l=1}^{\infty}\int_0^t f_l(x,s)(v_l^n(s) - v_l(s))ds\right| \to 0, \tag{5.13}$$

since $v_n \to v$ weakly in $L^2([0,T]:l_2)$ and

$$\sum_{l=1}^{\infty}\int_0^t |f_l(x,s)|^2 ds \le T||a||^{\sim}_{T,k,\delta} < \infty.$$

Furthermore from (5.4) (recall $k \ge 1$) we have that for some $c_1 \in (0,\infty)$ and all $x, y \in \mathbb{R}^d$, $0 \le t \le T$,

$$\left|\int_0^t \big(\hat{b}_{v_n}(x, s) - \hat{b}_{v_n}(y, s)\big)ds\right| \le |x-y|\int_0^t \big(||b_{v_n}(s)||_{k,\gamma} + ||b(s)||_{k,\gamma}\big)ds \le c_1|x-y|. \tag{5.14}$$

Using the Ascoli-Arzelà Theorem (in the spatial variable) and equations (5.13), (5.14) yield now that the expression on the left side of (5.13) converges to 0 uniformly for x in compact subsets of $\mathbb{R}^d$. Thus $L_2 \to 0$ as $n \to \infty$. Following similar arguments L_1 is bounded by $c_2 \sup_{0\le s\le T}|\xi_s^n - \xi_s|$, which converges to 0 as $n \to \infty$. Hence the expression in (5.12) converges to 0 as $n \to \infty$ and the result follows. $\square$

We next show the tightness of the family of probability measures $\{\mathbb{P}_{k-1}^n\}$. Key ingredients in the proof are the following uniform L^p–estimates on $\partial^\alpha \hat{F}^n(x,t)$ and $\partial^\alpha \phi_t^n(x)$.

Lemma 5.4.3. *For each $p \ge 1$ there exists $k_1 \in (0,\infty)$ such that for all $t, t' \in [0,T]$, $x \in$*

$\mathbb{R}^d$, $n \geq 1$, *and* $|\alpha| \leq k$*:*

$$\mathbb{E}\big|\partial^\alpha \hat{F}^n(x,t) - \partial^\alpha \hat{F}^n(x,t')\big|^p \leq k_1\, |t-t'|^{p/2}\,. \tag{5.15}$$

Proof. Fix a multi–index α such that $|\alpha| \leq k$ and $p \geq 1$. Using the Bürkholder–Davis–Gundy inequality for the martingale $\partial^\alpha M(x,\cdot)$ and the fact that $a \in \widetilde{\mathbb{C}}_T^{k,\delta}(\mathbb{R}^d)$, we obtain that for some $c_1 \in (0,\infty)$ and all $x \in \mathbb{R}^d,\ t,t' \in [0,T]$,

$$\mathbb{E}\big|\partial^\alpha M(x,t) - \partial^\alpha M(x,t')\big|^p \leq c_1\, |t-t'|^{p/2}\,. \tag{5.16}$$

Recalling that $\hat{b}_{u_n}(\cdot,t) \in \mathbb{C}^{k,\gamma}(\mathbb{R}^d)$ a.e. (t,ω) and using (5.4) we get

$$\int_0^t \sup_{x\in\mathbb{R}^d} |\partial^\alpha \hat{b}_{u_n}(x,r)|dr < \infty \text{ a.e.},$$

and thus $\partial^\alpha \int_0^t b_{u_n}(x,r)dr = \int_0^t \partial^\alpha b_{u_n}(x,r)dr$ a.e. An application of the Cauchy–Schwarz inequality and (5.4) now gives, for some $c_2 \in (0,\infty)$,

$$\mathbb{E}\left|\partial^\alpha \int_{t'}^t b_{u_n}(x,r)dr\right|^p \leq c_2\, |t-t'|^{p/2}\,. \tag{5.17}$$

Equation (5.15) is an immediate consequence of (5.16) and (5.17). $\square$

For $g : \mathbb{R}^d \times [0,T] \to \mathbb{R}^d$, let $\nabla_y g(y,r)$ be the $d \times d$ matrix with entries $[\nabla_y g(y,r)]_{ij} = \frac{\partial}{\partial y_j} g_i(y,r)$. Differentiating with respect to x_1 in (5.9) we obtain

$$\begin{aligned}
\partial_1 \phi_t^n(x) &= \partial_1 x + \int_0^t [\nabla_y \hat{b}_{u_n}(\phi_r^n(x),r)\cdot \partial_1 \phi_r^n(x)]dr + \sqrt{\varepsilon_n}\int_0^t \nabla_y M(\phi_r^n(x),dr)\cdot \partial_1\phi_r^n(x) \\
&= \partial_1 x + \int_0^t \nabla_y \hat{F}^n(\phi_r^n(x),dr)\cdot \partial_1 \phi_r^n(x).
\end{aligned}$$

By repeated differentiation one obtains the following lemma whose proof follows along the lines of Theorem 3.3.3 of [28]. Given $0 \leq m \leq k$, let Λ_m be the set of all multi–

indices α satisfying $|\alpha| \leq m$. For a multi–index γ, denote by $m(\gamma) = \sharp\{\gamma_0 : |\gamma_0| \leq |\gamma|\}$. Also for a $|\gamma|$–times differentiable function $\Psi : \mathbb{R}^d \to \mathbb{R}$, denote by $\partial^{\leq|\gamma|}\Psi(x)$ the $m(\gamma)-$dimensional vector with entries $\partial^{\gamma_0}\Psi(x)$, $|\gamma_0| \leq |\gamma|$. If $\Psi = (\Psi_1, \Psi_2, \ldots, \Psi_d) : \mathbb{R}^d \to \mathbb{R}^d$ is such that each Ψ_i is $|\gamma|$–times continuously differentiable then $\partial^{\leq|\gamma|}\Psi(x) \doteq (\partial^{\leq|\gamma|}\Psi_1(x), \ldots, \partial^{\leq|\gamma|}\Psi_d(x))$. We will call a map $P : \mathbb{R}^m \to \mathbb{R}^d$ a polynomial of degree at most $\wp$ if $P(x) = (P_1(x), \ldots, P_d(x))'$ and each $P_i : \mathbb{R}^m \to \mathbb{R}$ is a polynomial of degree at most $\wp$. Also for $u, v \in \mathbb{R}^l$ we define $u * v \doteq (u_1v_1, \ldots, u_lv_l)'$.

Lemma 5.4.4. *Let α, β, γ be multi–indices such that $|\alpha|, |\beta|, |\gamma| \leq k$. Then there exist subsets Λ^1_α, Λ^2_α of $\Lambda_{|\alpha|}$ and $\Lambda_{|\alpha|-1}$ respectively, a subset $\Gamma^\alpha_{\beta,\gamma}$ of $\Lambda_{|\gamma|}$ and polynomials $P^\alpha_{\beta,\gamma} : \mathbb{R}^{m(\gamma)} \to \mathbb{R}^d$ of degree at most $|\alpha|$, such that $\partial^\alpha\phi^n$ satisfies:*

$$\partial^\alpha\phi^n_t(x) = \partial^\alpha x + \int_0^t G^n\big(\partial^\alpha\phi^n_r(x), \phi^n_r(x), dr\big) + \sum_{(\beta,\gamma)\in\Lambda^1_\alpha\times\Lambda^2_\alpha} \int_0^t G^{\alpha,n}_{\beta,\gamma}\big(\partial^{\leq|\gamma|}\phi^n_r(x), \phi^n_r(x), dr\big), \tag{5.18}$$

where for $x, y \in \mathbb{R}^d$, $G^n(x, y, r) = \nabla_y\hat{F}^n(y, r)\cdot x$ and for $(x, y) \in \mathbb{R}^{m(\gamma)}\times\mathbb{R}^d$, $G^{\alpha,n}_{\beta,\gamma}(x, y, r) = P^\alpha_{\beta,\gamma}(x) * \partial^\beta_y\hat{F}^n(y, r)$.

Note in particular that in the third term on the right hand side of (5.18), one finds partial derivatives of $\phi^n_r(x)$ of order strictly less than $|\alpha|$.

Lemma 5.4.5. *For each $p \geq 1$, $L \in (0, \infty)$, there is a constant $k_1 \equiv k_1(k, p, L) \in (0, \infty)$ such that for every multi–index α, $|\alpha| \leq k$*

$$\sup_n \sup_{|x|\leq L} \mathbb{E} \sup_{0\leq t\leq T} |\partial^\alpha\phi^n_t(x)|^p \leq k_1 \tag{5.19}$$

$$\sup_n \sup_{|x|\leq L} \mathbb{E}\big|\partial^\alpha\phi^n_t(x) - \partial^\alpha\phi^n_{t'}(x)\big|^p \leq k_1\,|t - t'|^{p/2}. \tag{5.20}$$

Proof. Fix $L > 0$ and consider $x \in \mathbb{R}^d$ such that $|x| \leq L$. We will first show inequality (5.19). It suffices to prove (5.19) for $\alpha = 0$ and establish that if, for some $m < k$, it holds for $\partial^\alpha\phi^n_t$ with $|\alpha| \leq m$ and all $p \geq 1$ then it also holds for $\partial_i\partial^\alpha\phi^n_t$ with all $p \geq 1$

(with a possibly larger constant k_1) and $i = 1, \ldots, d$. The desired result then follows by induction. Consider first $\alpha = 0$. For this case the bound in (5.19) follows immediately on using (5.4) and applying the Bürkholder–Davis–Gundy inequality to the square integrable martingale $N_t = \int_0^t M(\phi_r^n(x), dr)$ [note that $<< N >>_t = \int_0^t a(\phi_r^n(x), \phi_r^n(x), r) dr$ and $a \in \widetilde{\mathbb{C}}_T^{k,\delta}(\mathbb{R}^{d\times d})$].

Now, suppose that (5.19) holds for all multi–indices α with $|\alpha| \leq m$, for some $m < k$. Fix α with $|\alpha| \leq m$, an $i \in \{1, 2, \ldots, d\}$, and consider the multi–index $\tilde{\alpha} = \alpha + 1_i$, where 1_i is a d–dimensional vector with 1 in the ith entry and 0 elsewhere. From Lemma 5.4.4, one finds that $\partial^{\tilde{\alpha}} \phi_t^n$ solves (5.18) for $\alpha = \tilde{\alpha}$. Note that for $\beta \in \Lambda_{\tilde{\alpha}}^1$,

$$\partial_y^\beta \hat{F}^n(y,t) = \int_0^t \partial_y^\beta b_u(y,s) ds + \sqrt{\varepsilon_n} \partial_y^\beta M(y,t).$$

From (5.4) and recalling that $(b, a) \in \mathbb{C}_T^{k,\delta}(\mathbb{R}^d) \times \widetilde{\mathbb{C}}_T^{k,\delta}(\mathbb{R}^{d\times d})$, we have that for some $c_1, c_2 \in (0, \infty)$,

$$\sup_{0\leq t\leq T} \sup_{y\in\mathbb{R}^d} \left| \int_0^t \partial_y^\beta b_u(y,s) ds \right| \leq c_1 \text{ and } \sup_{0\leq t\leq T} \sup_{y\in\mathbb{R}^d} \left| << \partial_y^\beta M(y,t) >>_t \right| \leq c_2.$$

This along with the assumption

$$\sup_n \sup_{|x|\leq L} \mathbb{E} \sup_{0\leq t\leq T} |\partial^\nu \phi_t^n(x)|^p \leq k_1 \text{ for } \nu,\ |\nu| \leq |\alpha|,$$

shows that for some $c_3 \in (0, \infty)$, for all $(\beta, \gamma) \in \Lambda_{\tilde{\alpha}}^1 \times \Lambda_{\tilde{\alpha}}^2$

$$\sup_n \sup_{|x|\leq L} \mathbb{E} \sup_{0\leq t\leq T} \left| \int_0^t G_{\beta,\gamma}^{\tilde{\alpha},n} \left(\partial^{\leq|\gamma|} \phi_r^n(x), \phi_r^n(x), dr \right) \right|^p \leq c_3.$$

Also, in a similar manner one has for some $c_4 \in (0, \infty)$

$$\mathbb{E} \sup_{0\leq t\leq T} \left| \int_0^s G^n \left(\partial^{\tilde{\alpha}} \phi_r^n(x), \phi_r^n(x), dr \right) \right|^p \leq c_4 \int_0^t \mathbb{E} \left(\sup_{0\leq r\leq s} |\partial^{\tilde{\alpha}} \phi_r^n(x)|^p \right) ds.$$

Combining the above inequalities we obtain

$$\sup_n \sup_{|x|\le L} \mathbb{E} \sup_{0\le s\le t} |\partial^{\tilde\alpha} \phi^n_s(x)|^p \le c_3 + c_4 \sup_n \sup_{|x|\le L} \int_0^t \mathbb{E}\left(\sup_{0\le r\le s} |\partial^{\tilde\alpha}\phi^n_r(x)|^p \right) ds.$$

Now an application of Gronwall's lemma shows that for some $c_5 \in (0,\infty)$

$$\sup_n \sup_{|x|\le L} \mathbb{E} \sup_{0\le t\le T} |\partial^{\tilde\alpha}\phi^n_t(x)|^p \le c_5.$$

This establishes (5.19) for all $\tilde\alpha$ with $|\tilde\alpha| \le |\alpha|+1$. Finally consider (5.20). For $t, t' \in [0,T],\ t' \le t$, we have from (5.18) that

$$\begin{aligned}\partial^\alpha \phi^n_t(x) - \partial^\alpha \phi^n_{t'}(x) &= \int_{t'}^t G^n\big(\partial^\alpha \phi^n_r(x), \phi^n_r(x), dr\big)\\ &\quad + \sum_{(\beta,\gamma)\in\Lambda^1_\alpha\times\Lambda^2_\alpha} \int_{t'}^t G^{\alpha,n}_{\beta,\gamma}\big(\partial^{\le|\gamma|}\phi^n_r(x), \phi^n_r(x), dr\big). \qquad (5.21)\end{aligned}$$

Using (5.19) on the right hand side of (5.21) we now have (5.20) via an application of Hölder's and Bürkholder–Davis–Gundy's inequalities. □

The proof of Theorem 5.3.2 proceeds along the lines of Section 5.4 of [28]. We begin by introducing certain Sobolev spaces. Let j be a non–negative integer and let $1 < p < \infty$. Let $B_N \equiv B(0,N)$ be the $\mathbb{R}^d$–ball with center the origin and radius N. Let $h : \mathbb{R}^d \to \mathbb{R}^d$ be a function such that the distributional derivative (see, e.g. Chapter 6 [35]) $\partial^\alpha h \in L^p(B_N)$ for all α such that $|\alpha| \le j$. Define

$$||h||_{j,p:N} = \left(\sum_{|\alpha|\le j} \int_{B_N} \big|\partial^\alpha h(x)\big|^p dx \right)^{1/p}.$$

The space $H^{loc}_{j,p} = \{h : \mathbb{R}^d \to \mathbb{R}^d,\ ||h||_{j,p:N} < \infty \text{ for all } N\}$ together with the seminorms defined above is a real separable semi–reflexive Fréchet space. By Sobolev's imbedding theorem, we have $H^{loc}_{j+1,p} \subset \mathbb{C}^j(\mathbb{R}^d) \subset H^{loc}_{j,p}$ if $p > d$. Furthermore the imbedding i :

$H^{loc}_{j+1,p} \to \mathbb{C}^j(\mathbb{R}^d)$ is a compact operator by the Rellich–Kondrachov theorem (see [1]).

Proposition 5.4.2. *The sequence* $\{(\phi^n, \hat{F}^n)\}_{n\geq 1}$ *is tight in* $W_{k-1} \times W_{k-1}$.

Proof. It suffices to show that both $\{\phi^n\}_{n\geq 1}$ and $\{\hat{F}^n\}_{n\geq 1}$ are tight in W_{k-1}. We will use Kolmogorov's tightness criterion (see, e.g., Theorem 1.4.7, p.38, [28]). From Lemmas 5.4.3 and 5.4.5, we have that for each $p \geq 1,\ N > 1$, there exist $c_1, c_2 \in (0,\infty)$ such that for all $t, t' \in [0,T]$

$$\sup_n \mathbb{E}||\phi^n_t - \phi^n_{t'}||^p_{k,p:N} \leq c_1\,|t-t'|^{p/2}\,,$$
$$\sup_n \mathbb{E}||\hat{F}^n(\cdot,t) - \hat{F}^n(\cdot,t')||^p_{k,p:N} \leq c_2\,|t-t'|^{p/2}\,.$$

Furthermore, since $\hat{F}^n(\cdot,0) = 0$ and $\phi^n_0(x) = x$, we get that for each $p \geq 1,\ N > 1$ there exist $c_3, c_4 \in (0,\infty)$ such that

$$\sup_n \mathbb{E}||\phi^n_t||^p_{k,p:N} \leq c_3 \text{ and } \sup_n \mathbb{E}||\hat{F}^n(t)||^p_{k,p:N} \leq c_4.$$

Theorem 1.4.7 of [28] now gives tightness in the semiweak topology on $H^{loc}_{k,p}$ (cf. [28]). Since the imbedding map $i : H^{loc}_{k,p} \to \mathbb{C}^{k-1}$ is compact, tightness in $W_{k-1} \times W_{k-1}$ with the topology introduced in Section 2 follows (see pp. 246–247 [28]). $\square$

Recall the definitions (5.2) and (5.3). For the proof of the following lemma we refer the reader to Section 2.1 of [4].

Lemma 5.4.6. *Let* $f_n, f \in \hat{W}_{k-1}$ *be such that* $\sup_{0\leq t\leq T} \lambda_{k-1}(f_n(t), f(t)) \to 0$, *as* $n \to \infty$. *Then* $\sup_{0\leq t\leq T} d_{k-1}(f_n(t), f(t)) \to 0$.

PROOF OF THEOREM 5.3.2. Convergence as $\mathbb{C}^{k-1}$–flows is immediate from Theorem 5.4.1, Proposition 5.4.1 and Proposition 5.4.2. Using Skorohod's representation theorem, one can find a sequence of pairs $\{(\tilde{\phi}^n, \tilde{F}^n)\}_{n\geq 1}$ which has the same distribution as

$\{(\phi^n, \hat{F}^n)\}_{n\geq 1}$ and $\{(\tilde{\phi}^0, \tilde{F}^0)\}$ which has the same distribution as $\{(\phi^0, F^0)\}$ and

$$\sup_{0\leq t\leq T}\left[\lambda_k(\tilde{\phi}^n_t, \tilde{\phi}^0_t) + \lambda_k(\tilde{F}^n(t), \tilde{F}^0(t))\right] \to 0, \text{ a.s.}$$

Since $\phi^n, \phi^0 \in \hat{W}_k$ a.s., the same holds for $\tilde{\phi}^n, \tilde{\phi}^0$. Thus from Lemma 5.4.6

$$\sup_{0\leq t\leq T} d_{k-1}(\tilde{\phi}^n_t, \tilde{\phi}^0_t) \to 0 \text{ a.s.}$$

Hence $(\phi^n, \hat{F}^n) \to (\phi^0, F^0)$ as G^{k-1}–flows. □

5.5 Application to image analysis

A common approach to image matching problems (see [22], [31], [16] and references therein) is to consider a $\mathbb{R}^p$–valued, continuous and bounded function $T(\cdot)$, referred to as the "template" function, defined on a bounded open set $\mathcal{O} \subseteq \mathbb{R}^3$, which represents some canonical example of a structure of interest. By considering all possible smooth transformations $h : \mathcal{O} \to \mathcal{O}$ one can generate a rich library of targets (or images) given by the form $T\big(h(\cdot)\big)$.

In typical situations we are given data generated by an a priori unknown function h, and the key question of image matching is that of estimating h from the observed data. A Bayesian approach to this problem requires a prior distribution on the space of transformations and a formulation of a noise/data model. The "maximum" of the posterior distribution on the space of transformations given the data can then be used as an estimate $\hat{h}$ for the underlying unknown transformation h. In certain applications (e.g., medical diagnosis), the goal is to obtain numerical approximations for certain key structures present in the image, such as volumes of subregions, curvatures and surface areas. If the prior distribution on the transformations (and in particular the estimated transformation) is on the space of diffeomorphisms, then this information can be recovered from the template. Motivated by such a Bayesian approach a variational problem on the space of diffeomorphic flows was formulated and analyzed in [16].

Before going in to the description of this variational problem, we note that although the chief motivation for the variation problem studied in [16] came from Bayesian considerations, no rigorous results on relationships between the two formulations (variational and Bayesian) were established. The goal of our study is to develop a rigorous asymptotic theory that connects a Bayesian formulation for such an image matching problem with the variational approach taken in [16]. The precise result that we will establish is Theorem 5.5.1, given at the end of this section.

Let $\mathcal{C}_0^\infty(\mathcal{O})$ be the space of infinitely differentiable, real–valued functions on $\mathcal{O}$ with compact support in $\mathcal{O}$. The starting point of the variational formulation is a differential operator L on $[\mathcal{C}_0^\infty(\mathcal{O})]^3$, the exact form of which is determined from specific features of the problem under study. The formulation, particularly for problems from biology, often uses principles from physics and continuum mechanics as a guide in the selection of L. We refer the reader to Christensen et. al. [10], [11], where natural choices of L in shape models from anatomy are provided.

Define the norm $||\cdot||_L$ on $[\mathcal{C}_0^\infty(\mathcal{O})]^3$ by

$$||f||_L^2 \doteq \sum_{i=1}^{3} \int_{\mathcal{O}} |(Lf)_i(u)|^2 du,$$

where we write a function $g \in [\mathcal{C}_0^\infty(\mathcal{O})]^3$ as $(g_1, g_2, g_3)'$. It is assumed that $||\cdot||_L$ generates an inner product on $[\mathcal{C}_0^\infty(\mathcal{O})]^3$ and that the Hilbert space H defined as the closure of $[\mathcal{C}_0^\infty(\mathcal{O})]^3$ with this inner product is separable. We will need the functions in H to have sufficient regularity and thus assume that the norm $||\cdot||_L$ dominates an appropriate Sobolev norm. More precisely, let $W_0^{m+2,2}(\mathcal{O})$ be the closure of $\mathcal{C}_0^\infty(\mathcal{O})$ with respect to the norm

$$||g||_{W_0^{m+2,2}(\mathcal{O})} \doteq \Big(\int_{\mathcal{O}} \sum_{|\alpha|\le m+2} |\partial^\alpha g(u)|^2 du\Big)^{1/2}, \quad g \in \mathcal{C}_0^\infty(\mathcal{O}), \tag{5.22}$$

where α denotes a multi-index and $m \ge 3$. Define $\mathcal{V}_m \doteq [W_0^{m+2,2}(\mathcal{O})]^{\otimes 3}$, where $\otimes$ is used to denote the usual tensor product of Hilbert spaces. We denote by $||\cdot||_{\mathcal{V}_m}$ the norm on $\mathcal{V}_m$. The main regularity condition on L is the following domination requirement on the $||\cdot||_L$ norm. There exists a constant $c \in (0,\infty)$ such that

$$||f||_L \geq c||f||_{\mathcal{V}_m} \text{ for all } f \in [\mathcal{C}_0^\infty(\mathcal{O})]^3.$$

This condition ensures that $H \subseteq \mathbb{C}^{m,1/2}(\overline{\mathcal{O}})$ (see Theorem 4.12 parts II and III, p.85 [1]). We denote by $\mathcal{H}$ the Hilbert space $L^2([0,1] : H)$. For a fixed $b \in \mathcal{H}$ let $\{\eta_{s,t}(x)\}_{s \leq t \leq 1}$ be the unique solution of the ordinary differential equation

$$\frac{\partial \eta_{s,t}(x)}{\partial t} \doteq b\big(\eta_{s,t}(x), t\big), \quad \eta_{s,s}(x) = x, \ 0 \leq s \leq t \leq 1. \tag{5.23}$$

Then it follows that $\{\eta_{s,t}, 0 \leq s \leq t \leq T\}$ is a forward flow of $\mathbb{C}^m$–diffeomorphisms on $\mathcal{O}$ (see Theorem 4.6.5, p.173 [28]). Since $b(\cdot,t)$ has a compact support in $\overline{\mathcal{O}}$, one can extend $\eta_{s,t}$ to all of $\mathbb{R}^3$ by setting $\eta_{s,t}(x) \equiv x$, if $x \in \mathcal{O}^c$. Extended in this way $\eta_{s,t}$ can be considered as an element of G^m, as defined in Section 5.3. Denoting $\eta_{0,1}$ by h_b, we can now generate a family of smooth transformations (diffeomorphisms) on $\mathcal{O}$ by varying $b \in \mathcal{H}$. Specifically, the library of transformations which is used in the variational formulation of the image matching problem is $\{h_b | b \in \mathcal{H}\}$.

We now describe the data that is used in selecting the transformation h_{b^*} for which the image $T\big(h_{b^*}(\cdot)\big)$ best matches the data. Let $\mathcal{L}$ be a finite index set and $\{\mathbb{X}_i\}_{i\in\mathcal{L}}$ be a collection of disjoint subsets of $\mathcal{O}$ such that $\cup_{i\in\mathcal{L}}\mathbb{X}_i = \mathcal{O}$. Collected data $\{d_i\}_{i\in\mathcal{L}}$ represents integrated responses over each of the subsets $\mathbb{X}_i, \ i \in \mathcal{L}$. More precisely, if $T\big(h(\cdot)\big)$ was the true underlying image and the data were completely error free and noiseless, then $d_i = \int_{\mathbb{X}_i} T(h(\sigma))d\sigma/\mathrm{vol}(\mathbb{X}_i), \ i \in \mathcal{L}$, where $\mathrm{vol}(\mathbb{X}_i)$ denotes the Lebesgue measure of $\mathbb{X}_i$. Let $d = (d_1, d_2, .., d_n)'$, where $n = |\mathcal{L}|$. Defining $Y_d(x) = d_i, \ x \in \mathbb{X}_i, \ i \in \mathcal{L}$, the expression

$$\frac{1}{2}\int_{\mathcal{O}} \big|T(h_b(x)) - Y_d(x)\big|^2 dx$$

is a measure of discrepancy between a candidate target image $T(h_b(\cdot))$ and the observations. This suggests a natural variational criterion for selecting the "best" transformation matching the data. The objective function that is minimized in the variational formulation of the image matching problem is a sum of two terms, the first reflecting the "likelihood" of the transformation or change–of–variable h_b and the second measuring the conformity of the transformed template with the observed data. More precisely, define for $b \in \mathcal{H}$

$$J_d(b) \doteq \frac{1}{2}\Big(||b||^2_{\mathcal{H}} + \int_{\mathcal{O}} \big| T(h_b(x)) - Y_d(x) \big|^2 dx \Big). \tag{5.24}$$

Then $b^* \in \operatorname{argmin}_{b \in \mathcal{H}} J_d(b)$, represents the "optimal" velocity field that matches the data d and for which the h_{b^*}, obtained by solving (5.23), gives the "optimal" transformation. This transformation then yields an estimate of the target image as $T(h_{b^*}(\cdot))$. Equivalently, defining for each $h \in G^0$

$$\hat{J}_d(h) \doteq \inf_{b \in \Psi_h} J_d(b) \qquad (\text{where } \Psi_h = \{b \in \mathcal{H} : h = h_b\}),$$

we see that an optimal transformation is $h^* = h_{b^*} \in \operatorname{argmin}_h \hat{J}_d(h)$.

Up to a relabelling of the time variable, the above variational formulation(in particular the cost function in (5.24)) was motivated in [16] through Bayesian considerations, but no rigorous justification was provided. [In [16] the orientation of time is consistent with the change–of–variable evolving toward the identity mapping at the terminal time. To relate the variational problem to stochastic flows it is more convenient to have the identity mapping at time zero.] We next introduce a stochastic Bayesian formulation of the image matching problem and describe the precise asymptotic result that we will establish.

Let $\{\phi_i\}$ be a complete orthonormal system in H and $\beta \equiv (\beta_i)_{i=1}^{\infty}$ be as in Section 5.2, a sequence of independent, standard, real–valued Brownian motions on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$. Recall the space and its associated Borel $\sigma-$field $(\mathcal{C}([0,T] : \mathbb{R}^\infty), \mathcal{B}(\mathcal{C}([0,T] : \mathbb{R}^\infty))) \equiv (S, \mathcal{S})$ as introduced in Section 3.4, and note that

β is a random variable with values in S. Consider the stochastic flow

$$d\psi_{s,t}(x) = \sqrt{\varepsilon}\sum_{i=1}^{\infty}\phi_i\big(\psi_{s,t}(x)\big)d\beta_i(t), \;\; \psi_{s,s}(x) = x, \; x \in \mathcal{O}, \; 0 \le s \le t \le 1, \tag{5.25}$$

where $\varepsilon \in (0, \infty)$ is fixed. From Maurin's theorem (see Theorem 6.61, p. 202 [1]) it follows that the imbedding map $H \to \mathcal{V}_{m-2}$ is Hilbert–Schmidt. Also, $\mathcal{V}_{m-2}$ is continuously embedded in $\mathbb{C}^{m-2,1/2}(\overline{\mathcal{O}})$. Thus for some $k_1, k_2 \in (0, \infty)$ and all $u, x, y \in \mathcal{O}$,

$$\sum_{i=1}^{\infty}|\phi_i(u)|^2 \le k_1\sum_{i=1}^{\infty}||\phi_i||^2_{\mathcal{V}_{m-2}} < \infty,$$
$$\sum_{i=1}^{\infty}|\phi_i(x) - \phi_i(y)|^2 \le k_1|x-y|^2\sum_{i=1}^{\infty}||\phi_i||^2_{\mathcal{V}_{m-2}} = k_2|x-y|^2.$$

One also has that if ϕ_l is extended to all of $\mathbb{R}^3$ by setting $\phi_l(u) = 0$, for all $x \in \mathcal{O}^c$, then $a(x, y) = \sum_{l=1}^{\infty}\phi_l(x)\phi_l'(y)$ is in $\widetilde{\mathbb{C}}_T^{m-2,1/2}(\mathbb{R}^{3\times 3})$. Thus it follows (cf. p.80 and p.106 [28]) that

$$F(x,t) = \sum_{l=1}^{\infty}\int_0^t \phi_i(x)d\beta_i(r)$$

is a $\mathbb{C}^{m-2,\nu}$–Brownian motion, $0 < \nu < 1/2$, with local characteristics $(a, 0)$. Also (5.25) admits a unique solution $\{\psi^{\varepsilon}_{s,t}(x), 0 \le s \le t \le 1\}$ for each $x \in \mathcal{O}$ and $\{\psi^{\varepsilon}_{s,t}\}_{0\le s\le t\le 1}$ is a forward flow of $\mathbb{C}^k$–diffeomorphisms, with $k = m-2$, (see Theorem 4.6.5 [28]). In particular, $X^{\varepsilon} \doteq \psi^{\varepsilon}_{0,1}$ is a random variable in the space of $\mathbb{C}^k$–diffeomorpshisms on $\mathcal{O}$. The law of X^{ε} (for a fixed $\varepsilon > 0$) on G^k will be used as the prior distribution on the transformation space G^k. Note that $T\big(X^{\varepsilon}(\cdot)\big)$ induces a measure on the space of target images.

We next consider the data model. Let $\mathcal{L}$ and n be as introduced below (5.23). We suppose that the data is given through an additive Gaussian noise model:

$$D_i = \int_{\mathbb{X}_i} T\big(X^{\varepsilon}(x)\big)dx + \sqrt{\varepsilon}\xi_i$$

where $\{\xi_i,\ i \in \mathcal{L}\}$ is a family of independent, p–dimensional standard normal random variables.

In the Bayesian approach to the image matching problem one considers the posterior distribution of X^ε given the data D and uses the "mode" of this distribution as an estimate for the underlying true transformation. More precisely, let $\{\Gamma^\varepsilon\}_{\varepsilon>0}$ be a family of measurable maps from $\mathbb{R}^{np}$ to $\mathcal{P}(G^k)$ (the space of probability measures on G^k), such that

$$\Gamma^\varepsilon(A|D) = \mathbb{P}[X^\varepsilon \in A|D] \text{ a.s. for all } A \in \mathcal{B}(G^k).$$

We refer to $\Gamma^\varepsilon(\cdot|d)$ as a regular conditional probability distribution (r.c.p.d.) of X^ε given $D = d$. In Theorem 5.5.1 below, we will show that there is a r.c.p.d. $\{\Gamma^\varepsilon(\cdot|d),\ d \in \mathbb{R}^{np}\}_{\varepsilon>0}$ such that for each $d \in \mathbb{R}^{np}$, the family $\{\Gamma^\varepsilon(\cdot|d)\}_{\varepsilon>0}$, regarded as elements of $\mathcal{P}(G^{k-1}) \supseteq \mathcal{P}(G^k)$, satisfies a LDP with rate function

$$I_d(h) = \hat{J}_d(h) - \lambda_d, \text{ where } \lambda_d = \inf_{h \in G^{k-1}} \hat{J}_d(h) = \inf_{b \in \mathcal{H}} J_d(b).$$

Formally writing $\Gamma^\varepsilon(A|d) \approx \int_A e^{-\frac{I_d(h)}{\varepsilon}} dh$, one sees that for small ε, the "mode" of the posterior distribution given $D = d$, which represents the "optimal transformation" in the Bayesian formulation, can be formally interpreted as $\operatorname{argmin}_h I_d(h)$. Note that $\hat{J}_d(h) = \infty$ if $h \notin G^m$ (recall $m = k + 2$). Theorem 5.5.1 in particular says that $h \in G^m$ is a δ–minimizer for $I_d(h)$ if and only if it is also a δ–minimizer for $\hat{J}_d(h)$. Thus Theorem 5.5.1 makes precise the asymptotic relationship between the variational and the Bayesian formulation of the above image matching problem.

We say a sequence $\{\mathcal{Q}^\epsilon, \epsilon > 0\}$ of probabilities measures satisfies a LDP (as $\epsilon \to 0$) on some Polish space $\mathcal{E}$ if the corresponding sequence of canonical $\mathcal{E}-$valued random variables satisfies a LDP.

Theorem 5.5.1. *There exists an r.c.p.d. Γ^ε such that for each $d \in \mathbb{R}^n$, the family of probability measures $\{\Gamma^\varepsilon(d)\}_{\varepsilon>0}$ on G^{k-1} satisfies a large deviation principle (as $\varepsilon \to 0$)*

with rate function

$$I_d(h) \doteq \hat{J}_d(h) - \lambda_d. \tag{5.26}$$

We begin with the following proposition. Let $\tilde{I} : G^{k-1} \to [0, \infty]$ be defined as

$$\tilde{I}(h) \doteq \inf_{b \in \Psi_h} \frac{1}{2} ||b||^2_{\mathcal{H}}.$$

Proposition 5.5.1. *The family $\{X^\varepsilon\}_{\varepsilon>0}$ satisfies a LDP in G^{k-1} with rate function $\tilde{I}$.*

Proof. From Theorem 5.3.1 and an application of the contraction principle we have that $\{X^\varepsilon\}_{\varepsilon>0}$ satisfies LDP in G^{k-1} with rate function

$$I^*(h) \doteq \inf_{u \in \mathcal{L}^*(h)} \frac{1}{2} \int_0^T ||u(s)||^2_{l_2} ds,$$

where $\mathcal{L}^*(h) = \left\{u \in L^2([0,1] : l_2) | h = \phi^{0,u}(1)\right\}$ and where $\phi^{0,u}$ is defined via (5.6), but with f_i there replaced by ϕ_i in defining b_u. Note that there is a one to one correspondence between $u \in L^2([0,1] : l_2)$ and $b \in \mathcal{H}$ given as $b(t,x) = \sum_{l=1}^{\infty} u_l(t)\phi_l(x)$ and $\int_0^T ||u(s)||^2_{l_2} ds = ||b||^2_{\mathcal{H}}$. In particular $u \in \mathcal{L}^*(h)$ if and only if $b \in \Psi_h$. Thus $I^*(h) = \tilde{I}(h)$ and the result follows. ☐

Remark 5.5.1. *Proposition 5.5.1 is consistent with results in Section 5.3 in that although the local characteristics are in $\mathbb{C}^k$ and $X^\varepsilon \in G^k$, the LDP is established in the larger space G^{k-1}. This is due to the tightness issues described in Chapter 1. Furthermore, as noted below (5.23), if $||b||_{\mathcal{H}} < \infty$ then b induces a flow of $\mathbb{C}^m$–diffeomorphisms on $\mathcal{O}$. Thus if $h \in G^{k-1} \setminus G^m$ then Ψ_h is empty, and consequently $\tilde{I}(h) = \infty$. Hence there is a further widening of the "gap" between the regularity needed for the rate function to be finite and the regularity associated with the space in which the LDP is set. This is due to the fact that the variational problem is formulated essentially in terms of L^2 norms of derivatives, while in the theory of stochastic flows as developed in [28] assumptions are phrased in terms of L^∞ norms.*

Proposition 5.5.2. *For each $d \in \mathbb{R}^n$, I_d defined in* (5.26) *is a rate function on G^{k-1}.*

Proof. From (5.26) and the definition of $\tilde{I}$ we have for $h \in G^{k-1}$ that

$$I_d(h) = \tilde{I}(h) + \frac{1}{2}\int_{\mathcal{O}} |T(h(x)) - Y_d(x)|^2 dx - \inf_{h\in G^{k-1}} \left\{ \tilde{I}(h) + \frac{1}{2}\int_{\mathcal{O}} |T(h(x)) - Y_d(x)|^2 dx \right\}.$$

From Proposition 5.5.1, $\tilde{I}$ is a rate function and therefore has compact level sets. Additionally T is a continuous and bounded function on $\mathcal{O}$. The result follows. $\square$

PROOF OF THEOREM 5.5.1. We begin by noting that $\Gamma^\varepsilon(\cdot|d)$ defined as

$$\Gamma^\varepsilon(A|d) \doteq \frac{\int_A e^{-\frac{1}{2\varepsilon}\sum_{i=1}^n \left|d_i - \int_{\mathbb{X}_i} T(h(y))dy\right|^2} \mu^\varepsilon(dh)}{\int_{G^{k-1}} e^{-\frac{1}{2\varepsilon}\sum_{i=1}^n \left|d_i - \int_{\mathbb{X}_i} T(h(y))dy\right|^2} \mu^\varepsilon(dh)},$$

where $\mu^\varepsilon = \mathbb{P}\circ(X^\varepsilon)^{-1} \in \mathcal{P}(G^{k-1})$, is a r.c.p.d. of X^ε given $D = d$. Using the equivalence between Laplace principle and large deviations principle (see Section 1.2 [15]) it suffices to show that for all continuous and bounded real functions F on G^{k-1},

$$-\varepsilon \log \int_{G^{k-1}} \exp\left[-\frac{1}{\varepsilon}F(v)\right] \Gamma^\varepsilon(dv|d) \tag{5.27}$$

converges to $\inf_{h\in G^{k-1}}\{F(h) + I_d(h)\}$. Note that (5.27) can be expressed as

$$\begin{aligned} &-\varepsilon \log \int_{G^{k-1}} e^{-\frac{1}{\varepsilon}\left[F(h) + \frac{1}{2}\sum_{i=1}^n \left|d_i - \int_{\mathbb{X}_i} T(h(y))dy\right|^2\right]} \mu^\varepsilon(dh) \\ &+ \varepsilon \log \int_{G^{k-1}} e^{-\frac{1}{\varepsilon}\left[\frac{1}{2}\sum_{i=1}^n \left|d_i - \int_{\mathbb{X}_i} T(h(y))dy\right|^2\right]} \mu^\varepsilon(dh). \end{aligned} \tag{5.28}$$

From Proposition 5.5.1 we see that the first term converges to

$$\begin{aligned} &\inf_{h\in G^{k-1}} \left\{ \tilde{I}(h) + F(h) + \frac{1}{2}\sum_{i=1}^n \left| d_i - \int_{\mathbb{X}_i} T(h(y))dy \right|^2 \right\} \\ &= \inf_{h\in G^{k-1}} \inf_{b\in\Psi_h} \left\{ F(h) + \frac{1}{2}||b||^2_{\mathcal{H}} + \frac{1}{2}\int_{\mathcal{O}} |T(h(y)) - Y_d(y)|^2 dy \right\}, \end{aligned}$$

which is $\hat{J}_d(h)$. Likewise, the second term of (5.28) converges to $-\lambda_d$. This shows the result. $\square$

APPENDIX A

Functional Analysis Background

In this appendix we present basic functional analysis terminology and definitions used in this work. Since only real vector spaces are used in this work we will limit our presentation to this setting.

Definition A.0.1. *(Normed Linear Space). A real vector space $\mathbb{V}$ is said to be a normed linear space if for each $x \in \mathbb{V}$ there is associated a nonnegative real number $||x||$, called the norm of x, such that*

1. $||x + y|| \leq ||x|| + ||y||$ *for all* x *and* $y \in \mathbb{V}$,
2. $||ax|| = |a|||x||$ *if* $x \in \mathbb{V}$ *and* $a \in \mathbb{R}$,
3. $||x|| = 0$ *implies* $x = 0$.

Definition A.0.2. *(Banach Space). A Banach space is a normed linear space which is complete in the metric defined by its norm.*

Definition A.0.3. *(Inner Product Space). A real vector space $\mathbb{V}$ is said to be an inner product space if to each pair of vectors x and y in $\mathbb{V}$, there is associated a real number $< x, y >$, the so–called inner product of x and y, such that the following rules hold:*

1. $< x, x > \geq 0$ *for all* $x \in \mathbb{V}$ *and* $< x, x >= 0$ *if and only if* $x = 0$,
2. $< x, y >=< y, x >$, *for all* $x, y \in \mathbb{V}$,
3. $< ax + by, z >= a < x, z > +b < y, z >$ *for all* $x, y, z \in \mathbb{V}$ *and* $a, b \in \mathbb{R}$.

Definition A.0.4. *(Hilbert Space). A Hilbert space H is an inner product space which is complete.*

Remark A.0.2. *Note that an inner product space is a normed linear space and a Hilbert space is a Banach space.*

Definition A.0.5. *(Linear/Closed Operator). A linear operator A on a Banach space L is a linear mapping whose domain $D(A)$ is a subspace of L and whose range $R(A)$ lies in L. The graph of A is given by $\mathcal{G}(A) = \{(f, Af) : f \in D(A)\} \subset L \times L$. A is said to be closed if $\mathcal{G}(A)$ is a closed subspace of $L \times L$. We say A is densely defined if $D(A)$ is dense in L.*

Definition A.0.6. *(Bounded Linear Operator). A linear operator A on a Banach space L is called bounded if $D(A) = L$ and the image under A of any bounded subset of L is bounded. Linear bounded operators A from L_1 to L_2, where L_1 and L_2 are Banach spaces are defined similarly. We denote by $\mathcal{L}(L_1, L_2)$ as the space of all bounded linear operators from L_1 to L_2.*

Definition A.0.7. *(Strongly Continuous Semigroup). Let $\{A(t), t \geq 0\}$ be a family of closed, densely defined, linear operators on a Banach space L. We say that $\{A(t), t \geq 0\}$ generates a two parameters strongly continuous semigroup $\{U(t,s) : 0 \leq s \leq t\}$ if and only if*

1. *for all $0 \leq s \leq t$, $U(t,s) \in \mathcal{L}(L, L)$,*
2. *for $0 \leq s_1 \leq s_2 \leq t$, $U(t, s_2)U(s_2, s_1) = U(t, s_1)$,*
3. *$U(t,t) = I$ for all $t \geq 0$,*
4. *$|U(t,s)x - x| \to 0$ as $|t - s| \to 0$ for all $x \in L$,*
5. *$D(A(t)) = \left\{x \in \mathbb{V} : \text{ the limit } \frac{U(t+h,t)x - x}{h} \text{ in } \mathbb{V} \text{ exists as } h \to 0+\right\}$ and $A(t)x = \lim_{h\to 0} \frac{U(t+h,t)x - x}{h}$ for all $x \in D(A)$.*

Definition A.0.8. *(Compact Operator). An operator $A \in \mathcal{L}(H, H)$ for some Hilbert space H is said to be a compact operator if the image of any bounded subset of H is pre–compact in H.*

Definition A.0.9. *(Orthogonal/Complete Orthogonal/Complete Orthonormal System). A subset S of a Hilbert space H is called an orthogonal system if for any $x, y \in S$ such that $x \neq y$ we have $< x, y >= 0$. S is called a complete orthogonal system (COS) if there exists no other orthogonal system which strictly contains S. S is a complete orthonormal system (CONS) if S is a COS and, for any $x \in S$, $||x|| = 1$.*

Remark A.0.3. *(Seperable Hilbert Space) A Hilbert space is called separable if it admits a countable CONS. All Hilbert spaces here will be separable and so the adjective "separable" will be dropped.*

Definition A.0.10. *(Hilbert–Schmidt Operator). An operator $A \in \mathcal{L}(H, H)$ for some Hilbert space H is said to be a Hilbert–Schmidt operator if $||A||_{(2)} < \infty$, where $||A||_{(2)} \doteq \left(\sum_{n=1}^{\infty} ||Ae_n||^2\right)^{1/2}$ and $\{e_n\}$ a CONS of H.*

Definition A.0.11. *(Trace Class Operator). Let H be Hilbert space and $A \in \mathcal{L}(H, H)$ be a compact operator. A is said to be a trace class operator if there exists a CONS $\{e_n\}$ of H such that $\sum_{n=1}^{\infty} ||Ae_n|| < \infty$.*

Definition A.0.12. *(Symmetric/Nonnegative/Positive Operator). An operator $A \in \mathcal{L}(H, H)$ is called symmetric if $< Ax, y >=< x, Ay >$ for all $x, y \in H$. Such an operator is called nonnegative if $< Ax), x > \geq 0$ for all $x \in H$, and positive if $< Ax, x > > 0$ for all $x \in H \backslash \{0\}$.*

Remark A.0.4. *A nonnegative compact operator admits a CONS of eigenvectors. If $\{e_n\}$ is a CONS of eigenvectors of a nonnegative compact operator A with $\{\lambda_n\}$ the corresponding eigenvalues then $\lambda_n \to 0$ as $n \to \infty$. Such an operator is trace class if $\sum \lambda_n < \infty$ and Hilbert–Schmidt if $\sum \lambda_n^2 < \infty$.*

Definition A.0.13. *(Adjoint operator). Let H_1, H_2 be two Hilbert spaces and $A \in \mathcal{L}(H_1, H_2)$. An operator $A^* \in \mathcal{L}(H_2, H_1)$ is called adjoint of A if and only if $< Ax, y >=< x, A^*y >$ for all $x \in H_1$ and $y \in H_2$.*

APPENDIX B

Frequently used notations and assumptions

- $\mathbb{N}_0$: the space of nonnegative integers.
- $\mathbb{N}$: the space of positive integers.
- $\mathbb{R}$: the space of real numbers.
- H, H_α: Hilbert spaces. All Hilbert spaces in this work will be real and separable.
- $< \cdot, \cdot >$, $< \cdot, \cdot >_\alpha$, $\| \cdot \|, \| \cdot \|_\alpha$: the inner products and the norms of the above Hilbert spaces respectively.
- For a real–valued function f, $|f|_\infty = \sup_x |f(x)|$.
- Given a k and m dimensional continuous local martingales M, N on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ we will write the cross–quadratic variation of M and N as $<< M, N >>_t$. This is a continuous $\mathbb{R}^{m \times k}$–valued $\{\mathcal{F}_t\}$–adapted process.
- $L^2([0,T] : H)$: Hilbert space of maps f from $[0,T]$ to the Hilbert space H such that $\int_0^T ||f(t)||^2 dt < \infty$.
- $\mathcal{C}([0,T] : \mathcal{E})$: The space of continuous functions from $[0,T]$ to a Polish space $\mathcal{E}$.
- $X_n \xrightarrow{d} X$: X_n converges to X in distribution. The terms, convergence in distribution, convergence in law or weak convergence for random variables, will be used interchangeably.
- $X_n \xrightarrow{\mathbb{P}} X$: X_n converges to X in probability.
- Generic constants will be denoted as $c_1, c_2, \cdots$. Their values may change from one proof to next.

- Transpose of a d–dimensional vector v will be denoted by v'.
- By convention infimum on an empty set will be taken to be ∞.
- All Polish spaces in this work will be considered as measurable spaces endowed with the corresponding Borel σ–fields. Borel σ–fields on a Polish space $\mathcal{E}$ will be written as $\mathcal{B}(\mathcal{E})$.

BIBLIOGRAPHY

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces.* Academic Press, second edition, 2003.

[2] R. Azencott. *Grandes Déviations et Applications*, volume 774 of *Lecture Notes in Math.* Springer, Berlin, 1980.

[3] P. Baldi and M. Sanz-Solé. Modulus of continuity for stochastic flows. In *Barcelona Seminar on Stochastic Analysis (St. Feliu de Guíxols, 1991)*, volume 32 of *Progr. Probab.*, pages 1–20. Birkhäuser, Basel, 1993.

[4] P. Baxendale. Brownian motions in the diffeomorphism group. I. *Compositio Math.*, 53(1):19–50, 1984.

[5] G. Ben Arous and F. Castell. Flow decomposition and large deviations. *J. Funct. Anal.*, 140(1):23–67, 1996.

[6] P. Billingsley. *Convergence of Probability Measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.

[7] J. M. Bismut. *Mécanique aléatoire*, volume 866 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1981.

[8] A. Budhiraja and P. Dupuis. A variational representation for positive functional of infinite dimensional Brownian motions. *Probability and Mathematical Statistics*, 20:39–61, 2000.

[9] A. Budhiraja, P. Dupuis, and V. Maroulas. Large deviations for infinite dimensional stochastic dynamical systems. *To appear in Annals of Probability.*

[10] G. Christensen. *Deformable shape models for Anatomy.* PhD thesis, Department of Electrical Engineering, Sever Institute of Technology, Washington University, St. Louis, MO, August 1994.

[11] G. E. Christensen, R. D. Rabbitt, and M. I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, 1996.

[12] G. Da Prato and J. Zabcyk. *Stochastic Equations in Infinite–Dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1992.

[13] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Jones and Bartlett Publishers, 1993.

[14] J. D. Deuschel and D. W. Stroock. *Large Deviations.* Academic Press, 1989.

[15] P. Dupuis and R. S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations.* John Wiley & Sons, 1997.

[16] P. Dupuis, U. Grenander, and M.I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly in Applied Mathematics*, 56(3):587–600, 1998.

[17] K. D. Elworthy. Stochastic dynamical systems and their flows. In *Stochastic analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978)*, pages 79–95. Academic Press, New York, 1978.

[18] J Feng. Large deviation for diffusions and Hamilton–Jacobi equation in Hilbert spaces. *Ann. Probab.*, 34(1):321–385, 2006.

[19] J. Feng and T. G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2006.

[20] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, New York, second edition, 1998.

[21] A. Garsia. *Continuity properties of Gaussian processes with multidimensional time parameter.* University California Press, Berkeley, California, 1972.

[22] U. Grenander and M.I. Miller. Representations of knowledge in complex systems. *Journal of the Royal Statistical Society B*, 56(3):549–603, 1994.

[23] I. Ibragimov. On smoothness conditions for trajectories of random functions. *Theory of Probability and its Applications*, 28(2):240–262, 1983.

[24] O. Kallenberg. *Foundations of Modern Probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[25] G. Kallianpur and J. Xiong. *Stochastic Differential Equations in Infinite–Dimensional Spaces.* Institute of Mathematical Statistics Lecture Notes—Monograph Series, 26. Institute of Mathematical Statistics, Hayward, CA, 1995.

[26] G. Kallianpur and J. Xiong. Large deviations for a class of stochastic partial differential equations. *The Annals of Probability*, 24(1):320–345, 1996.

[27] P. Kotelenez. Existence, uniqueness and smoothness for a class of function valued stochastic partial differential equations. *Stochastics Stochastics Rep.*, 41(3):177–199, 1992.

[28] H. Kunita. *Stochastic Flows and Stochastic Differential Equations*, volume 24 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1997.

[29] H. J. Kushner and P. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2001.

[30] Y. Le Jan. Flots de diffusion dans $\mathbf{R}^d$. *C. R. Acad. Sci. Paris Sér. I Math.*, 294(21):697–699, 1982.

[31] M.I. Miller, G. E. Christensen, Y. Amit, and U. Grenander. Mathematical textbook of deformable Neuroanatomics. *Proceedings of the National Academy of Science*, 90(24), 1993.

[32] A. Millet, D. Nualart, and M. Sanz-Solé. Large deviations for a class of anticipating stochastic differential equations. *Ann. Probab.*, 20(4):1902–1931, 1992.

[33] J. Ren and X. Zhang. Freidlin-Wentzell's large deviations for homeomorphism flows of non-Lipschitz SDEs. *Bull. Sci. Math.*, 129(8):643–655, 2005.

[34] J. Ren and X. Zhang. Schilder theorem for the Brownian motion on the diffeomorphism group of the circle. *J. Funct. Anal.*, 224(1):107–133, 2005.

[35] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, 1991.

[36] R.B. Sowers. Large deviations for a reaction-diffusion equation with non–Gaussian perturbations. *The Annals of Probability*, 20(1):504–537, 1992.

[37] S. S. Sritharan and P. Sundar. Large deviations for the two-dimensional Navier-Stokes equations with multiplicative noise. *Stochastic Process. Appl.*, 116(11):1636–1659, 2006.

[38] J.B. Walsh. *An Introduction to Stochastic Partial Differential Equations*, volume 1180 of *Lecture Notes in Mathematics*. Springer, Berlin, 1986.

[39] W. Wang and J. Duan. Reductions and deviations for stochastic differential equations under fast dynamical boundary conditions. *Preprint*.