5-2014

Asymptotic Behavior of a Class of SPDEs

Parisa Fatheddin
University of Tennessee - Knoxville, pfathedd@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Parisa Fatheddin entitled "Asymptotic Behavior of a Class of SPDEs." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

Jie Xiong, Major Professor

We have read this dissertation and recommend its acceptance:

Xia Chen, Seddik Djouadi, Jan Rosinski

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Asymptotic Behavior of a Class of SPDEs

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Parisa Fatheddin

May 2014

To be a philosopher is not merely to have subtle thoughts, nor even to found a school, but so to love wisdom as to live according to its dictates, a life of simplicity, independence, magnanimity, and trust. It is to solve some of the problems of life, not only theoretically, but practically. -Henry David Thoreau

To My Family

Acknowledgements

I would like to thank the many faculty and staff at the University of Tennessee's Mathematics department. I was especially inspired by Prof. Michael Frazier, Prof. Stefan Richter, Prof. Jan Rosinski and Prof. Jie Xiong for their immense knowledge and depth in mathematics and have looked up to them as role models.

My appreciation goes to Prof. Jie Xiong for his support and for advising me on my research problems and inspiring me to understand and learn many concepts in modern probability. I have always been in awe of his great understanding in mathematics. His knowledge and great problem solving skills inspired me to go deeper in probability theory and in the problems I worked on, knowing that Prof. Xiong would know how to tackle major obstacles in case I face them.

In addition, I would like to thank Prof. Xia Chen for his course on Brownian motion and for serving as one of my committee members. I am also thankful to Prof. Seddik Djouadi for his time for reading this dissertation and serving on the committee. In addition, I am grateful to Prof. Conrad Pluat, Prof. Grozdena Todorova, and Prof. Steven Wise for helping me think in and understand concepts in Topology, PDE and Numerical Analysis respectively and set good foundations.

Abstract

We establish the large and moderate deviation principles for a class of stochastic partial differential equations with a non-Lipschitz continuous coefficient. As an application we derive these principles for an important population model, Fleming-Viot Process. In addition, we establish the moderate deviation principle for another classical population model, super-Brownian motion.

Table of Contents

Chapter 1

Introduction

Large deviations is a very active area of research in modern probability theory. It is the study of rare events that have probabilities decaying to zero exponentially fast and is concerned with determining the exact form of this rate of decay. Here we begin by providing an exposition of the many applications of large deviations.

The roots of large deviations were developed by Harald Cramèr, at the time that he served as an insurance consultant. He answered the question of what is the likelihood that the amount of money the insurance company has to pay to settle claims exceeds the amount of income it has earned. His answer was as the number of claims grows sufficiently large, this probability converges to zero exponentially fast and he gave its exact rate of decay. His theorem is given in more detail in chapter three. It was the first result in large deviations and opened the door to many applications in finance. In mathematical finance they refer to his theorem as the classical ruin problem.

Many important decisions in investments can be made by utilizing large deviations results. Similar to Cramèr's theorem, large deviation can be applied to risk management to determine the probability of large losses of a portfolio subject to market risk or the probability of a portfolio under credit risk. It is also used to approximate option pricing and the probability of stock market behaving in an unusual way that would devastate the investor. For more information on the applications of large deviations to finance and insurance, we refer the reader to the article written by H. Pham included in [8].

Applications of large deviations began with mathematical finance; however, it did not become limited to this area. It expanded to many applications in modeling, including queues and communication theory. Consider the situation of having n employees answering calls for a company. Large deviations can be used to find the probability of the number of calls coming in a time interval to exceed n. Many models in queuing theory, such as the one just described, can be given by jump Markov processes and large deviations provides the probabilities of errors and breakdowns of the system so that they can be avoided and the system can provide faster and less inclined to error service.

An important problem in large deviation theory is the exit problem introduced by Freidlin and Wentzell. It can be given as follows. Suppose a process $z_n(t)$ has a strong tendency to stay near a point q. We call this point a global attractor and let B be a ball of radius one around q. The question is how long does it take for the process to exit this ball. Related to queues, we let 0 be the attraction point and consider a ball of radius n around 0. Suppose the company can provide service for customers for busy periods of time less than n. Using the exit problem we can find the probability of a busy period being larger than n. To formulate this, let λ be the arrival rate to a queue, μ be the service rate, and $x(t)$ denote the number of customers in the queue. For a queue to be stable, that is for it to have a nondegenerate steady rate, we require $\lambda/\mu < 1$. Then for a stable queue,

$$\lim_{n\to\infty} \frac{1}{n} \log P_0 (\text{ busy period} \geq n) = -\left(\sqrt{\mu} - \sqrt{\lambda}\right)^2$$

Similarly, large deviations can be applied to determine how large the process can get during long busy periods. Shwartz and Weiss in [57] provide a great introduction and many examples to applications of large deviations to queues and communication models.

In physics and biology, large deviations has also proved to be fruitful. The main connection is entropy, which is the study of disorder and randomness in a system. The first person to connect probability theory to physical systems by studying entropy was Boltzmann. In many situations we are interested in the behavior of solids, liquids and gas at the molecular level, referred to as microscopic level. Thermodynamics provides properties of physical systems such as pressure and volume at the macroscopic level. The main goal of statistical mechanics is to study properties from a probability distribution, referred to as an ensemble, which provides information on its microscopic properties. It identifies macroscopic variables with ensemble averages of microscopic sums. For ideal gas, large deviations is used to prove that the microscopic sums converge exponentially to their ensemble averages as the number of particles increases to infinity. For many other applications of large deviations to statistical mechanics we recommend [18].

As seen in above examples, large deviations is used to examine rare events in large systems by offering analytic and less costly method to approximate the probabilities than the previous methods of simulation and numerical approximations. In models there are events that have a very small chance of occurring but have severe consequences when they do occur. These events are the core of studies of large deviations.

The techniques in large deviations can be applied to many processes in probability theory to give useful estimates. The law of large numbers states that for a sequence $\{X_n\}$ of i.i.d. random variables with $\mu = \mathbb{E}(X_1)$ we have

$$P\left(\frac{X_1 + X_2 + ... + X_n}{n} \to \mu\right) = 1$$

However, if we have a large number of X_i's then rare events will occur and change the true mean. For example, for a random walk, large deviations gives the tools to find, $P\left(x_0 + x_1 + ... + x_n \geq na\right)$ for some $a \geq \mathbb{E}(x_1)$, which can be interpreted as probability of arriving at an unlikely position. In general, for a sequence of random variables, $\{X_n\}$, large deviations principle is satisfied if the probability of X_n being in a class of Borel sets A, converges to zero exponentially fast as $n \to 0$ and the upperbound and lowerbound of these probabilities given by functions, referred to as rate functions, provide the exact rate of exponential convergence of these events. Therefore, this class of Borel sets, usually Borel measurable sets in the space of study, serve as sets of normal or usual events. For applications of large deviations to random walks, Markov chains and Markov processes and many classical results in the area see [14] and [16].

Among applications to processes in Biology, large deviations can be applied to population models as we have done in this manuscript. We have studied two of the most commonly used population models called super-Brownian motion(SBM) and Fleming-Viot Process(FVP). Here we study the large and moderate deviation principles for these models as the branching rate for SBM and the mutation rate for FVP tend to zero. We denote the population models as μ_t^ϵ and find the rate of convergence of these processes as ϵ is set to go to zero. Based on context, ϵ represents branching or mutation rate. For large deviations we consider the family $\{\mu_t^\epsilon\}$ itself, whereas for the case of moderate deviations we study this family multiplied by a collection $\{a(\epsilon)\}$ satisfying $0 \leq a(\epsilon) \to 0$ and $\frac{a(\epsilon)}{\sqrt{\epsilon}} \to \infty$ as $\epsilon \to 0$ and prove the large deviation principle for this sequence multiplied by the centered process.

LDP for Measure-Valued Processes (MVP) has been studied by many authors. Fleischmann and Kaj [25] proved the LDP for SBM for a fixed time t. Later on, sample path LDP for SBM was derived independently by Fleischmann *et al* [24], and Schied [54] while the rate function was expressed by a variational form. To obtain an explicit expression for the rate function, [24] assumed a *local blow-up condition* which was not proven. On the other hand, [54] obtained the explicit expression of the rate function when the term representing the movements of the particles also tends to zero. The local blow-up condition of [24] was recently removed by Xiang for SBM with finite and infinite initial measure, [67], [66] respectively, and the same explicit expression was established. Fleischmann and Xiong [26] proved an LDP for catalytic SBM with a single point catalyst. Making use of the Brownian snake representation introduced by LeGall, Serlet in [52] and [53] also obtained large deviation estimates for SBM. The successes of the LDP for SBM depend on the branching property of this process. This property implies the weak LDP directly, and hence the problem diminishes to showing the exponential tightness of SBM, which yields the LDP and identifying its rate function.

Since FVP does not possess the branching property, the derivation of LDP depends on new ideas. Dawson and Feng [12], [13], and Feng and Xiong [23] considered the LDP for FVP when the mutation is neutral. In [12], LDP was shown to hold when the process remains in the interior of the simplex, and in [13] the authors proved that if the process starts from the interior, it will not reach the boundary. On the other hand, authors in [23] focused on the singular case when the process starts from the boundary. For non-neutral case, Xiang and Zhang [68] derived an LDP for FVP when the mutation operator also tends to zero by projecting to the finite dimensional case. Our LDP for FVP contributes to the literature, by not requiring the neutrality and vanishing of mutation. We note that our method only applies to the case of superprocesses with spatial dimension one.

Authors in [31, 32, 55, 74, 76] have also considered moderate deviations for SBM. Hong investigated moderate deviations for SBM with super-Brownian immigration (SBMSBI) in [32] and for this process' quenched mean in [31]. In both cases he considered the space,

$$M_p(\mathbb{R}^d) := \left\{ \mu \in M(\mathbb{R}^d) : \langle \mu, f \rangle := \int f(x)\mu(dx) < \infty, \forall f \in \mathcal{C}_p(\mathbb{R}^d) \right\}$$

where

$$\mathcal{C}_p(\mathbb{R}^d) := \left\{ f \in \mathcal{C}(\mathbb{R}^d) : \sup \frac{|f(x)|}{\phi_p(x)} < \infty \text{ for } p > d, \phi_p(x) := (1 + |x|^2)^{-\frac{p}{2}} \right\}$$

Endowing space $M_p(\mathbb{R}^d)$ with the p-vague topology, that is $\mu_k \to \mu$ iff $\langle \mu_k, f \rangle \to \langle \mu, f \rangle$ for all $f \in M_p(\mathbb{R}^d)$, he established MDP in dimensions $d \geq 3$ for SBMSBI and in dimensions $3 \leq d \leq 6$ for the quenched mean of SBMSBI. The rate function for the two cases turn out to have a similar form. Schied [55] considered the SBM itself and established the MDP. In [55] he applied this result to prove the law of iterated logarithm for SBM, as well. He showed the MDP in space $\mathcal{C}([0, 1]; M(\mathbb{R}^d))$ equipped with compact open topology, where $M(\mathbb{R})$ is the space of finite signed measures on $\mathbb{R}^d$ with the coarsest topology such that $\mu \mapsto \langle \mu, f \rangle$ are continuous for every bounded Lipschitz function on $\mathbb{R}^d$. The main tool he used is the G$\ddot{a}$rtner-Ellis Theorem (cf. Theorem 4.6.1 of [14]). In this paper we also prove MDP for SBM obtaining the same result as Schied's, however with a different method. Yang and Zhang also proved MDP in [74] by applying the G$\ddot{a}$rtner-Ellis Theorem but for the occupation density process of the single point catalytic SBM in space $M(\mathbb{R})$ of all nonnegative measures μ on $\mathbb{R}$ such that $\langle \mu, f \rangle < \infty$ for all $f \in \Phi_+(\mathbb{R})$. Space $\Phi_+(\mathbb{R})$ is a separable Banach space composed of all continuous functions $f : \mathbb{R} \to \mathbb{R}$ such that $\lim_{|a| \to \infty} e^{|a|} f(a)$ exists in $\mathbb{R}$ with norm

$$\|f\| = \sup_{a \in \mathbb{R}} e^{|a|} |f(a)|$$

Using the above mentioned theorem the authors first showed the MDP for the occupation measure then with an application of the contraction principle (cf. [14] Theorem 4.2.1) proved the MDP for the occupation density field. Zhang [76] achieved the MDP for an immigration

SBM where the immigration is governed by the Lebesgue measure. Her setting was in $M_\rho(\mathbb{R}^d)$, the set of Borel measures μ on $\mathbb{R}^d$ such that $\langle \mu, \rho \rangle < \infty$, where ρ is a positive bounded function on $\mathbb{R}^d$ satisfying

$$e^{-\alpha t} P_t \rho(x) \to \rho(x) \text{ as } t \to 0$$

for $\alpha > 0$ and $x \in \mathbb{R}^d$. She equipped this space with the topology having the convention $\mu_k \to \mu$ iff $\langle \mu_k, f \rangle \to \langle \mu, f \rangle$ for all f in space,

$$\mathcal{C}_\rho(\mathbb{R}^d) := \left\{ f \in \mathcal{C}(\mathbb{R}^d) : \|f(x)\| \leq \rho \text{ for all } x \in \mathbb{R}^d \right\}$$

This dissertation is organized as follows. In chapter one we introduce the two population models of our study. In chapter three we establish the large deviations for FVP and achieve moderate deviations for both models FVP and SBM in chapter three. Our method is to first represent SBM and FVP in a form of a stochastic partial differential equations (SPDE) and prove the principles for this SPDE by applying the powerful technique offered by Budhiraja *et al* given in [7]. We then use the contraction principle (cf. Appendix B) to obtain the LDP and MDP for population models and we perform some calculations to derive the exact form of their rate functions. To our knowledge, MDP has not yet been proven for FVP. We note that here only superprocesses having spatial dimension one are considered. For information on LDP we refer the reader to the books of Dembo and Zeitouni [14] and Dupuis and Ellis [16] and for introductory material for the two population growth models studied in this article we recommend [11, 19].

Chapter 2

Super-Brownian Motion and Fleming-Viot Process

2.1 Introduction

Classes of measure-valued stochastic processes used to model evolving populations are referred to as superprocesses. These processes observe not only the size of the population, but the location of the individuals, as well. Superprocesses have been very useful in studying infinite-dimensional Markov processes and discovering many of their important sample properties such as hitting probabilities and moment functionals. There are also links between superprocesses and stochastic partial differential equations, which enable one to investigate the asymptotic behavior of superprocesses as done in this manuscript in Chapters Three, and Four.

The main advantage of using superprocesses to model populations is their general setting. One can consider $\mathbb{R}^d$ to characterize the spatial position of the individuals in the population. Two fundamental superprocesses that are commonly studied are super-Brownian motion (SBM) and Fleming-Viot Process (FVP). SBM considers the location of the individuals and views the population as a "cloud" that evolves through time. Since the size of the population is measured as finite measures on $\mathbb{R}^d$, then SBM is a measure-valued process. It is assumed that in this process each individual leaves behind a random number of offspring upon death.

FVP, on the other hand, observes the gene type of the individuals and keeps the population size fixed throughout time. These superprocesses are derived as the scaled limits of the appropriate discrete models using the Feller approximation.

In 1951, Feller suggested that to study a small population, one should investigate the discrete particle system and then use a continuous approximation to draw conclusions about the population. The main steps used in this technique are as follows. First the rescaled (discrete) process is proved to be tight. The rescaled process is $\{X^{(n)}\}_{n\geq 1}$ in which each individual has mass $1/n$ and at stage n an initial measure of $O(n)$ is taken. As a consequence of the complete and separable properties of the spaces that these models, including the ones mentioned above, take values in, it can be concluded that tightness implies relative compactness. By the definition of relative compactness, every subsequence has a further subsequence that converges in distribution. To obtain a continuous version by this approximation, one needs to justify that the limits of these subsequences are the same. To do so, a form of the limit is given by a martingale using the martingale representation theorem. This martingale forms a martingale problem and it is shown that it has a unique solution which becomes the limit and eventually the martingale characterization of the continuous version of the population model. In order to achieve the uniqueness of solutions to the martingale problem, the method of duality is implemented. This martingale characterization is a way to define the population models rigorously.

For more details and complete steps in achieving these approximations for SBM and FVP we refer the reader to [19] Chapter 1. The discrete model for SBM is the branching Brownian motion and for FVP is the stepwise mutation process. In this chapter we offer an introduction to the mentioned discrete models and give different ways to define the population models including their martingale characterization.

2.2 Super-Brownian Motion

One of the most frequently used superprocesses, was first introduced independently by S. Watanabe (1968) and D. A. Dawson (1975) and was originally referred to as Dawson-

Watanabe superprocess. In the late 1980s, E.B. Dynkin gave this superprocess a new name, super-Brownian motion (SBM). Like Brownian motion, SBM is used in models from various fields, not limited to the studies of population evolution. These areas include combinatorics (lattice tree and algebraic series), statistical mechanics, mathematical biology, interacting particle systems, and nonlinear partial differential equations.

The discrete particle counterpart of SBM is the branching Brownian motion, the oldest and best known branching process, in which individuals reproduce following a discrete process called the Galton-Watson process. We describe both of these discrete processes below.

In branching Brownian motion the spatial motion of the particles are studied in $\mathbb{R}^d$ in addition to the number of particles. As encoded in the name, each particle is assumed to move around following a Brownian motion. There is also an associated branching rate denoted as V and the lifetime of every particle is exponentially distributed with parameter V. The branching mechanism, Φ, is the probability generating function, $\Phi(s) = \sum_{k=0}^{\infty} p_k s^k$, where p_k is the probability that the number of particles in the n^{th} generation is k. This probability generating function is a useful tool in calculations regarding this process. Also offsprings are assumed to evolve independently of each other if conditioned on their time and place of birth. To model this process, let the unit point mass, δ_x, at x denote a particle at point $x \in \mathbb{R}^d$ and Y_t^i be the position of the i^{th} member of the population at time t. Then,

$$\xi_t = \sum \delta_{Y_t^i}$$

where the sum is over all individuals alive at time t, is the representation of the whole population at time t. Therefore, branching Brownian motion takes values in purely atomic measures on $\mathbb{R}^d$ by viewing each particle as an atom, and so is a measure-valued process. Since the number of offspring of each generation depends only on the number of offspring of the generation before, then this process has the Markov property. It also has the branching property given as follows: if $P_t(., \nu)$ is the transition probability with initial measure, ν, then,

$$P_t(., \nu_1 + \nu_2) = P_t(., \nu_1) * P_t(., \nu_2)$$

where $*$ is the convolution of measures. This means that the distribution of the process with initial value $\nu_1 + \nu_2$ is the same as the distribution of the sum of two independent copies of the process with initial value ν_1 and ν_2, respectively. Intuitively, the branching property states that the process can decompose into superprocesses that are identically distributed with each other and with the entire process. For more information on branching processes we refer the reader to [37] and [36].

In Galton-Watson process, each particle is assumed to live for precisely one unit of time and at the time of its death leave behind a random number of offspring at exactly the place of its death. Therefore, generations do not overlap and grow like a random tree. Galton-Watson process can be modeled via two ways: by backward or forward equation. The notion of backward equation is that every particle except the ancestor can be assigned to a subprocess traceable to a first-generation offspring of the ancestor. Let Z_{n+1} be the number of particles in the $(n+1)^{\text{st}}$ generation. There are Z_1 subprocesses from the ancestor. Let $Z_{1,n+1}^{(j)}$ denote the number of individuals in generation $n+1$ in the process starting by the ancestor, where (j) indicates the j^{th} i.i.d. copy. Then,

$$Z_{n+1} = \begin{cases} \sum_{j=1}^{Z_1} Z_{1,n+1}^{(j)} & Z_1 > 0 \\ 0 & Z_1 < 0 \end{cases}$$

As for forward equation, the notion is that every particle in the $(n+1)^{\text{st}}$ generation can be traced to its parent in the n^{th} generation. Let ξ_n^i denote the number of offsprings of the i^{th} particle in generation n. Then

$$\begin{aligned} Z_0 &= 1 \\ Z_{n+1} &= \begin{cases} \sum_{j=1}^{Z_n} \xi_n^i & Z_n > 0 \\ 0 & Z_n = 0 \end{cases} \end{aligned}$$

Now we give the key steps in attaining SBM from the branching Brownian motion with the underlying Galton-Watson process. First we rescale the branching Brownian motion by considering a large number, n of particles each having mass $1/n$ and lifetime $1/(nV)$. We assume that the number of offsprings of each particle has a Poisson distribution with parameter one. Let $X_t^{(n)}$ be the n^{th} rescaled process and $\mathcal{M}(\mathbb{R}^d)$ be the space of finite measures on $\mathbb{R}^d$ equipped with the weak topology. One shows that $\{X_t^{(n)}\}_{n\geq 1}$ is tight in $\mathcal{M}_F(\hat{\mathbb{R}}^d)$ where $\hat{\mathbb{R}}^d$ is the one point compactification of $\mathbb{R}^d$. Then any infinite subsequence contains a convergent subsequence and using the martingale characterization of branching Brownian motion, we obtain a unique limit point and so we conclude that the process converges and form its martingale characterization. The following definition gives the various ways of defining the SBM.

Definition 2.2.1 (super-Brownian motion)**.** *Denoted by μ_t^ϵ, SBM with branching rate, ϵ, is a measure-valued Markov process that can be characterized by one of the following.*
i) (μ_t^ϵ) having laplace transform,

$$\mathbb{E}_{\mu_0^\epsilon} \exp(- < \mu_t^\epsilon, f >) = \exp(- < \mu_0^\epsilon, v(t, \cdot) >)$$

where $v(\cdot, \cdot)$ is the unique mild solution of the evolution equation:

$$\begin{cases} \dot{v}(t, x) = \frac{1}{2}\Delta v(t, x) - v^2(t, x) \\ v(0, x) = f(x) \end{cases}$$

for $f \in \mathcal{C}_p^+(\mathbb{R}^d)$ where $\mathcal{C}_p(\mathbb{R}^d)$ was defined in the introduction.
ii) (μ_t^ϵ) as the unique solution to a martingale problem given as: for all $f \in \mathcal{C}_b^2(\mathbb{R})$

$$M_t(f) := \langle \mu_t^\epsilon, f \rangle - \langle \mu_0^\epsilon, f \rangle - \int_0^t \left\langle \mu_s^\epsilon, \frac{1}{2}\Delta f \right\rangle ds$$

is a square-integrable martingale with quadratic variation,

$$< M(f) >_t = \epsilon \int_0^t < \mu_s^\epsilon, f^2 > ds$$

iii) its "distribution" function-valued process u_t^ϵ defined as

$$u_t^\epsilon(y) = \int_0^y \mu_t^\epsilon(dx), \qquad \forall y \in \mathbb{R} \tag{2.1}$$

which can be used to present SBM by the following SPDE,

$$u_t^\epsilon(y) = F(y) + \int_0^t \int_0^{u_s^\epsilon(y)} W(dsda) + \int_0^t \frac{1}{2}\Delta u_s^\epsilon(y) ds \tag{2.2}$$

where $F(y) = \int_0^y \mu_0^\epsilon(dx)$ is the "distribution" function of μ_0, W is an $\mathcal{F}_t$-adapted space-time white noise random measure on $\mathbb{R}^+ \times U$ with intensity measure $ds\lambda(da)$, and $(U, \mathcal{U}, \lambda)$ is a measure space with λ denoting the Lebesgue measure. (This formulation is given in [69]).

Additional material on SBM can be found in [11, 17, 19].

2.3 Fleming-Viot Process

In 1979 W. Fleming and M. Viot introduced a class of probability measure-valued diffusion process, which became known as the Fleming-Viot process. This process observes the evolution of population based on the genetic type of individuals. The discrete version of Fleming-Viot process is called the step-wise mutation model, which is the continuous time version of Moran model. In step-wise mutation model, individuals move in $\mathbb{Z}^d$ according to a continuous-time simple random walk. As described by Ethier and Kurtz [20], since the population size is assumed to be constant, generations overlap in the sense that at the place of an individual's death, another person is born.

Step-wise mutation model and FVP have been used extensively in biology to model populations. In biology, mutation is the term given to the change in copies of DNA of parents to their offspring. DNA is a polymer having two long complementary strands with each strand having bases A, T, C and G. These strands are paired up with base A connected to base T and base C connected to base G. So base A is said to be complementary to base T and base C complementary to base G. In the copy of DNA from parent to offspring, if the sequence of bases is not identical or complementary then we say a mutation has occurred. Mutations can also happen in cell division causing abnormal growth such as the development of tumors. [36] offers an excellent discussion and background on the biological aspect of population models based on gene types.

In order to provide a rigorous mathematical definition of step-wise mutation process, some terminology and notations are required. As explained by Fleming and Viot in their original work [27], consider a population of large but finitely many individuals, each of whom has a "type" (usually genetic type) given by an element x in some set E. Not only are we concerned about the types of individuals but also with the distributions of types in the whole population. Given that E is a finite set, with J elements: $x_1, ..., x_J$, let p_j be the frequency of type x_j such that $p_j \geq 0, \sum_{j=1}^{J} p_j = 1$. Furthermore, the type of distribution is given by the vector, $p = (p_1, ..., p_J)$. Mutation in types is given by an operator $\mathcal{L}$, which acts as a linear, deterministic mechanism for change of type.

In analogue to the branching Brownian motion, step-wise mutation model assumes each individual to have an independent exponential "clock" with parameter γ, referred to as the sampling rate. During his lifetime, when this "clock" rings, he gets relocated to a position chosen at random from the empirical distribution of the population, where his mutation continues from this new position. This change in position is referred to as the sampling mechanism. If we assume that changes in type distribution are caused by mutation and chance fluctuations in the type distribution, then we are assuming "selective neutrality." On the other hand, if we let nature make the selections, we are said to consider the "non-neutral" case.

We now offer a mathematical formulation of step-wise mutation model. Let each individual be represented by an atom of mass $1/N$, where N is a fixed constant denoting the number of individuals in the whole population. Then the population at time t is considered by this model to be a probability measure, $\mathcal{P}(t)$ on $\mathbb{Z}^d$ and is given by the vector,

$$\mathcal{P}(t) = \left\{ p(t, z) : z \in \mathbb{Z}^d \right\}$$

where $p(t, z)$ is the proportion of population at point z and time t. In other words, $p(t, z)$ is the number of individuals at position z and time t divided by N. By this characterization, step-wise mutation model and so Fleming-Viot process are time homogeneous continuous time Markov processes.

Like the SBM, the rescaled process of stepwise mutation model can be passed to a limit to obtain the FVP. First tightness is proved for the rescaled process and then the generator of the rescaled process is used to find the limit as $N \to \infty$ and by the martingale characterization of the stepwise mutation process one is able to determine the martingale characterization of the FVP. Here the rescaled process, $Y_N(t)$ is the probability measure on the rescaled lattice, $\mathbb{Z}^d/\sqrt{N}$ where the mass assigned to the set A, is given by

$$Y_N(t, A) = \sum_{\frac{j}{\sqrt{N}} \in A} p(Nt, j)$$

[19] states these steps in more detail. The following provides three ways the FVP can be defined.

Definition 2.3.1 (Fleming-Viot Process)**.** *Denoted as (μ_t^ϵ), Fleming-Viot Process is a probability measure-valued Markov process and can be characterized by one of the following.*
i) a family of Markov process generated by $\mathcal{L}^\epsilon$ defined as

$$\begin{aligned}\mathcal{L}^\epsilon F(\mu_t^\epsilon) &= f'(<\mu_t^\epsilon, \phi>) <\mu_t^\epsilon, A\phi> \\ &\quad + \frac{\epsilon}{2}\int\int f''(<\mu_t^\epsilon,\phi>)\phi(x)\phi(y)Q(\mu_t; dx, dy)\end{aligned}$$

for $\epsilon > 0$ where

$$Q(\mu_t^\epsilon; dx, dy) = \mu_t^\epsilon(dx)\delta_x(dy) - \mu_t^\epsilon(dx)\mu_t^\epsilon(dy)$$

with δ_x denoting the Dirac measure at x and A is the generator of a Feller process. The operator $\mathcal{L}^\epsilon$ is given on the set,

$$\mathcal{D} = \{F(\mu_t^\epsilon) = f(<\mu_t^\epsilon, \phi>) : f \in \mathcal{C}_b^2(\mathbb{R}), \phi \in \mathcal{C}(\mathbb{R})\}.$$

(see [12] and [23] for this formulation).
ii) as the unique solution to the following martingale problem: for $f \in \mathcal{C}_c^2(\mathbb{R})$,

$$M_t(f) =<\mu_t^\epsilon, f> - <\mu_0^\epsilon, f> - \int_0^t <\mu_s^\epsilon, \frac{1}{2}\Delta f> ds$$

is a continuous square-integrable martingale with quadratic variation,

$$<M(f)>_t = \epsilon\int_0^t \left(<\mu_s^\epsilon, f^2> - <\mu_s^\epsilon, f>^2\right) ds$$

iii) using

$$u_t^\epsilon(y) = \mu_t^\epsilon((-\infty, y]) \tag{2.3}$$

FVP can be given by the following SPDE,

$$u_t^\epsilon(y) = F(y) + \int_0^t\int_0^1 \left(1_{a\le u_s^\epsilon(y)} - u_s^\epsilon(y)\right) W(dsda) + \int_0^t \frac{1}{2}\Delta u_s^\epsilon(y) ds \tag{2.4}$$

where the variables are the same as those for the SPDE representing the SBM given above.

More material on the FVP is given in [19, 22].

Chapter 3

Large Deviations

3.1 Introduction

As described earlier, large deviations is the study of probabilities of events that largely "deviate" from typical events. By typical events we mean those that satisfy the classical strong law of large numbers and central limit theorem. In other words, this theory is concerned with very rare events whose probabilities converge to zero exponentially fast and its goal is to determine asymptotic estimates for such probabilities.

The idea of large deviations was first introduced in 1937 by a Swedish probabilist, Harald Cramér. He was an insurance consultant at the time and to understand the problem he worked on, suppose an insurance company's income, p, and the number of claims from premium payments are fixed for each day. Since the size of each claim, X_t, is random, there is the risk that at the end of a period of length T, the total amount paid in setting the claims, $\sum_{t=1}^{T} X_t$, becomes greater than the total income from premium payments over the period. Cramer was concerned in determining the probability of such an event. That is finding $P\left(\sum_{t=1}^{T} X_t > pT\right)$ where pT is the total income in period T and the sizes of claims, $\{X_t\}_t$, are assumed to be independent, identically distributed (i.i.d.) random variables. He discovered that this probability decays exponentially fast as the number of claims becomes sufficiently large and provided an explicit form of this rate as a power series using methods in complex analysis. His theorem is stated as follows.

Theorem 3.1.1 (Cramér)**.** *Suppose $\{X_j\}$ is a sequence of i.i.d. $\mathbb{R}$-valued random variables and let $\{\mu_n\}$ denote the distribution law of $\hat{S}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$, then $\{\mu_n\}$ satisfies the following inequalities:*

a. For every open set $U \subset \mathbb{R}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(U) \geq -\inf_{x\in U} \Lambda^*(x)$$

b. For every closed set $C \subset \mathbb{R}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(C) \leq -\inf_{x\in C} \Lambda^*(x)$$

where $\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}} \{\lambda x - \log \mathbb{E}(\exp(\lambda X_1))\}$.

An extension of this result to $\mathbb{R}^d$ is shown in [14] Theorem 6.1.3 and Corollary 6.1.6. The types of estimates in above theorem set the stage for the theory of Large deviations and later in 1966, S.R.S Varadhan formulated this theory in a unified form, for which he received the 2007 Abel prize.

Large deviations is concerned with finding the rate at which probabilities of very unusual events go to zero and determining their asymptotic behavior. This rate of decay of probabilities is given by a lower semicontinuous map $I : \mathcal{E} \to [0, \infty]$ called a rate function where $\mathcal{E}$ is a Polish space (complete separable metric space). Recall that $f : X \to [0, \infty]$ is a lower semicontinuous map if for all $\alpha \in [0, \infty)$ the level sets $\{x : f(x) \leq \alpha\}$ are closed. When these level sets are compact, the rate function is referred to as a good rate function. Furthermore, if $\{X_n\}_{n\in\mathbb{N}}$ is a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$, taking values in a Polish space, $\mathcal{E}$, then the asymptotic estimates of the probability of their events that converge to zero exponentially fast are given by the Large Deviation Principle (LDP).

Definition 3.1.1 (Large Deviation Principle)**.** *The sequence $\{X_n\}_{n\in\mathbb{N}}$ satisfies the LDP on $\mathcal{E}$ with rate function I if it fulfills the following two conditions.*

a. LDP lower bound: for every open set $U \subset \mathcal{E}$,

$$-\inf_{x\in U} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \log P(X_n \in U)$$

b. LDP upper bound: for every closed set $C \subset \mathcal{E}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log P(X_n \in C) \leq - \inf_{x\in C} I(x)$$

Here we give some important facts in Large Deviations theory. It is known that if a sequence satisfies the LDP with a rate function, I, then the rate function is unique. The proof of this uniqueness can be found in [16] Theorem 1.3.1. If the LDP upperbound holds for all compact sets instead of closed sets, then the above principle is referred to as the Weak Large Deviation Principle. Also if the Weak Large Deviation Principle holds for a sequence and the sequence is exponentially tight (cf. Appendix), then LDP also holds with the same rate function, proof of which is given in [14] Lemma 1.2.18. Another important principle in this area is called the Laplace Principle given below.

Definition 3.1.2 (Laplace Principle)**.** *The sequence $\{X_n\}_{n\in\mathbb{N}}$ satisfies the Laplace Principle on $\mathcal{E}$ with rate function I if for every bounded continuous function $h : \mathcal{E} \to \mathbb{R}$, the following two conditions hold.*

a. Laplace Principle lowerbound:

$$- \inf_{x\in\mathcal{E}} \{h(x) + I(x)\} \leq \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left(exp(-nh(X_n))\right)$$

b. Laplace Principle upperbound:

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left(exp(-nh(X_n))\right) \leq - \inf_{x\in\mathcal{E}} \{h(x) + I(x)\}$$

It is shown in Corollary 1.2.5 of [16] that to obtain the Laplace principle, it is sufficient to satisfy the conditions in above definition for every bounded Lipschitz continuous functions, $h : \mathcal{E} \to \mathbb{R}$. Varadhan proved the implication of Laplace principle from LDP (cf. [16] Theorem 1.2.1) and W. Bryc established the equivalence between the two principles by showing the reverse implication (cf. Theorem 1.2.3 in [16] for a proof). As a consequence of this equivalence, Laplace principle shares many properties of LDP such as the uniqueness of rate function (cf. Theorem 1.3.1 of [16]). A stronger version of Laplace Principle is Uniform Laplace Principle. For a family of rate functions, $\{I_y\}$, parameterized by $y \in \mathcal{Y}$, if for all compact sets $K \subset \mathcal{Y}$, and $M < \infty$, the set

$$\bigcup_{y\in K} \{x \in \mathcal{E} : I_y(x) \leq M\}$$

is a compact subset of $\mathcal{E}$, then this family is said to have compact level sets on compacts.

Definition 3.1.3 (Uniform Laplace Principle). *Suppose $\{I_y\}$ is a family of rate functions on $\mathcal{E}$ parametrized by $y \in \mathcal{Y}$ where $\mathcal{Y}$ is a Polish space and assume this family has compact level sets on compacts, then a sequence of random variables, $\{X_n\}$ taking values on $\mathcal{E}$ satisfies the Laplace Principle on $\mathcal{E}$ with rate function I_y uniformly on compacts if for every compact set $K \subset \mathcal{Y}$ and continuous, bounded function, $h : \mathcal{E} \to \mathbb{R}$ the following condition holds.*

$$\lim_{n\to\infty} \sup_{y\in K} \left| \frac{1}{n} \log \mathbb{E}_y \left\{ \exp\left(-nh\left(X^n\right)\right)\right\} + \inf_{x\in\mathcal{E}} \left\{h(x) + I_y(x)\right\} \right| = 0$$

The classical approach to applying large deviations to stochastic dynamic system was introduced by Freidlin and Wentzell. For an overview of this method, suppose we have the small perturbed dynamic system

$$\dot{X}_t^\epsilon = b(X_t^\epsilon) + \epsilon\psi_t$$

where ψ_t is a stationary Gaussian process and suppose that the trajectories of the perturbed system, $\dot{x}_t = b(x_t)$ begin at points in a bounded domain D and never leave it and are attracted to a stable equilibrium point as t tends to infinity. The problem is to determine if the trajectories of the perturbed system will also stay in D. Since sup of $|\psi_t|$ is infinite for $t \in [0, \infty)$ then time is divided into countable number of intervals of lengths T and it is found that the probability of X_t^ϵ leaving D on any given time interval is very small. They use exponential estimates to obtain the rate functions given in large deviations. This method is also used to show that $P\left(\sup_{0\le t\le T} |X_t^\epsilon - x_t| \ge \delta\right)$ converges to zero for any $\delta > 0$. For more details on this approach see [28] and for examples of results applying this technique see [59, 61].

The above method of time discretization and finding the appropriate exponential estimates proves to be very technical and difficult for many dynamic systems. An alternative approach given in [16], is called the weak convergence method and is based on variational representation formulas introduced by Budhiraja and Dupuis in [5]. To understand this approach, suppose $\{X_n\}$ is a sequence of random variables taking values in Polish space χ and let $\mathcal{P}(\chi)$ be the set of probability measures on χ. For γ and θ in $\mathcal{P}(\chi)$, let $R(\gamma\|\theta)$ be the relative entropy of γ with respect to θ. Further, let θ^n be the distribution of X_n, then by a variational formula we have,

$$\begin{aligned} -\frac{1}{n}\log\mathbb{E}\left(e^{-nh(X_n)}\right) &= -\frac{1}{n}\log\int_{\chi} e^{-nh}d\theta^n \\ &= \inf_{\mu\in\mathcal{P}(\chi)}\left(\int_{\chi} hd\mu + \frac{1}{n}R(\mu\|\theta^n)\right) \end{aligned} \tag{3.1}$$

Note that this is what we need for Laplace principle and it is sufficient to determine the rate function I on χ such that as $n\to\infty$, (3.1) converges to $\inf_{x\in\chi}\{h(x)+I(x)\}$. In [16] dynamic programming is used to evaluate this limit as the minimal cost function of a stochastic optimal control problem.

The key step in the weak convergence approach is the use of variational representation formulas. In [5] the authors established a variational representation formula for functionals of Brownian motion and gave weak convergence conditions that ensure the LDP of the process. Some useful examples of variational representation formulas include one for an infinite sequence $\beta=\{\beta_i\}$ of independent, standard real Brownian motions, given by

$$-\log\mathbb{E}\left(\exp(-f(\beta))\right) = \inf_{u\in\mathcal{P}_2(\ell_2)}\mathbb{E}\left(\frac{1}{2}\int_0^T\|u(s)\|_{\ell_2}^2ds + f\left(\beta+\int_0^{\cdot}u(s)ds\right)\right)$$

for bounded, Borel measurable function, $f:\mathcal{C}\left([0,T];\mathbb{R}^\infty\right)\to\mathbb{R}$, where $\mathcal{P}_2(\ell_2)$ is the family of all ℓ_2 valued predictable processes u for which
$\sum_{i=1}^\infty\int_0^T|u_i(s)|^2ds<\infty$ a.s. Also for a Brownian sheet, B, we have the following variational representation,

$$-\log\mathbb{E}\left(\exp\left(-f(B)\right)\right) = \inf_{u\in\mathcal{P}_2}\mathbb{E}\left(\frac{1}{2}\int_0^T\int_{\mathcal{O}}u^2(s,r)drds + f(B^u)\right)$$

where $\mathcal{O}$ is a bounded set in $\mathbb{R}$, $f:\mathcal{C}\left([0,T]\times\mathcal{O}:\mathbb{R}\right)\to\mathbb{R}$ is a bounded measurable map, $\mathcal{P}_2$ is the class of all predictable processes f such that $\int_{[0,T]\times\mathcal{O}}f^2(s,x)dsdx$ is finite a.s. Also

$$B^u(t,x) = B(t,x) + \int_0^t\int_{(-\infty,x]\cap\mathcal{O}}u(s,y)dyds$$

For variational representation for a function of a Q-Wiener process see [7]. In addition, we mention that other useful variational representation formulas such as those for functionals of Poisson random measures, were also established in [6, 9, 10].

Since the weak convergence approach based on the variational representation formulas did not require the exponential estimates as the previous technique, it was an important and ground breaking result that generated many publications in LDP theory on dynamic systems. From among them are the LDP for stochastic quasi-geostrophic equations by Liu, Röckner and Zhu in [40], LDP for optimal filtering with fractional Brownian motion by Maroulas and Xiong in [41], LDP for multivalued SDEs with monotone drifs by Ren, Xu and Zhang [46], LDP for two-dimensional Navier-Stokes equations with multiplicative noise in bounded and unbounded domains by Sritharan and Sundar in [58], and LDP for stochastic tamed three-dimensional Navier-Stokes equations with small noise by Röckner, T. Zhang and X. Zhang in [48]. In addition, Ren and Zhang [47], Liu [39] and Arani and Zangeneh [1] used the weak convergence approach to achieve LDP for different classes of stochastic evolution equations. The results in [5] were extended in [7] by obtaining LDP for functionals of infinite sequence of independent standard real Brownian motions and functional of a Brownian sheet.

A common assumption that papers using weak convergence approach make is the Lipschitz continuity of the coefficient of the noise term so to achieve the uniqueness of solutions. Since we do not meet this condition, we have altered our space to achieve a form of uniqueness. For more depth on the weak convergence approach we refer the reader to the excellent source [16] by Dupuis and Ellis.

Our goal in this chapter is to establish the LDP first for a class of SPDEs and as an application prove the LDP for the two population models described in chapter one. More precisely, our aim is to determine the limiting behavior of SBM and FVP as the branching rate V for SBM, and the mutation rate γ for FVP, converge to zero.

We begin with some notations used for the current chapter and chapter four. Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $\{\mathcal{F}_t\}$ is a family of non-decreasing right continuous sub σ-fields of $\mathcal{F}$ such that $\mathcal{F}_0$ contains all P-null subsets of Ω. We denote $\mathcal{C}_b(\mathbb{R})$ to be the space of continuous bounded functions on $\mathbb{R}$ and $\mathcal{C}_c(\mathbb{R})$ to be composed of continuous functions in $\mathbb{R}$ with compact support. In addition, for $0 < \beta \in \mathbb{R}$, we let $\mathcal{M}_\beta(\mathbb{R})$ denote the set of σ-finite measures μ on $\mathbb{R}$ such that

$$\int e^{-\beta|x|} d\mu(x) < \infty. \tag{3.2}$$

The topology of $\mathcal{M}_\beta(\mathbb{R})$ is defined by the following modified Wasserstein distance,

$$\begin{aligned} & \rho_\beta(\mu,\nu) \\ := \; & \sup\left\{\left|\int_{\mathbb{R}} f(x)e^{-\beta|x|}(\mu(dx)-\nu(dx))\right| : f\in\mathcal{C}_b^1(\mathbb{R}), \|f\|_\infty \vee \|f'\|_\infty \le 1\right\} \end{aligned} \tag{3.3}$$

Similarly, let $\mathcal{P}_\beta(\mathbb{R})$ be the set of probability measures with distance given by (3.3). We denote $(S,\mathcal{S})$ to be the measurable space defined as

$$(S,\mathcal{S}) := (\mathcal{C}([0,1];\mathbb{R}^\infty), \mathbb{B}(\mathcal{C}([0,1];\mathbb{R}^\infty)))$$

where $\mathbb{R}^\infty$ is the Polish space with the metric given as

$$d(\{x_i\},\{y_i\}) := \sum_{i=1}^\infty 2^{-i}(|x_i-y_i|\wedge 1)$$

where $\{x\}$ denotes the least integer greater than $|x|$.

Throughout this manuscript, we assume K to be a constant that can change from place to place. For $\alpha\in(0,1)$, we consider the space $\mathbb{B}_{\alpha,\beta}$ composed of all functions $f:\mathbb{R}\to\mathbb{R}$ such that for all $m\in\mathbb{N}$,

$$|f(y_1)-f(y_2)| \le Ke^{\beta m}|y_1-y_2|^\alpha, \quad \forall |y_1|,|y_2|\le m \tag{3.4}$$

$$|f(y)| \le Ke^{\beta|y|}, \quad \forall y\in\mathbb{R} \tag{3.5}$$

with the metric,

$$d_{\alpha,\beta}(u,v) = \sum_{m=1}^\infty 2^{-m}(\|u-v\|_{m,\alpha,\beta}\wedge 1), \qquad u,\ v\in\mathbb{B}_{\alpha,\beta}$$

where

$$\|u\|_{m,\alpha,\beta} = \sup_{x\in\mathbb{R}} e^{-\beta|x|}|u(x)| + \sup_{y_1\ne y_2}\frac{|u(y_1)-u(y_2)|}{|y_1-y_2|^\alpha}e^{-\beta m}.$$

Note that the collection of continuous functions on $\mathbb{R}$ satisfying (3.5), referred to as $\mathbb{B}_\beta$, is a Banach space with norm,

$$\|f\|_\beta = \sup_{x\in\mathbb{R}} e^{-\beta|x|}|f(x)|$$

Recall for SBM, μ_t^ϵ, we have the following representation:

$$u_t^\epsilon(y) = F(y) + \int_0^t\int_0^{u_s^\epsilon(y)} W(dads) + \int_0^t \frac{1}{2}\Delta u_s^\epsilon(y)ds \tag{3.6}$$

where $u^\epsilon_t(y) := \mu^\epsilon_t([0,y])$ for all $y \in \mathbb{R}$, $F(y) = \int_0^y \mu^\epsilon_0(dx)$, μ^ϵ_0 being the initial measure of the process and W is a white noise random measure on $\mathbb{R}^+ \times \mathbb{R}$ with intensity measure $ds\lambda(da)$.

Similarly, FVP, μ^ϵ_t, is given as the unique solution to

$$u^\epsilon_t(y) = F(y) + \int_0^t \int_0^1 \left(1_{a \le u^\epsilon_s(y)} - u^\epsilon_s(y)\right) W(dads) + \int_0^t \frac{1}{2}\Delta u^\epsilon_s(y) ds \tag{3.7}$$

where $u^\epsilon_t(y) = \mu^\epsilon_t((-\infty, y])$ for all $y \in \mathbb{R}$, W is a measure here on $\mathbb{R}^+ \times [0,1]$ and $F(y)$ is defined the same as that for SBM. Note that the main difference between (3.6) and (3.7) is in the second term; therefore, we consider a general stochastic partial differential equations (SPDE) with small noise term of the form

$$u^\epsilon_t(y) = F(y) + \sqrt{\epsilon}\int_0^t \int_U G(a, y, u^\epsilon_s(y)) W(dads) + \int_0^t \frac{1}{2}\Delta u^\epsilon_s(y) ds \tag{3.8}$$

with conditions,

$$\int_U \left|G(a,y,u_1) - G(a,y,u_2)\right|^2 \lambda(da) \quad \le \quad K|u_1 - u_2| \tag{3.9}$$

$$\int_U |G(a,y,u)|^2 \lambda(da) \quad \le \quad K(1 + |u|^2) \tag{3.10}$$

where $u_1, u_2, u, y \in \mathbb{R}$, F is a function on $\mathbb{R}$ and $G : U \times \mathbb{R}^2 \to \mathbb{R}$. The white noise W is a random measure on $\mathbb{R}^+ \times U$. Also we require its control PDE, defined for every $h \in L^2\left([0,1] \times U, ds\lambda(da)\right)$ as

$$u_t(y) = F(y) + \int_0^t \int_U G(a, y, u_s(y)) h_s(a) \lambda(da) ds + \int_0^t \frac{1}{2}\Delta u_s(y) ds \tag{3.11}$$

having the property that if there exists $h_1 \in L^2\left([0,1] \times U, ds\lambda(da)\right)$ with which both u_1, u_2 satisfy (3.11) and a function $h_2 \in L^2\left([0,1] \times U, ds\lambda(da)\right)$ with which both u_2, u_3 satisfy the above controlled PDE, then there exists an $h_3 \in L^2\left([0,1] \times U, ds\lambda(da)\right)$ such that both u_1, u_3 satisfy (3.11) with h_3.

It is not difficult to see that the middle term in both (3.6) and (3.7) satisfy the conditions imposed on G. Also by the choice of spaces $\mathcal{M}_\beta(\mathbb{R})$ and $\mathcal{P}_\beta(\mathbb{R})$, we have that $F(y)$ in both SBM and FVP is an element of $\mathbb{B}_{\alpha,\beta_0}$ space. More precisely, for SBM,

$$\begin{aligned} e^{-\beta_0 m}|F(y_1) - F(y_2)| &= e^{-\beta_0 m}\left|\int_0^{y_1} \mu^\epsilon_0(dx) - \int_0^{y_2} \mu^\epsilon_0(dx)\right| \\ &\le \int_{y_1}^{y_2} e^{-\beta_0 m} \mu^\epsilon_0(dx) \\ &\le K \end{aligned}$$

Similarly,

$$|F(y)| \leq Ke^{\beta|y|}$$

The same reasoning shows that the function F in FVP also is an element of $\mathbb{B}_{\alpha,\beta_0}$. Thus, we begin our study by proving the LDP for SPDE (3.8) with $F \in \mathbb{B}_{\alpha,\beta_0}$, which we refer to as the general SPDE and then apply our results to the population models. Specifically, SBM satisfies the general SPDE with

$$F(y) = \int_0^y \mu_0^\epsilon(dx), \;\; U = \mathbb{R}, \;\; \lambda(da) = da \quad \text{and } G(a,y,u) = 1_{0\leq a\leq u} + 1_{u\leq a\leq 0} \tag{3.12}$$

and for the case of FVP we have

$$F(y) = \mu_0((-\infty, y]), \;\; U = [0,1], \;\; \lambda(da) = da \quad \text{and } G(a,y,u) = 1_{a<u} - u. \tag{3.13}$$

Throughout this chapter $\beta_0 \in (0,\beta)$ and $\beta_1 \in (\beta_0,\beta)$.

3.2 Large Deviations for the General SPDE

Our goal in this section is to achieve the LDP for SPDE (3.8). We begin by showing that the solution to SPDE (3.8) takes values in space $\mathcal{C}([0,1];\mathbb{B}_\beta)$. Inspired by Shiga [56], to obtain this regularity of the solution and for its tightness to be used in a later section, the following refined version of Kolmogorov's criterion is proved and applied.

Lemma 3.2.1. *Let $\beta_0 \in (0,\beta)$ and $\beta_1 \in (\beta_0,\beta)$. Suppose $\{v_t^\epsilon(y)\}$ is a sequence of random fields. If there exist constants n, q, $K > 0$ such that*

$$\mathbb{E}\left|v_{t_1}^\epsilon(y_1) - v_{t_2}^\epsilon(y_2)\right|^n \leq Ke^{n\beta_1(|y_1|\vee|y_2|)}\left(|y_1-y_2| + |t_1-t_2|\right)^{2+q}, \tag{3.14}$$

then, there exists a constant $\alpha > 0$ such that

$$\sup_{\epsilon>0}\mathbb{E}\left|\sup_{m\in\mathbb{N}}\sup_{t_i\in[0,1],|y_i|\leq m,i=1,2}\frac{\left|u_{t_1}^\epsilon(y_1) - u_{t_2}^\epsilon(y_2)\right|}{\left(|y_1-y_2|+|t_1-t_2|\right)^\alpha}e^{-\beta m}\right|^n < \infty \tag{3.15}$$

As a consequence, $v_\cdot^\epsilon \in \mathcal{C}([0,1];\mathbb{B}_\beta)$ a.s.

Furthermore, if condition (3.14) holds and $\sup_{\epsilon>0}\mathbb{E}|v_{t_0}^\epsilon(y_0)|^n < \infty$ for some $(t_0,y_0) \in [0,1]\times\mathbb{R}$ then

$$\sup_{\epsilon>0}\mathbb{E}\left|\sup_{(t,y)\in[0,1]\times\mathbb{R}} e^{-\beta|y|}\left|v_t^\epsilon(y)\right|\right|^n<\infty \tag{3.16}$$

and the family $\{v_\cdot^\epsilon\}$ *is tight in* $\mathcal{C}\left([0,1];\mathbb{B}_\beta\right)$.

Proof. Let $m\in\mathbb{N}$. For $i=1,2$ and $|y_i|\leq m$, let $y_i':=\frac{1}{m}y_i$ and $\tilde{v}_t^\epsilon(y_i'):=v_t^\epsilon(y_i)$. By the hypothesis,

$$\begin{aligned}\mathbb{E}\left|\tilde{v}_{t_1}^\epsilon(y_1')-\tilde{v}_{t_2}^\epsilon(y_2')\right|^n &= \mathbb{E}\left|u_{t_1}^\epsilon(my_1')-u_{t_2}^\epsilon(my_2')\right|^n\\ &\leq Ke^{n\beta_1(|y_1|\vee|y_2|)}\left(m\left|y_1'-y_2'\right|+|t_1-t_2|\right)^{2+q}\\ &\leq Km^{2+q}e^{n\beta_1 m}\left(\left|y_1'-y_2'\right|+|t_1-t_2|\right)^{2+q}. \end{aligned}\tag{3.17}$$

By Kolmogorov's criterion (cf. Appendix), there exists a random variable Y_m such that $\mathbb{E}Y_m^n\leq Km^{2+q}e^{n\beta_1 m}$ and

$$\left|\tilde{v}_{t_1}^\epsilon(y_1')-\tilde{v}_{t_2}^\epsilon(y_2')\right|\leq KY_m\left(\left|y_1'-y_2'\right|+|t_1-t_2|\right)^{q/n}$$

therefore,

$$\left|v_{t_1}^\epsilon(y_1)-v_{t_2}^\epsilon(y_2)\right|\leq KY_m\left(|y_1-y_2|+|t_1-t_2|\right)^{q/n}. \tag{3.18}$$

Let $Y:=\sup_m\{Y_m e^{-\beta m}\}$. Then,

$$\begin{aligned}\mathbb{E}Y^n &\leq \mathbb{E}\sum_m Y_m^n e^{-\beta mn}\\ &= \sum_m \mathbb{E}Y_m^n e^{-\beta mn}\\ &\leq \sum_m Km^{2+q}e^{-(\beta-\beta_1)mn}<\infty.\end{aligned}\tag{3.19}$$

Thus, Y is a finite random variable, and (3.18) implies

$$\left|v_{t_1}^\epsilon(y_1)-v_{t_2}^\epsilon(y_2)\right|\leq KYe^{\beta m}\left(|y_1-y_2|+|t_1-t_2|\right)^{q/n}. \tag{3.20}$$

This proves (4.19) with $\alpha=\frac{q}{n}$. Taking $y_1=y_2=y$ and m such that $m-1<|y|\leq m$ in (3.20), we have

$$\|v_{t_1}^\epsilon-v_{t_2}^\epsilon\|_\beta\leq KYe^{\beta m}|t_1-t_2|^\alpha \tag{3.21}$$

and hence $v_\cdot^\epsilon\in\mathcal{C}\left([0,1];\mathbb{B}_\beta\right)$ a.s.

Now, we suppose there exists $(t_0, y_0) \in [0,1] \times \mathbb{R}$ such that

$$\sup_{\epsilon>0} \mathbb{E}|v^\epsilon_{t_0}(y_0)|^n < \infty.$$

Note that (3.20) remains true with β replaced by $\beta_2 \in (\beta_1, \beta)$. To simplify the presentation, we choose $t_0 = y_0 = 0$. Taking $t_1 = t$, $y_1 = y$ and $t_2 = y_2 = 0$ in (3.20), gives

$$|v^\epsilon_t(y)| \leq |v^\epsilon_0(0)| + Y e^{\beta_2 y}(|y| + |t|)^\alpha$$

where $\{y\}$ is the least integer greater than $|y|$. Suppose that $|y| \leq m$. Let $\beta_3 \in (\beta_2, \beta)$ be fixed. Then,

$$\begin{aligned} e^{-\beta_3|y|}|v^\epsilon_t(y)| &\leq e^{-\beta_3|y|}|v^\epsilon_0(0)| + Y e^{-(\beta_3-\beta_2)y}\,(|y| + |t|)^\alpha \\ &\leq K(e^{-\beta_0|y|}|v^\epsilon_0(0)| + Y) \end{aligned} \tag{3.22}$$

for a suitable constant K. Inequality (4.20) then follows easily.

Now we achieve that tightness of the random fields. For any $\delta > 0$, let $L > 0$ be such that

$$\sup_{\epsilon>0} \left\{ P\left(K\left(e^{-\beta_0|y|}|v^\epsilon_0(0)| + Y\right) > L\right) + P\left(YKe^{\beta m} > L\right)\right\} < \delta \tag{3.23}$$

Let K_L be the set of all functions f in $\mathbb{B}_\beta$ such that for all $m \geq 1$,

$$\sup_{x\in\mathbb{R}} |f(x)|e^{-\beta_3|x|} \leq L$$

and

$$\sup_{|x_1|,|x_2|\leq m} \frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|^\alpha} \leq Le^{\beta m}$$

Then K_L is a compact subset of $\mathbb{B}_\beta$ by Arzela-Ascolli theorem (cf Appendix). Combining (3.20), (3.22) and (3.23), we have

$$P\left(\exists t \in [0,1], v^\epsilon_t \notin K_L\right) < \delta \tag{3.24}$$

let

$$\mathbb{K}_L = \left\{ v \in \mathcal{C}_\beta : v_t \in K_L, \forall t \in [0,1], \sup_{t\neq s} \frac{\|v_t - v_s\|_\beta}{|t-s|^\alpha} \leq L \right\}$$

then $\mathbb{K}_L$ is a compact subset of $\mathcal{C}([0,1];\mathbb{B}_\beta)$ again by Arzela-Ascolli theorem. By (3.21) and (3.24) we obtain,

$$P(v^\epsilon \notin \mathbb{K}_L) < \delta$$

which proves tightness by applying the Prohorov theorem (see Appendix) and noting that since our setting is in a metric space, compactness implies sequentially compactness. $\square$

Since Xiong [69] proved the uniqueness of a strong solution to the general SPDE (3.8), then its mild solution (cf. Appendix) is equivalent to (3.8). To obtain a mild solution of (3.8), we consider the Brownian semigroup $\{P_t\}$ given by $P_t f(y) = \int_{\mathbb{R}} p_t(x-y)f(x)dx$ with heat kernel, $p_t(x) = \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{x^2}{2t}\right)$. For the simplicity of notation, we take $\epsilon = 1$ and denote $u_t^\epsilon(y)$ by $u_t(y)$. Then (3.8) can be written in the following mild form,

$$u_t(y) = \int_{\mathbb{R}} p_t(y-x)F(x)dx + \int_0^t \int_U \int_{\mathbb{R}} p_{t-s}(y-x)G(a,x,u_s(x))dxW(dads) \tag{3.25}$$

We apply the following lemma, the proof of which is very similar to the proof of Lemma 2.3 in [69] so is omitted.

Lemma 3.2.2. *For any $n \geq 2$ and $\beta_1 \in (\beta_0, \beta)$ we have,*

$$M := \mathbb{E}\left(\sup_{0\leq s\leq 1}\int_{\mathbb{R}} |u_s(x)|^2 e^{-2\beta_1|x|}dx\right)^n < \infty \tag{3.26}$$

$\square$

Theorem 3.2.1. *For any $\alpha \in (0, \frac{1}{2})$, and $F \in \mathbb{B}_{\alpha,\beta_0}$, there exists a measurable map, $g^\epsilon : \mathbb{B}_{\alpha,\beta_0} \times \mathcal{S} \to \mathcal{C}([0,1];\mathbb{B}_\beta)$ such that $u^\epsilon = g^\epsilon(F, \sqrt{\epsilon}B)$ is the unique mild solution of (3.8) where B is a Brownian sheet.*

Proof. The uniqueness and existence of a strong solution and so of mild solution was established by Xiong [69]. We proceed by showing the mild solution (3.25) takes values in $\mathcal{C}([0,1];\mathbb{B}_\beta)$. For simplicity of the presentation, we refer to the first term on the RHS of (3.25) by $u_t^0(y)$, and the second term by $v_t(y)$. For both terms, our goal is to apply Lemma 3.2.1 to achieve the desired result.

We begin by proving $u^0_\cdot$ is an element of $\mathcal{C}([0,1];\mathbb{B}_\beta)$ by considering two cases. Let B_t be a Brownian motion. Recall,

$$u_t^0(y) = \int_{\mathbb{R}} p_t(y-x)F(x)dx$$

so we have, by change of variables,

$$\begin{aligned}
\mathbb{E}F(y-B_t) &= \int_{\mathbb{R}} F(y-x)p_t(x)dx \\
&= -\int_{\infty}^{-\infty} F(z)p_t(y-z)dz \\
&= \int_{-\infty}^{\infty} F(z)p_t(y-z)dz = u_t^0(y)
\end{aligned}$$

Therefore,

$$u_t^0(y) = \mathbb{E}F(y-B_t)$$

Let $\delta > 0$ be such that $(1+\delta)\beta_0 \le \beta$, $y \in \mathbb{R}$ be fixed such that $m-1 < |y| \le m$ and let $0 < t_1 \le t_2 < 1$. Using above new expression for $u_t^0(y)$ and applying condition (3.5) of $\mathbb{B}_{\alpha,\beta_0}$ we have,

$$\begin{aligned}
&\left|u_{t_1}^0(y) - u_{t_2}^0(y)\right| \\
\le\ & \mathbb{E}\left|F(y-B_{t_1}) - F(y-B_{t_2})\right| \\
=\ & \sum_{j_1,j_2} \mathbb{E}\left|F(y-B_{t_1}) - F(y-B_{t_2})\right| 1_{j_1 m\delta \le \left|B_{t_1}\right| \le (j_1+1)m\delta} 1_{j_2 m\delta \le \left|B_{t_2}\right| \le (j_2+1)m\delta} \\
\le\ & \sum_{j_1,j_2} K e^{(m+(j_1\vee j_2+1)m\delta)\beta_0} \mathbb{E}\left(\left|B_{t_1} - B_{t_2}\right|^\alpha 1_{|B_{t_1}|\ge j_1 m\delta} 1_{|B_{t_2}|\ge j_2 m\delta}\right) \\
\le\ & \sum_{j_1,j_2} K e^{(m+(j_1\vee j_2+1)m\delta)\beta_0} \left[\mathbb{E}\left|B_{t_1} - B_{t_2}\right|^{2\alpha}\right]^{\frac{1}{2}} P\left(|B_{t_1}| \ge j_1 m\delta, |B_{t_2}| \ge j_2 m\delta\right)^{\frac{1}{2}}
\end{aligned}$$

Note that since $\alpha \in (0,\frac{1}{2})$, then $f(x) = x^\alpha$ is a concave function enabling us to continue with

$$\le \sum_{j_1,j_2} K e^{(m+(j_1\vee j_2+1)m\delta)\beta_0} \left|t_1 - t_2\right|^{\frac{\alpha}{2}} P\left(|B_{t_1}| \ge j_1 m\delta, |B_{t_2}| \ge j_2 m\delta\right)^{\frac{1}{2}}$$

By the independence of the increments we can calculate,

$$P\left(|B_{t_1}| \ge j_1 m\delta, |B_{t_2}| \ge j_2 m\delta\right)^{\frac{1}{2}} = P\left(|B_{t_1}| \ge j_1 m\delta\right)^{\frac{1}{2}} P\left(|B_{t_2}| \ge j_2 m\delta\right)^{\frac{1}{2}}$$

where

$$
\begin{aligned}
P\left(|B_{t_1}| \geq j_1 m\delta\right) &= \int_{j_1 m\delta}^{\infty} |x| p_{t_1}(x) dx \\
&= K \int_{j_1 m\delta}^{\infty} |x| e^{-\frac{x^2}{2t_1}} dx = K e^{-j_1^2 m^2 \delta^2} 2t_1 \leq K e^{-\frac{j_1^2 m^2 \delta^2}{4}}
\end{aligned}
$$

since $t_1 \leq 1$. We can obtain a similar estimate for $P(|B_{t_2}| \geq j_2 m\delta)$. Therefore,

$$
\begin{aligned}
& \left|u_{t_1}^0(y) - u_{t_2}^0(y)\right| \\
\leq & \sum_{j_1, j_2} K e^{(m+(j_1 \vee j_2+1)m\delta)\beta_0} \left|t_1 - t_2\right|^{\frac{\alpha}{2}} e^{-\frac{1}{4} m^2 \delta^2 (j_1^2 + j_2^2)} \\
\leq & K \int_{\mathbb{R}} \int_{\mathbb{R}} K \exp\left((m + (x \vee y + 1)m\delta)\beta_0 - \frac{1}{4} m^2 \delta^2 (x^2 + y^2) \right) dx dy \left|t_1 - t_2\right|^{\frac{\alpha}{2}} \\
\leq & K e^{\beta_0 (1+\delta) m} \left|t_1 - t_2\right|^{\frac{\alpha}{2}} \\
\leq & K e^{m\beta} \left|t_1 - t_2\right|^{\frac{\alpha}{2}}
\end{aligned}
$$

So

$$
\left\| u_{t_1}^0 - u_{t_2}^0 \right\|_{\beta} \leq K \left|t_1 - t_2\right|^{\frac{\alpha}{2}}
$$

Now we fix $t \in [0,1]$ and let $y_1, y_2 \in \mathbb{R}$ be arbitrary such that $|y_1|, |y_2| \leq m$, then

$$
\begin{aligned}
\left|u_t^0(y_1) - u_t^0(y_2)\right| &\leq \mathbb{E}\left|F(y_1 - B_t) - F(y_2 - B_t)\right| \\
&\leq K e^{\beta_0 m} |y_1 - y_2|
\end{aligned}
$$

Therefore condition 3.14 in lemma 3.2.1 is satisfied. As $u_0^0 = F \in \mathbb{B}_{\alpha,\beta_0} \subset \mathbb{B}_\beta$, we have $u_t^0 \in \mathbb{B}_\beta$ and therefore, $u_\cdot^0 \in \mathcal{C}\left([0,1]; \mathbb{B}_\beta\right)$.

Now we turn our attention to second term, $v_\cdot$ and prove the two cases. For case one, let $t \in [0,1]$ be fixed, while $y_1, y_2 \in \mathbb{R}$ are arbitrary numbers such that $|y_i| \leq m$ for $i = 1, 2$ and denote

$$
P_1 := p_{t-s}(y_1 - x) - p_{t-s}(y_2 - x).
$$

Applying Burkholder-Davis-Gundy (cf. Appendix) and Hölder's inequalities and (4.4), we obtain for $n \geq 2$,

$$
\begin{aligned}
& \mathbb{E}\left|v_t(y_1)-v_t(y_2)\right|^n \\
= \; & \mathbb{E}\left|\int_0^t \int_U \int_{\mathbb{R}} P_1 G(a,x,u_s(x))dxW(dads)\right|^n \\
\leq \; & K\mathbb{E}\left|\int_0^t \int_U \left|\int_{\mathbb{R}} P_1 G(a,x,u_s(x))dx\right|^2 \lambda(da)ds\right|^{n/2} \\
\leq \; & K\mathbb{E}\left|\int_0^t \int_U \int_{\mathbb{R}} |P_1|^2 e^{2\beta_1|x|}dx \int_{\mathbb{R}} G(a,x,u_s(x))^2 e^{-2\beta_1|x|}dx\lambda(da)ds\right|^{n/2} \\
\leq \; & K\mathbb{E}\left|\int_0^t J_{t-s}(y_1,y_2)\int_{\mathbb{R}}(1+|u_s(x)|^2)e^{-2\beta_1|x|}dxds\right|^{n/2} \qquad (3.27)
\end{aligned}
$$

where

$$
J_s(y_1,y_2) := \int_{\mathbb{R}} |p_s(y_1-x)-p_s(y_2-x)|^2 e^{2\beta_1|x|}dx
$$

is estimated below by applying inequality (4.4), the fact that for all $a,b>0$,

$$
|e^a - e^b| \leq |a-b| \qquad (3.28)
$$

and using the simplified notation,

$$
P_2 := p_s(y_1-x)-p_s(y_2-x)
$$

$$
\begin{aligned}
J_s(y_1,y_2) \; = \; & \int_{\mathbb{R}} |P_2|^{\alpha} |P_2|^{2-\alpha} e^{2\beta_1|x|}dx \qquad (3.29) \\
\leq \; & \int_{\mathbb{R}} \left|\frac{1}{\sqrt{2\pi s}}\right|^{\alpha} \left|\frac{(y_1-x)^2-(y_2-x)^2}{2s}\right|^{\alpha} |P_2|^{2-\alpha} e^{2\beta_1|x|}dx \\
\leq \; & K\int_{\mathbb{R}} \left|\frac{1}{\sqrt{2\pi s}}\right|^{\alpha} \frac{|y_1-y_2|^{\alpha}|y_1+y_2-2x|^{\alpha}}{(2s)^{\alpha}(2\pi s)^{(2-\alpha)/2}} e^{-\frac{(2-\alpha)(y_1-x)^2}{2s}} e^{2\beta_1|x|}dx \\
& +K\int_{\mathbb{R}} \left|\frac{1}{\sqrt{2\pi s}}\right|^{\alpha} \frac{|y_1-y_2|^{\alpha}|y_1+y_2-2x|^{\alpha}}{(2s)^{\alpha}(2\pi s)^{(2-\alpha)/2}} e^{-\frac{(2-\alpha)(y_2-x)^2}{2s}} e^{2\beta_1|x|}dx \\
\leq \; & K|y_1-y_2|^{\alpha}s^{-(1+\alpha)}\int_{\mathbb{R}} |y_1+y_2-2x|^{\alpha} e^{-\frac{(2-\alpha)(y_1-x)^2}{2s}} e^{2\beta_1|x|}dx \\
& +K|y_1-y_2|^{\alpha}s^{-(1+\alpha)}\int_{\mathbb{R}} |y_1+y_2-2x|^{\alpha} e^{-\frac{(2-\alpha)(y_2-x)^2}{2s}} e^{2\beta_1|x|}dx
\end{aligned}
$$

Since for $i = 1, 2$,

$$
\begin{aligned}
& K|y_1 - y_2|^\alpha s^{-(1+\alpha)} \int_{\mathbb{R}} |y_1 + y_2 - 2x|^\alpha e^{-\frac{(2-\alpha)(y_i - x)^2}{2s}} e^{2\beta_1|x|} dx \\
\leq \quad & K|y_1 - y_2|^\alpha s^{-(1+\alpha)} e^{2\beta_1|y_i|} \int_{\mathbb{R}} |y_1 + y_2 - 2x|^\alpha e^{-\frac{(2-\alpha)(y_i - x)^2}{2s}} e^{2\beta_1|x - y_i|} dx \\
\leq \quad & K|y_1 - y_2|^\alpha s^{-(\frac{1}{2}+\alpha)} e^{2\beta_1|y_i|}
\end{aligned}
$$

therefore,

$$
J_s(y_1, y_2) \leq K e^{2\beta_1(|y_1| \vee |y_2|)} s^{-(\frac{1}{2}+\alpha)} |y_1 - y_2|^\alpha
$$

Before plugging back $J_s(y_1, y_2)$ in (3.29), note that by (3.26)

$$
\begin{aligned}
& \mathbb{E}\left(\int_0^t (t-s)^{-(\frac{1}{2}+\alpha)} \int_{\mathbb{R}} (1 + |u_s(x)|^2) e^{-2\beta_1|x|} dx ds \right)^{n/2} \\
\leq \quad & \mathbb{E}\left(\sup_{0 \leq s \leq 1} \int_{\mathbb{R}} \left(1 + |u_s(x)|^2\right) e^{-2\beta_1|x|} dx \right)^{\frac{n}{2}} \left(\int_0^t (t-s)^{-(\frac{1}{2}+\alpha)} ds \right)^{\frac{n}{2}} \\
\leq \quad & M^{\frac{1}{2}} K \leq K
\end{aligned}
\tag{3.30}
$$

Using this inequality and (3.26) we obtain,

$$
\mathbb{E}|v_t(y_1) - v_t(y_2)|^n \leq K e^{n\beta_1(|y_1| \vee |y_2|)} |y_1 - y_2|^{\frac{\alpha n}{2}}.
$$

Next to prove case two, let $y \in \mathbb{R}$ and choose any $0 \leq t_1 < t_2 \leq 1$. Note that with

$$
P_2 := p_{t_1 - s}(y - x) - p_{t_2 - s}(y - x)
$$

we can deduce for $n \geq 2$,

$$
\begin{aligned}
& \mathbb{E}\left|v_{t_1}(y) - v_{t_2}(y)\right|^n \qquad (3.31) \\
= \quad & \left| \int_0^{t_1} \int_U \int_{\mathbb{R}} p_{t_1 - s}(y - x) G(a, x, u_s(x)) dx W(dads) \right. \\
& \left. - \int_0^{t_2} \int_U \int_{\mathbb{R}} p_{t_2 - s}(y - x) G(a, x, u_s(x)) dx W(dads) \right|^n \\
\leq \quad & K\mathbb{E} \left| \int_0^{t_1} \int_U \int_{\mathbb{R}} P_2 G(a, x, u_s(x)) dx W(dads) \right|^n \\
& + K\mathbb{E} \left| \int_{t_1}^{t_2} \int_U \int_{\mathbb{R}} p_{t_2 - s}(y - x) G(a, x, u_s(x)) dx W(dads) \right|^n \\
\leq \quad & K\mathbb{E} \left| \int_0^{t_1} \int_{\mathbb{R}} P_2^2 e^{2\beta_1|x|} dx \int_{\mathbb{R}} (1 + |u_s(x)|^2) e^{-2\beta_1|x|} dx ds \right|^{\frac{n}{2}} \\
& + K\mathbb{E} \left| \int_{t_1}^{t_2} \int_{\mathbb{R}} p_{t_2 - s}^2(y - x) e^{2\beta_1|x|} dx \int_{\mathbb{R}} (1 + |u_s(x)|^2) e^{-2\beta_1|x|} dx ds \right|^{\frac{n}{2}}
\end{aligned}
$$

We perform the following calculations to arrive at the desired estimate for $\int_{\mathbb{R}} P_2^2 e^{2\beta_1|x|}dx$. Let

$$\begin{aligned} I_s(t_1,t_2) &:= \int_{\mathbb{R}} P_2^2 e^{2\beta_1|x|}dx \\ &= \int_{\mathbb{R}} P_2^{\alpha} P_2^{2-\alpha} e^{2\beta_1|x|}dx = I_s^1(t_1,t_2) + I_s^2(t_1,t_2), \\ I_s^i(t_1,t_2) &:= \int_{\mathbb{R}} |p_{t_1-s}(y-x) - p_{t_2-s}(y-x)|^{\alpha}\, p_{t_i-s}(y-x)^{2-\alpha} e^{2\beta_1|x|}dx \end{aligned}$$

for $i = 1,2$. Since,

$$\begin{aligned} &|p_{t_1-s}(y-x) - p_{t_2-s}(y-x)|^{\alpha} \\ = &\left| p_{t_1-s}(y-x) - \frac{1}{\sqrt{t_2-s}} e^{-\frac{|y-x|^2}{2(t_1-s)}} + \frac{1}{\sqrt{t_2-s}} e^{-\frac{|y-x|^2}{2(t_1-s)}} - \frac{1}{\sqrt{t_2-s}} e^{-\frac{|y-x|^2}{2(t_2-s)}} \right|^{\alpha} \end{aligned}$$

then term $I_s^1(t_1,t_2)$ can be approximated by $K\left(I_s^{11}(t_1,t_2) + I_s^{12}(t_1,t_2)\right)$ where

$$I_s^{11}(t_1,t_2) := \int_{\mathbb{R}} \left| \frac{1}{\sqrt{t_1-s}} - \frac{1}{\sqrt{t_2-s}} \right|^{\alpha} p_{t_1-s}(y-x)^{2-\alpha} e^{2\beta_1|x|}dx$$

and by using (3.28),

$$I_s^{12}(t_1,t_2) := \int_{\mathbb{R}} \left| \frac{1}{\sqrt{t_2-s}} \left| \frac{1}{t_1-s} - \frac{1}{t_2-s} \right| (y-x)^2 \right|^{\alpha} p_{t_1-s}(y-x)^{2-\alpha} e^{2\beta_1|x|}dx$$

Now we continue with

$$\begin{aligned} I_s^{11}(t_1,t_2) &\leq K \int_{\mathbb{R}} \left| \frac{t_2-t_1}{\sqrt{t_1-s}(t_2-s)} \right|^{\alpha} \frac{e^{2\beta_1|x|}}{\sqrt{t_1-s}^{1-\alpha}} p_{(t_1-s)/(2-\alpha)}(y-x)dx \\ &\leq K \frac{|t_1-t_2|^{\alpha}}{\sqrt{t_1-s}(t_2-s)^{\alpha}} e^{2\beta_1|y|} \end{aligned}$$

and

$$\begin{aligned} I_s^{12}(t_1,t_2) &\leq K \int_{\mathbb{R}} \frac{|t_1-t_2|^{\alpha}(y-x)^{2\alpha}}{(t_2-s)^{\frac{3\alpha}{2}}(t_1-s)^{1+\frac{\alpha}{2}}} p_{(t_1-s)/(2-\alpha)}(y-x) e^{2\beta_1|x|}dx \\ &\leq K \frac{|t_1-t_2|^{\alpha}}{(t_2-s)^{\frac{3\alpha}{2}}(t_1-s)^{1+\frac{\alpha}{2}}} e^{2\beta_1|y|} \end{aligned}$$

Recall $0 \leq t_1 < t_2 \leq 1$ so for $\alpha \in (0, \frac{1}{2})$,

$$\begin{aligned} \int_0^{t_1} I_s^{11}(t_1,t_2)ds &\leq K e^{2\beta_1|y|}|t_1-t_2|^{\alpha} \int_0^{t_1} (t_1-s)^{-(\frac{1}{2}+\alpha)}ds \\ &\leq K e^{2\beta_1|y|}|t_1-t_2|^{\alpha} \end{aligned}$$

and

$$\begin{aligned}\int_0^{t_1} I_s^{12}(t_1,t_2)ds &\le Ke^{2\beta_1|y|}|t_1-t_2|^{\alpha}\int_0^{t_1}(t_1-s)^{-(\frac{1}{2}+2\alpha)}ds\\ &\le Ke^{2\beta_1|y|}|t_1-t_2|^{\alpha},\end{aligned}$$

where we used the fact that $t_2 - s > t_1 - s$. Terms I^{21} and I^{22} are defined and estimated similarly. Therefore,

$$\begin{aligned}&K\left(\int_0^{t_1} I_s(t_1,t_2)ds\right)^{\frac{n}{2}}\\ \le\ &K\left(\int_0^{t_1}\left(I_s^{11}(t_1,t_2)+I_s^{12}(t_1,t_2)+I_s^{21}(t_1,t_2)+I_s^{22}(t_1,t_2)\right)ds\right)^{n/2}\\ \le\ &Ke^{n\beta_1|y|}|t_1-t_2|^{\frac{\alpha n}{2}}\end{aligned}$$

which leads to

$$K\mathbb{E}\left|\int_0^{t_1} I_s(t_1,t_2)\int_{\mathbb{R}}(1+|u_s(x)|^2)e^{-2\beta_1|x|}dxds\right|^{\frac{n}{2}} \le Ke^{n\beta_1|y|}|t_1-t_2|^{\frac{\alpha n}{2}}.$$

For the second term of (3.31), since $\alpha \in (0,1)$ we obtain,

$$\begin{aligned}\int_{t_1}^{t_2}\int_{\mathbb{R}} p_{t_2-s}^2(y-x)e^{2\beta_1|x|}dxds &\le K\int_{t_1}^{t_2}\int_{\mathbb{R}}\frac{e^{2\beta_1|x|}}{\sqrt{t_2-s}}p_{\frac{1}{2}(t_2-s)}(y-x)dxds\\ &\le Ke^{2\beta_1|y|}\int_{t_1}^{t_2}\frac{ds}{\sqrt{t_2-s}}\\ &\le K|t_1-t_2|^{\frac{\alpha}{2}}e^{2\beta_1|y|}\end{aligned}$$

Thus, we see that

$$\begin{aligned}&K\mathbb{E}\left|\int_{t_1}^{t_2}\int_{\mathbb{R}} p_{t_2-s}^2(y-x)e^{2\beta_1|x|}dx\int_{\mathbb{R}}(1+|u_s(x)|^2)e^{-2\beta_1|x|}dxds\right|^{\frac{n}{2}}\\ \le\ &Ke^{n\beta_1|y|}|t_1-t_2|^{\frac{n\alpha}{4}}\end{aligned}$$

Since both terms of (3.25) satisfy estimate (3.14) given in Lemma 3.2.1, then by the conclusion of this lemma, the mild solution (3.25) is in space $\mathcal{C}([0,1];\mathbb{B}_\beta)$ and so letting g^ϵ represent (3.25) establishes this theorem. □

Now we derive the LDP for SPDE (3.8) by using the powerful technique developed by Budhiraja *et al* [7]. For completeness we include their result below, with the following changes in their notation: $X^{\epsilon,x} := u_t^\epsilon$, $\mathcal{E}_0 := \mathbb{B}_{\alpha,\beta_0}$, $\mathcal{E} := \mathcal{C}([0,1];\mathbb{B}_\beta)$, $x := F$ and $u(s) := k_s$.

Assumption 1. *There exists a measurable map $g^0 : \mathbb{B}_{\alpha,\beta_0} \times \mathcal{S} \to \mathcal{C}([0,1]; \mathbb{B}_{\alpha,\beta})$ such that the following hold:*

1. For all $M < \infty$ and compact set $K \subseteq \mathbb{B}_{\alpha,\beta_0}$, the set

$$\Gamma_{M,K} = \left\{ g^0 \left(F, \int_0^{\cdot} k_s ds da \right) : k_s \in \mathcal{S}^N(\ell_2), F \in K \right\} \tag{3.32}$$

is a compact subset of $\mathcal{C}([0,1]; \mathbb{B}_{\alpha,\beta})$ where

$$\mathcal{S}^N(\ell_2) := \left\{ k_s \in L^2([0,1] : \ell_2) : \int_0^1 \|k_s\|_{\ell_2}^2 ds \leq N \right\}.$$

2. If $M < \infty$ and for families, $\{k^\epsilon\} \subset \mathcal{S}^N(\ell_2), \{F^\epsilon\} \subset \mathbb{B}_{\alpha,\beta_0}$, $k^\epsilon \xrightarrow{d} k, F^\epsilon \to F$ as $\epsilon \to 0$ then

$$g^\epsilon \left(F^\epsilon, \sqrt{\epsilon}\beta + \int_0^{\cdot} k_s^\epsilon(a) ds \right) \xrightarrow{d} g^0 \left(F, \int_0^{\cdot} k_s(a) ds \right) \tag{3.33}$$

Theorem 3.2.2. *Suppose $u_t^\epsilon := g^\epsilon(F, \sqrt{\epsilon}\beta)$, where $\beta := \{\beta_j\}$ is an infinite sequence of independent standard real Brownian motions. For $F \in \mathbb{B}_{\alpha,\beta_0}$, and $f \in \mathcal{C}([0,1]; \mathbb{B}_\beta)$, let*

$$I_F(f) = \inf_{\{k_s \in L^2([0,1]:\ell_2) : f = g^0\left(F, \int_0^{\cdot} k_s ds\right)\}} \left\{ \frac{1}{2} \int_0^1 \|k_s\|_{\ell_2}^2 ds \right\}$$

and assume that for all $f \in \mathcal{C}([0,1]; \mathbb{B}_\beta)$, $I_F(f)$ is a lower semicontinuous map from $\mathbb{B}_{\alpha,\beta_0}$ to $[0, \infty]$. If the assumption above holds, then for all $F \in \mathbb{B}_{\alpha,\beta_0}$, $f \mapsto I_F(f)$ is a rate function on $\mathcal{C}([0,1]; \mathbb{B}_\beta)$, the family $\{I_F(.) : F \in \mathbb{B}_{\alpha,\beta_0}\}$ has compact level sets on compacts and the sequence, $\{u_t^\epsilon\}$, satisfies the Laplace Principle on $\mathcal{C}([0,1]; \mathbb{B}_\beta)$ with rate function I_F uniformly on compact subsets of $\mathbb{B}_{\alpha,\beta_0}$.

We remark here that the authors in [7] also established the LDP for $u_t^\epsilon := g^\epsilon(F, \sqrt{\epsilon}B)$ where B is a Brownian sheet (see Theorem 7 in [7]). This form of u_t^ϵ better represents our solutions; however, their Brownian sheet is a measure on $[0,T] \times \mathcal{O}$ where $\mathcal{O}$ is a bounded open set in $\mathbb{R}^d$. In our case $\mathcal{O}$ is the set U and for example in the case of SBM, U is $\mathbb{R}$ and not bounded. Therefore, we use Theorem 3.2.2 above, instead, and aim to modify our solutions to match its setup.

Since $\{u_t^\epsilon\}$ in above theorem is in the form $g^\epsilon(F, \sqrt{\epsilon}\beta)$ where β is an infinite sequence of independent standard real Brownian motions, then we replace the white noise in SPDE (3.8) by a sequence of independent Brownian motions. Let

$$B_t^j := \int_0^t \int_U \phi_j(a) W(dads), \qquad j = 1, 2, \cdots . \tag{3.34}$$

where $\{\phi_j\}_{j\geq 1}$ is a complete orthonormal system (CONS) of $L^2(U, \mathcal{U}, \lambda)$. Since this sequence has quadratic variation,

$$[B_t^j, B_t^j] = \int_0^t \int_U \phi_j^2(a) dads = t$$

and $\int_0^t \int_U \mathbb{E}\left(\phi_j(a)\right)^2 dads = t \in [0, 1]$ implying that B_t^j is a continuous martingale, then by Lévy's characterization of Brownian motions (see Appendix), $\{B_t^j\}$ is a sequence of standard Brownian motions. Another way to see $\{B_t^j\}$ is a Brownian motion is to notice that $\{B_t^j\}$ is a Gaussian process with zero mean and covariance t. Using the fact that Brownian motion is a Gaussian process and here for each $i, j \in \mathbb{N}$, B_t^i, B_t^j are uncorrelated, we have that $\{B_t^j\}$ is a sequence of independent standard Brownian motions. We next apply the following theorem, the proof of which can be found in Kallianpur/Xiong [34] Theorem 1.1.10.

Theorem 3.2.3. *If X is a separable Hilbert space, then for any CONS $\{\phi_j\}$, and for all $x \in X$, the following equalities hold,*

$$x = \sum_{j=1}^{\infty} < x, \phi_j > \phi_j \tag{3.35}$$

$$\|x\|^2 = \sum_j < x, \phi_j >^2 \tag{3.36}$$

$$< x, y > = \sum_j < x, \phi_j >< y, \phi_j > \tag{3.37}$$

□

Using (3.35) we obtain,

$$\begin{aligned} & \sqrt{\epsilon} \int_0^t \int_U G(a, y, u_s^\epsilon(y)) W(dads) \\ = \ & \sqrt{\epsilon} \sum_j \int_0^t \int_U \left\langle G(a, y, u_s^\epsilon(y)), \phi_j(a) \right\rangle_{L^2} \phi_j(a) W(dads) \\ = \ & \sqrt{\epsilon} \sum_j \int_0^t G_j(a, y, u_s^\epsilon) dB_s^j \end{aligned}$$

where

$$G_j(a,y,u) := \langle G(a,y,u), \phi_j(a)\rangle_{L^2} = \int_U G(a,y,u)\phi_j(a)\lambda(da), \;\; j = 1,2,\cdots$$

Therefore, our SPDE (3.8) can be written as

$$u_t^\epsilon(y) = F(y) + \sqrt{\epsilon}\sum_j \int_0^t G_j(a,y,u_s^\epsilon(y))dB_s^j + \int_0^t \frac{1}{2}\Delta u_s^\epsilon(y)ds, \tag{3.38}$$

We also need to consider the controlled PDE of (3.8) with noise replaced by the control as was given in (3.11). Recall for any $h \in L^2([0,1]\times U, ds\lambda(da))$, this version has the following deterministic form,

$$u_t(y) = F(y) + \int_0^t \int_U G(a,y,u_s(y))h_s(a)\lambda(da)ds + \int_0^t \frac{1}{2}\Delta u_s(y)ds. \tag{3.39}$$

Applying (3.35) we have,

$$\int_0^t \int_U G(a,y,u_s(y))h_s(a)\lambda(da)ds = \sum_j \int_0^t \int_U G(a,y,u_s(y))\phi_j(a)k_s^j\lambda(da)ds$$

where

$$k_s^j := \langle h_s(a), \phi_j(a)\rangle_{L^2} = \int_U h_s(a)\phi_j(a)\lambda(da) \tag{3.40}$$

To apply Theorem 3.2.2 we need the uniqueness of mild solutions to (3.39). We cannot prove the uniqueness of strong solutions by using the classic results since the middle term is non-Lipschitz continuous. If we denote the derivative of u_t^ϵ with respect to t as u_t then we can write this equation in the form of a heat equation,

$$\begin{aligned} u_t - \frac{1}{2}\Delta u &= \int_U G(a,y,u_t(y))h_t(a)\lambda(da) \\ u(0) &= F(y) \end{aligned}$$

where the nonhomogenous term is Hölder continuous with order $\frac{1}{2}$. But using the Banach's fixed point theorem, classic texts such as [21] require this term to be Lipschitz continuous to achieve the uniqueness of weak solutions (cf. Section 9.2.1 in [21]). As for uniqueness of mild solutions, recently there has been extensive investigation on the existence and uniqueness of solutions with non-Lipschitz continuity of coefficients. Many papers have considered the reaction-diffusion form,

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dB_t$$

with initial condition $X_0 \in L^2(\mathbb{R}^d)$ and B a d-dimensional Brownian motion. Notice that we are interested in proving the uniqueness of mild solution. That is uniqueness of solutions to

$$u_t(y) = P_t F(y) + \int_0^t \int_U P_{t-s} G(a, y, u_s(y)) h_s(a) \lambda(da) ds$$

with P_t denoting the Brownian semigroup. Therefore, to apply their results we can denote

$$b(t, u_t) := \int_U P_{t-s} G(a, y, u_s(y)) h_s(a) \lambda(da)$$

and let $\sigma(t, u_t) = 0$. The authors of these papers such as [60, 62, 63, 70] were able to achieve the existence and uniqueness of solutions by applying the Picard's iteration method, also referred to as the successive approximation. The key assumption on the coefficients is as follows: suppose there exists a nonnegative, monotone, nondecreasing, concave function $\phi(t, u)$, continuous in u for each fixed $t \in [0, T]$ and is locally integrable in t for each fixed u such that

$$|b(t, X) - b(t, Y)|^2 + |\sigma(t, X) - \sigma(t, Y)|^2 \leq \phi(t, |X - Y|^2) \tag{3.41}$$

for all $t \in [t_0, T]$ and X, Y in the space of the problem of study. Further, if there exists a nonnegative, continuous function $z(t)$ such that

$$\begin{aligned} z(t_0) &= 0 \\ z(t) &\leq A \int_{t_0}^T \phi(s, z(s)) ds \end{aligned}$$

for all $t \in [t_0, \tilde{T}]$, where $\tilde{T} \in (t_0, T]$ and A is a constant, then $z(t) = 0$ for all $t \in [t_0, \tilde{T}]$.

If $\phi(t, u) = \lambda(t)\alpha(u)$ for $t \geq 0$, $u \geq 0$ where $\lambda(t) \geq 0$ is locally integrable and $\alpha(u)$ is nonnegative, nondecreasing, function with $\alpha(0) = 0$ and satisfies the Osgood's condition given as $\int_{0+} 1/\alpha(u) du = \infty$, then the z-condition is satisfied. This fact can be shown by using the Bihari's inequality (see [2]).

In our case $\phi(x) = \sqrt{x}$ and thus the Osgood's condition is not satisfied. Therefore, their results cannot be applied and instead of proving the uniqueness of mild solutions to the controlled PDE, we modify the topology of the state space, $\mathcal{C}([0, 1]; \mathbb{B}_\beta)$. We do so by introducing the following equivalent classes.

Definition 3.2.1. *We define a relation $\sim$ between $u, v \in \mathcal{C}([0,1];\mathbb{B}_\beta)$ by there exists an $h \in L^2([0,1] \times U, ds\lambda(da))$ such that both u, v are solutions to (3.39). Then this relation is an equivalence relation if we assume the transitive property. If u is not a solution to equation (3.39) for a suitable h, then u belongs to the equivalent class consisting of itself only.*

Denote the equivalence relation described above as $\sim_1$. Then the quotient space, $\mathcal{C}([0,1];\mathbb{B}_\beta) / \sim_1$ is a pseudo-metric space. For $u \in \mathcal{C}([0,1];\mathbb{B}_\beta)$ let $[u]_1 = \{v \in \mathcal{C}([0,1];\mathbb{B}_\beta) : u \sim_1 v\}$ be the equivalence class of u. Note that elements of the quotient space $\mathcal{C}([0,1];\mathbb{B}_\beta) / \sim_1$ are equivalence classes of relation $\sim_1$. Furthermore, there is a map from the space $\mathcal{C}([0,1];\mathbb{B}_\beta)$ to its quotient space $\mathcal{C}([0,1];\mathbb{B}_\beta) / \sim_1$ called the natural projection map taking an element $u \in \mathcal{C}([0,1];\mathbb{B}_\beta)$ to the equivalence class in the quotient space containing it. For a quick background on general quotient spaces and their pseudo-metric see Appendix. For our case, recall the norm of $\mathbb{B}_\beta$,

$$\|f\|_\beta = \sup_{x \in \mathbb{R}} e^{-\beta|x|}|f(x)|$$

Let $\|.\|_{\mathcal{C}_\beta}$ and $d_{\mathcal{C}_\beta}$ be the norm and its corresponding metric for $\mathcal{C}([0,1];\mathbb{B}_\beta)$ respectively. Then the pseudo-metric $\tilde{d}$ of the quotient space $\mathcal{C}([0,1];\mathbb{B}_\beta) / \sim_1$ is defined as follows: for any $[u]_1, [v]_1 \in \mathcal{C}([0,1];\mathbb{B}_\beta) / \sim_1$,

$$\tilde{d}([u]_1, [v]_1) = \inf \sum_{i=1}^{n} d_{\mathcal{C}_\beta}(p_i, q_i)$$

where the infimum is taken over all finite sequences $(p_1, ..., p_n)$ and $(q_1, ..., q_n)$ such that $p_1 \sim_1 u, q_n \sim_1 v$ and $q_i \sim_1 p_{i+1}$ for $i = 1, 2, ..., n-1$.

The classic results on LDP including Theorem 3.2.2 offered by [7], which we are using, are studied in Polish spaces (complete, separable, metric spaces). To convert our pseudo-metric space to a metric space so that it would become a Polish space, we use the usual technique of introducing a second equivalence relation, $\sim_2$. This new equivalence relation is defined as $[u]_1 \sim_2 [v]_1$ if $\tilde{d}([u]_1, [v]_1) = 0$. Then ρ given as

$$\rho([[u]_1]_2, [[v]_1]_2) = \tilde{d}([u]_1, [v]_1)$$

is a metric on $(\mathcal{C}([0,1];\mathbb{B}_\beta) / \sim_1) / \sim_2$. This is indeed a metric space since if $\rho([[u]_1]_2, [[v]_1]_2) = 0$ then $\tilde{d}([u]_1, [v]_1) = 0$ which by the definition of relation $\sim_2$ implies, $[u]_1 \sim_2 [v]_1$ so $[u]_1$ and

$[v]_1$ are in the same class and $[[u]_1]_2 = [[v]_1]_2$. The other properties of metric follow from $\tilde{d}$ being a pseudometric. To avoid confusion in notation, we denote the double quotient space, $(\mathcal{C}([0,1];\mathbb{B}_\beta)/\sim_1)/\sim_2$ as $\tilde{\tilde{\mathcal{C}}}([0,1];\mathbb{B}_\beta)$. Note that $\mathcal{C}([0,1];\mathbb{B}_\beta)$ has a stronger topology than its double quotient space $\tilde{\tilde{\mathcal{C}}}([0,1];\mathbb{B}_\beta)$. Therefore, the tightness and the regularity of solutions that we have already shown in space $\mathcal{C}([0,1];\mathbb{B}_\beta)$ automatically hold in double quotient space $\tilde{\tilde{\mathcal{C}}}([0,1];\mathbb{B}_\beta)$ as well.

We establish the LDP of $u^\epsilon_\cdot$ in the double quotient space $\tilde{\tilde{\mathcal{C}}}([0,1];\mathbb{B}_\beta)$ under the equivalence relations $\sim_1$ and $\sim_2$ given above. Note that when $h=0$, equation (3.39) has a unique solution, $u^0_t(y)$. Therefore, this modification of topology does not affect the exponential rate of decay in LDP.

Let $\gamma_0 : L^2([0,1]\times U, ds\lambda(da)) \to \mathcal{C}([0,1];\mathbb{B}_\beta)/\sim_1$ be a map whose domain consists of all h such that (3.39) has a solution and let $q : \mathcal{C}([0,1];\mathbb{B}_\beta)/\sim_1 \to \tilde{\tilde{\mathcal{C}}}([0,1];\mathbb{B}_\beta)$ be the natural projection mapping to the second quotient space. Then $\gamma := q \circ \gamma_0$ maps $h \in L^2([0,1]\times U; ds\lambda(da))$ to the equivalence class of the solution based on that h in the double quotient space $\tilde{\tilde{\mathcal{C}}}([0,1];\mathbb{B}_\beta)$. We denote this equivalence class of the solution as $u = \gamma(h)$. Furthermore, let g^ϵ be the map given in Theorem 3.2.1. Define a map ζ from $k \in \mathcal{S}^N(\ell_2)$ to $h = \zeta(k) \in L^2([0,1]\times U)$ as follows:

$$h_s(a) = \sum_j k^j_s \phi_j(a).$$

Let $g^0 : \mathbb{B}_{\alpha,\beta_0} \times \mathcal{S}^N(\ell_2) \to \mathcal{C}([0,1] : \mathbb{B}_{\alpha,\beta})$ be given by,

$$g^0\left(F, \int_0^\cdot k_s ds\right) = \gamma\left(F, \zeta\left(k\right)\right). \tag{3.42}$$

Now to obtain the LDP, it is sufficient to verify Assumption 1. Suppose $\{k^\epsilon\}$ is a family of random variables taking values in $\mathcal{S}^N(\ell_2)$ such that $k^\epsilon \to k$ in distribution and $F^\epsilon \to F$ as $\epsilon \to 0$. Denote the solution to

$$\begin{aligned} u^{\theta,\epsilon}_t(y) &= \int_{\mathbb{R}} p_t(y-x)F^\epsilon(x)dx + \theta\sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x)G_j(a,x,u^\epsilon_s(x))dxdB^j_s \\ &\quad + \sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x)G_j(a,x,u^\epsilon_s(x))k^{\epsilon,j}_s dxds \end{aligned} \tag{3.43}$$

as $u_t^{\theta,\epsilon}(y)$.

In [42], Mitoma applied the fact that if

$$\rho(x) = \begin{cases} C\exp\left(\frac{-1}{1-|x|^2}\right) & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

where C is determined by $\int_{\mathbb{R}} \rho(x)dx = 1$, then $g(x) = \int_{\mathbb{R}} e^{-|y|}\rho(x-y)dy$ satisfies

$$Ke^{-|x|} \leq g^{(n)}(x) \leq Ke^{-|x|}.$$

where $g_\beta^{(n)}$ denotes the nth derivative of function g. Therefore, in (3.2), we may replace $e^{-\beta|x|}$ by $J_\beta(x) := \int_{\mathbb{R}} e^{-\beta_1|y|}\rho(x-y)dy$ which satisfies,

$$Ke^{-\beta_1|y|} \leq J^{(n)}(y) \leq Ke^{-\beta_1|y|}$$

Theorem 3.2.4. *Suppose $F \in \mathbb{B}_{\alpha,\beta_0}$, then the family $\{u^\epsilon\}$ satisfies the LDP in $\tilde{\tilde{\mathcal{C}}}([0,1];\mathbb{B}_\beta)$ with rate function,*

$$I(u) = \begin{cases} \frac{1}{2}\inf\left\{\int_0^1\int_U |h_s(a)|^2\lambda(da)ds : u = \gamma(F,h)\right\} & \exists h \ s.t. \ u = \gamma(F,h) \\ \infty & otherwise. \end{cases} \tag{3.44}$$

Proof. We begin by proving the tightness of $\{u^{\theta,\epsilon}\}$ in space $\mathcal{C}([0,1];\mathbb{B}_\beta)$ by verifying the assumptions of Lemma 3.2.1. First let $J(x) = J_{2\beta_1}(x)$ and denote the Hilbert space $L^2(\mathbb{R}, J(x)dx)$ by χ_0. We note that for $f \in \mathcal{C}^2(\mathbb{R}) \cap \chi_0$,

$$\begin{aligned} \left\langle u_t^{\theta,\epsilon}, f\right\rangle_{\chi_0} &= \langle F, f\rangle_{\chi_0} + \int_0^t \left\langle u_s^{\theta,\epsilon}, \frac{1}{2}f''\right\rangle_{\chi_0} ds \\ &+ \theta\int_0^t\int_U\int_{\mathbb{R}} G(a,x,u_s^{\theta,\epsilon}(x))f(x)J(x)dxW(dads) \\ &+ \sum_j \int_0^t\int_{\mathbb{R}} G_j(x,u_s^{\theta,\epsilon}(x))f(x)J(x)dxk_s^{\epsilon,j}ds \end{aligned}$$

Applying Ito's formula, we obtain,

$$
\begin{aligned}
& \left\langle u_t^{\theta,\epsilon}, f\right\rangle_{\chi_0}^2 \\
= \; & \langle F, f\rangle_{\chi_0}^2 + \int_0^t 2\left\langle u_s^{\theta,\epsilon}, f\right\rangle_{\chi_0}\left\langle u_s^{\theta,\epsilon}, \frac{1}{2}f''\right\rangle_{\chi_0} ds \\
+ \; & \theta\int_0^t\int_U 2\left\langle u_s^{\theta,\epsilon}, f\right\rangle_{\chi_0}\int_{\mathbb{R}} G(a,x,u_s^{\theta,\epsilon}(x))f(x)J(x)dxW(dads) \\
+ \; & \sum_j\int_0^t 2\left\langle u_s^{\theta,\epsilon}, f\right\rangle_{\chi_0}\int_{\mathbb{R}} G_j(x,u_s^{\theta,\epsilon}(x))f(x)J(x)dxk_s^{\epsilon,j}ds \\
+ \; & \theta^2\int_0^t\int_U\left|\int_{\mathbb{R}} G(a,x,u_s^{\theta,\epsilon}(x))f(x)J(x)dx\right|^2\lambda(da)ds
\end{aligned}
$$

Summing on f over a CONS of χ_0 we arrive at,

$$
\begin{aligned}
\|u_t^{\theta,\epsilon}\|_{\chi_0}^2 \;\le\; & \|F\|_{\chi_0}^2 + \int_0^t 2\left\langle u_s^{\theta,\epsilon}, \Delta u_s^{\theta,\epsilon}\right\rangle_{\chi_0} ds \\
& +2\theta\int_0^t\int_U\left\langle u_s^{\theta,\epsilon}, G(a,.,u_s^{\theta,\epsilon})\right\rangle_{\chi_0} W(dads) \\
& +2\sum_j\int_0^t\left\langle G_j(.,u_s^{\theta,\epsilon}), u_s^{\theta,\epsilon}\right\rangle_{\chi_0} k_s^{\epsilon,j}ds \\
& +\theta^2\int_0^t\int_U\|G(a,.,u_s^{\theta,\epsilon})\|_{\chi_0}^2\lambda(da)ds
\end{aligned}
$$

We note that using Cauchy-Schwartz inequality, the one before the last term can be estimated as follows,

$$
\begin{aligned}
& 2\sum_j\int_0^t\left\langle G_j\left(.,u_s^{\theta,\epsilon}\right), u_s^{\theta,\epsilon}\right\rangle_{\chi_0} k_s^{\epsilon,j}ds \\
\le \; & 2\int_0^t\left(\sum_j\left\langle G_j\left(.,u_s^{\theta,\epsilon}\right), u_s^{\theta,\epsilon}\right\rangle_{\chi_0}^2\right)^{1/2}\|k_s^\epsilon\|_{\ell_2}ds \\
\le \; & 2\int_0^t\left(\sum_j\left(\int_{\mathbb{R}} G_j\left(x,u_s^{\theta,\epsilon}(x)\right)u_s^{\theta,\epsilon}(x)J(x)dx\right)^2\right)^{1/2}\|k_s^\epsilon\|_{\ell_2}ds \\
= \; & 2\int_0^t\left(\int_{\mathbb{R}}\int_U G\left(a,x,u_s^{\theta,\epsilon}(x)\right)^2\lambda(da)J(x)dx\right)^{1/2}\|u_s^{\theta,\epsilon}\|_{\chi_0}\|k_s^\epsilon\|_{\ell_2}ds \\
\le \; & 2\int_0^t\|k_s^\epsilon\|_{\ell_2}\|u_s^{\theta,\epsilon}\|_{\chi_0}\left(\int_{\mathbb{R}}\left(1+|u_s^{\theta,\epsilon}(x)|^2\right)J(x)dx\right)^{1/2}ds \\
\le \; & KN\int_0^t\left(1+\|u_s^{\theta,\epsilon}\|_{\chi_0}\right)ds
\end{aligned}
$$

where K is a constant depending on N but not on $\theta \leq 1$ and $k \in \mathcal{S}^N$. The other terms can be estimated similar to those in Lemma 2.3 of Xiong [69]. By Burkholder-Davis-Gundy inequality, we then obtain,

$$\mathbb{E} \sup_{s \leq t} \|u_t^{\theta,\epsilon}\|_{\chi_0}^{2n} \leq KN\mathbb{E} \int_0^t \left(1 + \|u_s^{\theta,\epsilon}\|_{\chi_0}^{2n}\right) ds$$

It follows from Gronwall's inequality (see Appendix) that

$$\mathbb{E} \left(\sup_{t \leq 1} \int_{\mathbb{R}} |u_s^{\theta,\epsilon}|^2 e^{-2\beta_1 |x|} dx \right)^n < \infty \tag{3.45}$$

Finally we estimate $\mathbb{E} \left| u_{t_1}^{\theta,\epsilon}(y_1) - u_{t_2}^{\theta,\epsilon}(y_2) \right|^n$ using (3.43) and (3.45). Since the main difference between $\{u^{\theta,\epsilon}\}$ and $\{u_t^\epsilon\}$ is in the last term, we restrict our attention to

$$w_t(y) := \sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x) G_j(a, x, u_s(x)) k_s^{\epsilon,j} dx ds. \tag{3.46}$$

and find estimates for (3.46) similar to those obtained in the proof of Theorem 3.2.1. Using (3.36), Cauchy-Schwarz Inequality and

$$P_1 := p_{t-s}(y_1 - x) - p_{t-s}(y_2 - x),$$

we have,

$$
\begin{aligned}
& \mathbb{E}\left|w_t(y_1) - w_t(y_2)\right|^n \\
= \ & \mathbb{E}\left|\int_0^t \int_{\mathbb{R}} P_1 \sum_j G_j(a, x, u_s(x)) k_s^{\epsilon, j} dx ds\right|^n \\
\leq \ & \mathbb{E}\left|\int_0^t \int_{\mathbb{R}} P_1 \left(\sum_j G_j(a, x, u_s(x))^2\right)^{\frac{1}{2}} \|k_s^{\epsilon}\|_{\ell_2} dx ds\right|^n \\
\leq \ & K\mathbb{E}\left|\int_0^t \int_{\mathbb{R}} P_1 \left(\sum_j < G(a, x, u_s(x)), \phi_j(a) >_{L^2}^2\right)^{\frac{1}{2}} \|k_s^{\epsilon}\|_{\ell_2} dx ds\right|^n \\
= \ & K\mathbb{E}\left|\int_0^t \int_{\mathbb{R}} P_1 \left(\|G(a, x, u_s(x))\|_{L^2}^2\right)^{\frac{1}{2}} \|k_s^{\epsilon}\|_{\ell_2} dx ds\right|^n \\
\leq \ & \mathbb{E}\left|\left(\int_0^t \left(\int_{\mathbb{R}} |P_1| \sqrt{K\left(1 + |u_s(y)|^2\right)} dx\right)^2 dx\right)^{\frac{1}{2}} \left(\int_0^t \|k_s^{\epsilon}\|_{\ell_2}^2 ds\right)^{\frac{1}{2}}\right|^n \\
\leq \ & \mathbb{E}\left|\int_0^t \left(\int_{\mathbb{R}} |P_1| \sqrt{K\left(1 + |u_s(x)|^2\right)} dx\right)^2 ds\right|^{\frac{n}{2}} N^{\frac{n}{2}} \\
\leq \ & N^{\frac{n}{2}} M K e^{n\beta_1(|y_1| \vee |y_2|)} |y_1 - y_2|^{\frac{\alpha n}{2}} \\
\leq \ & K e^{n\beta_1(|y_1| \vee |y_2|)} |y_1 - y_2|^{\frac{\alpha n}{2}}
\end{aligned}
$$

where the final step follows from an analogous argument as in the proof of Theorem 3.2.1 and M is given by (3.26). Similarly for a fixed $y \in \mathbb{R}$ and $0 \leq t_1 \leq t_2 \leq 1$ arbitrary, let

$$P_2 := p_{t_1 - s}(y - x) - p_{t_2 - s}(y - x)$$

then

$$
\begin{aligned}
& \mathbb{E}\left|w_{t_1}(y)-w_{t_2}(y)\right|^n \\
= \;& \mathbb{E}\left|\sum_j \int_0^{t_1}\int_{\mathbb{R}} p_{t_1-s}(y-x)G_j(a,x,u_s(x))k_s^{\epsilon,j}dxds\right. \\
& \left. -\sum_j \int_0^{t_2}\int_{\mathbb{R}} p_{t_2-s}(y-x)G_j(a,x,u_s(x))k_s^{\epsilon,j}dxds\right|^n \\
= \;& \mathbb{E}\left|\sum_j \int_0^{t_1}\int_{\mathbb{R}} P_2 G_j(a,x,u_s(x))k_s^{\epsilon,j}dxds\right. \\
& \left. +\sum_j \int_{t_1}^{t_2}\int_{\mathbb{R}} p_{t_2-s}(y-x)G_j(a,x,u_s(x))k_s^{\epsilon,j}dxds\right|^n \\
\leq \;& K\mathbb{E}\left|\int_0^{t_1}\int_{\mathbb{R}} P_2\left(\sum_j G_j(x,u_s(x))^2\right)^{\frac{1}{2}}\|k_s^{\epsilon}\|_{\ell_2}dxds\right|^n \\
& +K\mathbb{E}\left|\int_{t_1}^{t_2}\int_{\mathbb{R}} p_{t_2-s}(y-x)\left(\sum_j G_j(x,u_s(x))^2\right)^{\frac{1}{2}}\|k_s^{\epsilon}\|_{\ell_2}dxds\right|^n \\
\leq \;& \mathbb{E}\left|\left(\int_0^{t_1}\left(\int_{\mathbb{R}}|P_2|\sqrt{K\left(1+|u_s(x)|^2\right)}dx\right)^2 ds\right)^{\frac{1}{2}}\left(\int_0^{t_1}\|k_s^{\epsilon}\|_{\ell_2}^2 ds\right)^{\frac{1}{2}}\right|^n \\
& +\mathbb{E}\left|\left(\int_{t_1}^{t_2}\left(\int_{\mathbb{R}}|p_{t_2-s}(y-x)|\sqrt{K\left(1+|u_s(x)|^2\right)}dx\right)^2 ds\right)^{\frac{1}{2}}\left(\int_{t_1}^{t_2}\|k_s^{\epsilon}\|_{\ell_2}^2 ds\right)^{\frac{1}{2}}\right|^n \\
\leq \;& \mathbb{E}\left|\int_0^{t_1}\left(\int_{\mathbb{R}}|P_2|\sqrt{K\left(1+|u_s(x)|^2\right)}dx\right)^2 ds\right|^{\frac{n}{2}} N^{\frac{n}{2}} \\
& +\mathbb{E}\left|\int_{t_1}^{t_2}\left(\int_{\mathbb{R}}|p_{t_2-s}(y-x)|\sqrt{K\left(1+|u_s(x)|^2\right)}dx\right)^2 ds\right|^{\frac{n}{2}} N^{\frac{n}{2}} \\
\leq \;& Ke^{n\beta_1|y|}\left|t_1-t_2\right|^{\frac{\alpha n}{2}}
\end{aligned}
$$

By (3.45) and the two estimates above, $\{u^{\epsilon,\theta}\}$ satisfies the two conditions given in Lemma 3.2.1 and thus $u^{\epsilon,\theta}$ is tight. If we set $\theta = 0$ then $u_t^{\theta,\epsilon}(y)$ becomes,

$$
u_t^{0,\epsilon}(y) = \int_{\mathbb{R}} p_t(y-x)F^{\epsilon}(x)dx + \sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x)G_j(a,x,u_s^{\epsilon}(x))k_s^{\epsilon,j}dxds \tag{3.47}
$$

We let $g^0 : \mathbb{B}_{\alpha,\beta_0} \times \mathcal{S} \to \tilde{\tilde{\mathcal{C}}}\left([0,1];\mathbb{B}_{\beta}\right)$ be given by (3.47). By the tightness of $\{u_{\cdot}^{0,\epsilon}\}$ we apply

Prohorov's Theorem (cf. Appendix) to obtain a convergent subsequence in distribution but g^0 is a deterministic equation so we obtain sequential compactness and $\tilde{\tilde{\mathcal{C}}}([0,1];\mathbb{B}_\beta)$ being a metric space we have compactness of g^0 and hence part one of Assumption 1 holds. For part two we let $\theta = \sqrt{\epsilon}$. That is we consider,

$$\begin{aligned} u_t^{\sqrt{\epsilon},\epsilon}(y) &= \int_{\mathbb{R}} p_t(y-x)F^\epsilon(x)dx + \sqrt{\epsilon}\sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x)G_j(a,x,u_s(x))dxdB_s^j \\ &\quad + \sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x)G_j(a,x,u_s(x))k_s^{\epsilon,j}dxds \end{aligned} \tag{3.48}$$

Notice that this is $g^\epsilon\left(F^\epsilon, \sqrt{\epsilon}\beta + \int_0^\cdot k_s^\epsilon ds\right)$ and by letting $\epsilon \to 0$ we obtain,

$$u_t^{0,0}(y) = \int_{\mathbb{R}} p_t(y-x)F(x)dx + \sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x)G_j(a,x,u_s^{0,0}(x))k_s^j dxds \tag{3.49}$$

which is exactly $g^0\left(F, \int_0^\cdot k_s ds\right)$ as in the assumption using the notation $h = \zeta(k)$. Since by introducing equivalence classes we obtain uniqueness of solution to (3.49), then a subsequence converging in distribution to a unique limit implies the convergence in distribution of the sequence. Therefore, part two of Assumption 1 is also established. Thus based on Theorem 3.2.2, the rate function for SPDE (3.8) is given as,

$$\tilde{I}(u) = \begin{cases} \frac{1}{2}\inf\left\{\int_0^1 \|k_s\|_{\ell_2}^2 ds : u = \gamma(F,\zeta(k))\right\} & \exists k \text{ s.t.} u = \gamma(F,\zeta(k)) \\ \infty & \text{otherwise.} \end{cases}$$

Notice that

$$\begin{aligned} \|k_s\|_{\ell_2}^2 &= \sum_j \left(\int_U h_s(a)\phi_j(a)\lambda(da)\right)^2 \\ &= \sum_j \langle h_s, \phi_j\rangle^2 = \|h_s\|_{L^2}^2 = \int_U |h_s(a)|^2\lambda(da) \end{aligned}$$

which gives the form in (3.44). □

We note that we were unable to establish the transitive property of the equivalence relation $\sim_1$ for our quotient space $\mathcal{C}([0,1],\mathbb{B}_\beta)/\sim_1$ for SBM. Hence the LDP result of this section cannot be applied to SBM requiring further investigation.

3.3 Large Deviations for Fleming-Viot Process

We begin by confirming the transitive property of the relation $\sim_1$ for FVP.

Lemma 3.3.1. *The controlled PDE for FVP satisfies the transitive property of relation $\sim_1$ given by Definition 3.2.1.*

Proof. For FVP $G(a,y,u) = 1_{a<u} - u$ thus plugging it into the controlled version of the general SPDE we obtain for $h \in L^2([0,1] \times U, ds\lambda(da))$,

$$u_t(y) = F(y) + \int_0^t \int_0^{u_s(y)} h_s(a)dads - \int_0^t \int_0^1 u_s(y)h_s(a)dads + \int_0^t \frac{1}{2}\Delta u_s(y)ds \tag{3.50}$$

Suppose u^1, u^2, u^3 are solutions to (3.50) such that $u^1 \sim u^2$ with h^1 and $u^2 \sim u^3$ with h^2. Notice that these two relations have solution u^2 in common with different functions $h \in L^2([0,1] \times U, ds\lambda(da))$, hence subtracting these two solutions gives,

$$0 = \int_0^t \int_0^{u_s^2(y)} \left(h_s^1(a) - h_s^2(a)\right) dads - \int_0^t \int_0^1 u_s^2(y)\left(h_s^1(a) - h_s^2(a)\right) dads$$

further we have,

$$0 = \int_0^{u_t^2(y)} (H_t(a))\, da - u_t^2(y) \int_0^1 (H_t(a))\, da$$

where $H_t(a) := h_t^1(a) - h_t^2(a)$. We now introduce,

$$\begin{aligned} \tilde{h}_t^1(a) &:= h_t^1(a) - \int_0^1 h_t^1(b)db \\ \tilde{H}_t(a) &:= \tilde{h}_t^1(a) - \tilde{h}_t^2(a) \end{aligned}$$

then we obtain the relation,

$$\tilde{H}_t(a) = H_t(a) - \int_0^1 H_t(b)db$$

which helps us to arrive at,

$$\int_0^{u_t^2(y)} \tilde{H}_t(a)da = \int_0^{u_t^2(y)} H_t(a)da - u_t^2(y)\int_0^1 H_t(a)da = 0$$

Since this is true for all $u_t^2(y) \in [0,1]$ then $\tilde{H}_t(a) = 0$ a.e. Hence, using the definition of $\tilde{h}_t^i, i = 1, 2$, and solving $\tilde{H}_t(a) = 0$ for $h_t^1(a)$ we obtain,

$$h_t^1(a) = \int_0^1 h_t^1(b)db + h_t^2(a) - \int_0^1 h_t^2(b)db \tag{3.51}$$

Note that to achieve the transitive property it is sufficient to show that $u_t^1(y)$ satisfies the controlled PDE (3.50) with h^1 and h^2. We plug in the form $h_t^1(a)$ found in (3.51) in the PDE for u_t^1 to get,

$$\begin{aligned} u_t^1(y) &= F(y) + \int_0^t u_s^1(y) \int_0^1 h_s^1(b) db ds \\ &+ \int_0^t \int_0^{u_s^1(y)} h_s^2(a) da ds - \int_0^t u_s^1(y) \int_0^1 h_s^2(b) db ds \\ &- \int_0^t \int_0^1 u_s^1(y) h_s^1(a) da ds + \frac{1}{2} \int_0^t \Delta u_s^1(y) ds \end{aligned}$$

which is equivalent to the controlled PDE for u_t^1 with h^2 if

$$\int_0^t u_s^1(y) \int_0^1 h_s^1(b) db ds - \int_0^t \int_0^1 u_s^1(y) h_s^1(a) da ds = 0$$

but this is true since $u_s^1(y)$ does not depend on a. $\square$

We remark here that since $U = \mathbb{R}$ in the case of SBM then the above argument does not lead to $\tilde{H}_t(a) = 0$ a.e. to show the transitive property. Therefore, another approach and reasoning are required.

Similar to [24] we consider the Cameron-Martin space which is defined as follows. Let $\mathcal{D}$ be the Schwartz space of test functions with compact support in $\mathbb{R}$ and continuous derivatives of all orders. Denote the dual space of real distributions on $\mathbb{R}$ by $\mathcal{D}^*$ then for a fixed $\nu \in \mathcal{M}_\beta(\mathbb{R})$, the Cameron-Martin space, H_ν, is the set of measures $\mu_t \in \mathcal{C}([0,1]; \mathcal{M}_\beta(\mathbb{R}))$ satisfying the conditions below.

1. $\mu_0 = \nu$,
2. the $\mathcal{D}^*$-valued map $t \mapsto \mu_t$ defined on [0,1] is absolutely continuous with respect to time. Let $\dot{\mu}$ and $\Delta^*\mu$ be its generalized derivative and Laplacian respectively,
3. for every $t \in [0,1]$, $\dot{\mu}_t - \frac{1}{2}\Delta^*\mu_t \in \mathcal{D}^*$ is absolutely continuous with respect to ω_t with $\frac{d\left(\dot{\mu}_t - \frac{1}{2}\Delta^*\mu_t\right)}{d\mu_t}$ being the (generalized) Radon Nikodym derivative,
4. $\frac{d\left(\dot{\mu}_t - \frac{1}{2}\Delta^*\mu_t\right)}{d\mu_t}$ is in $L^2([0,1] \times \mathbb{R}, ds\mu(dy))$.

Let $\tilde{H}_\nu$ be the space for which conditions for H_ν hold with $\mathcal{M}_\beta(\mathbb{R})$ replaced by the space of probability measures $\mathcal{P}(\mathbb{R})$, and with the additional assumption,

$$\left\langle \mu_t^0, \frac{d\left(\dot{\mu}_t - \frac{1}{2}\Delta^*\mu_t\right)}{d\mu_t^0} \right\rangle = 0.$$

Let $\mathcal{C}\left([0,1];\mathcal{P}_\beta(\mathbb{R})\right)/\sim$ be the quotient space of $\mathcal{C}\left([0,1];\mathcal{P}_\beta(\mathbb{R})\right)$ defined as: $\mu \sim \nu$ if $\mu_0 = \nu_0, \mu, \nu \in \tilde{H}_{\mu_0}$ and

$$\frac{d\left(\dot{\mu}_t - \frac{1}{2}\Delta^*\mu_t\right)}{d\mu_t} = \frac{d\left(\dot{\nu}_t - \frac{1}{2}\Delta^*\nu_t\right)}{d\nu_t}$$

and if μ is not in $\tilde{H}_{\mu_0}$ then μ's equivalence class is composed of itself only.

Now we proceed to establish the LDP for FVP by applying the results from the previous section.

Lemma 3.3.2. *Let $\mathcal{A}$ be the set of all nondecreasing functions, then map $\xi : \mathbb{B}_\beta \cap \mathcal{A} \to \mathcal{P}_\beta(\mathbb{R})$ defined as $\xi(u)(B) = \int 1_B(y)du(y)$ for all $B \in \mathbb{B}(\mathbb{R})$, is Lipschitz continuous.*

Proof. Let $\mathcal{W}$ be the collection of all $f \in \mathcal{C}_b^1(\mathbb{R})$ such that $|f(x)| \leq 1$ and $|f'(x)| \leq 1$ for all $x \in \mathbb{R}$. For any $u, v \in \mathbb{B}_\beta \cap \mathcal{A}$ we have,

$$\begin{aligned}
\rho_\beta(\xi(u), \xi(v)) &= \sup_{f \in \mathcal{W}} \left| \int_{\mathbb{R}} f(x) J_\beta(x)(\xi(u) - \xi(v))(dx) \right| \\
&= \sup_{f \in \mathcal{W}} \left| \int_{\mathbb{R}} (f J_\beta)'(x)(u(x) - v(x))(dx) \right| \\
&\leq K\|u - v\|_\beta
\end{aligned}$$

which proves the Lipschitz continuity of map, ξ. $\square$

Next we define the map $\eta : \tilde{\tilde{\mathcal{C}}}\left([0,1];\mathbb{B}_\beta\right) \to \tilde{\tilde{\mathcal{C}}}\left([0,1];\mathcal{P}_\beta(\mathbb{R})\right)$ by $\eta(u)_t = \xi(u_t)$. It is then clear that η is Lipschitz continuous and $\eta(u) \sim \eta(v)$ whenever $u \sim v$. Therefore, η can be regarded as a map from $\tilde{\tilde{\mathcal{C}}}\left([0,1];\mathbb{B}_\beta\right)$ to $\tilde{\tilde{\mathcal{C}}}\left([0,1];\mathcal{P}_\beta\right)$.

Lemma 3.3.3. $\eta : \tilde{\tilde{\mathcal{C}}}\left([0,1];\mathbb{B}_\beta\right) \to \tilde{\tilde{\mathcal{C}}}\left([0,1];\mathcal{P}_\beta\right)$ *is continuous.*

Proof. For better presentation, we let $\tilde{\tilde{u}} := [[u]_1]_2, \tilde{\tilde{v}} := [[v]_1]_2$ and denote the metrics induced by $\tilde{\tilde{\mathcal{C}}}\left([0,1];\mathbb{B}_\beta\right)$ and $\tilde{\tilde{\mathcal{C}}}\left([0,1];\mathcal{P}_\beta\right)$ by $\rho_{\mathbb{B}_\beta}$ and $\rho_{\mathcal{P}_\beta}$, respectively. So we have

$$\rho_{\mathbb{B}_\beta}\left(\tilde{\tilde{u}}, \tilde{\tilde{v}}\right) = \tilde{d}_{\mathbb{B}_\beta}\left([u]_1, [v]_1\right)$$

If $\tilde{d}_{\mathbb{B}_\beta}\left([u]_1,[v]_1\right)=0$ then $\tilde{\tilde{u}}=\tilde{\tilde{v}}$ so $\eta\left(\tilde{\tilde{u}}\right)=\eta\left(\tilde{\tilde{v}}\right)$. Thus in this case η is continuous. If on the other hand, $\tilde{d}_{\mathbb{B}_\beta}\left([u]_1,[v]_1\right)\neq 0$ then using the definition of pseudo-metric $\tilde{d}$, for any $\epsilon>0$ let $p_1\sim u, q_n\sim v$ and $q_i\sim p_{i+1}$, $i=1,...,n-1$ be such that

$$\sum_{j=1}^{n}\|p_i-q_i\|_{\mathcal{C}_\beta}\leq \tilde{d}_\beta(u,v)+\epsilon$$

where $\|.\|_{\mathcal{C}_\beta}$ is the norm in space $\mathcal{C}\left([0,1];\mathbb{B}_\beta\right)$ and $d_{\mathcal{C}_\beta}$ is its corresponding metric. By last lemma, we have

$$\rho_{\mathcal{P}_\beta}\left(\eta(p_i),\eta(q_i)\right)\leq K\left\|p_i-q_i\right\|_\beta$$

where

$$\rho_{\mathcal{P}_\beta}(\mu,\nu)=\sup_{t\leq 1}\rho_\beta(\mu_t,\nu_t)$$

is the metric in $\mathcal{C}\left([0,1];\mathcal{P}_\beta(\mathbb{R})\right)$. Hence,

$$\tilde{\rho}_{\mathcal{P}_\beta}\left(\eta(u),\eta(v)\right)\leq\sum_{j=1}^{\infty}\rho_{\mathcal{P}_\beta}\left(\eta(p_i),\eta(q_i)\right)\leq K\tilde{d}_\beta(u,v)+K\epsilon$$

The conclusion now follows since ϵ was arbitrary. □

Theorem 3.3.1. *Suppose $\mu_0\in\mathcal{P}_\beta(\mathbb{R})$ such that $F\in\mathbb{B}_{\alpha,\beta_0}$. Then, $\{\mu^\epsilon\}$ satisfies the LDP on $\tilde{\tilde{\mathcal{C}}}([0,1];\mathcal{P}_\beta(\mathbb{R}))$ with rate function,*

$$I(\mu)=\begin{cases}\frac{1}{2}\int_0^1\int_{\mathbb{R}}\left|\frac{d\left(\dot{\mu}_t-\frac{1}{2}\Delta^*\mu_t\right)(y)}{d\mu_t(y)}\right|^2 d\mu_t(y)dt & \text{if } \mu\in\tilde{H}_{\mu_0}\\ \infty & \text{otherwise.}\end{cases}\tag{3.52}$$

Proof. Recall FVP can be represented by

$$u_t^\epsilon(y)=F(y)+\int_0^t\int_0^1\left(1_{a\leq u_s^\epsilon(y)}-u_s^\epsilon(y)\right)W(dads)+\int_0^t\frac{1}{2}\Delta u_s^\epsilon(y)ds.\tag{3.53}$$

By Theorem 3.2.4, $\{u_t^\epsilon\}$ satisfies the LDP on $\tilde{\tilde{\mathcal{C}}}\left([0,1];\mathbb{B}_\beta\right)$ and because $u_t^\epsilon\in\mathcal{A}$ a.s. for all t, we see that u_t^ϵ obeys the LDP on $\tilde{\tilde{\mathcal{C}}}\left([0,1];\mathbb{B}_\beta\cap\mathcal{A}\right)$ as well. Since for FVP, $\mu_t^\epsilon=\xi(u_t^\epsilon)$, then by Lemma 3.3.3 and the contraction principle (cf. Appendix), LDP holds for μ_t^ϵ in $\tilde{\tilde{\mathcal{C}}}\left([0,1];\mathcal{P}_\beta(\mathbb{R})\right)$ with the rate function determined as follows.

If $I(\mu) < \infty$, then there exists $h \in L^2([0,1] \times \mathbb{R}_+, dsda)$ such that (3.39) holds. Let $\mathcal{C}_c(\mathbb{R})$ be the collection of functions with compact support on $\mathbb{R}$, then for $f \in \mathcal{C}_c^1(\mathbb{R})$,

$$< \mu_t, f >= \int f(y)\mu_t(dy) = -\int f'(y)u_t(y)dy = - < u_t, f' >_{L^2(\mathbb{R})} .$$

Therefore using the controlled PDE (3.39),

$$u_t(y) = F(y) + \int_0^t \int_U G(a, y, u_s(y))h_s(a)\lambda(da)ds + \int_0^t \frac{1}{2}\Delta u_s(y)ds. \tag{3.54}$$

with $G(a, y, u) = 1_{a<u} - u$ we have for every $f \in \mathcal{C}_c^3(\mathbb{R})$,

$$\begin{aligned}
< \mu_t, f > &= - < F, f' > + \int_0^t < u_s', \frac{1}{2}\Delta f > ds \\
&\quad - \int_0^t \int_{\mathbb{R}} \int_0^{u_s(y)} h_s(a)da f'(y)dyds - \int_0^t < \mu_s, f > \int_0^1 h_s(a)dads \\
&= < \mu_0, f > + \int_0^t < \mu_s, \frac{1}{2}\Delta f > ds + \int_0^t \int_{\mathbb{R}} f(u_s^{-1}(a))h_s(a)dads \\
&\quad - \int_0^t < \mu_s, f > \int_0^1 h_s(a)dads \\
&= < \mu_0, f > + \int_0^t < \frac{1}{2}\Delta^*\mu_s, f > ds + \int_0^t \int_{\mathbb{R}} f(y)h_s(u_s(y))du_s(y)ds \\
&\quad - \int_0^t < \mu_s, f > \int_0^1 h_s(a)dads \\
&= < \mu_0, f > + \int_0^t < \frac{1}{2}\Delta^*\mu_s, f > ds + \int_0^t < \mu_s, fh_s(u_s) > ds \\
&\quad - \int_0^t < \mu_s, f > \int_0^1 h_s(a)dads
\end{aligned}$$

Summing on a CONS, $\{f_j\}$,

$$\frac{\left(\dot{\mu}_t - \frac{1}{2}\Delta^*\mu_t\right)(dy)}{\mu_t(dy)} = h_t(u_t(y)) - \int_0^1 h_t(a)da.$$

hence, $\mu \in H_{\mu_0}$. If h satisfies (3.39) then $\bar{h}_s(a) \equiv h_s(a) - \int_0^1 h_s(a)da$ also satisfies the same equation. To minimize $\int_0^1 |h_s(a)|^2 da$, we choose h such that $\int_0^1 h_s(a)da = 0$. Therefore, $\mu \in \tilde{H}_{\mu_0}$ and

$$\frac{d\left(\dot{\mu}_t - \frac{1}{2}\Delta^*\mu_t\right)(y)}{d\mu_t(y)} = h_t(u_t(y)).$$

Denote the right hand side of (4.38) by $I_0(\mu)$ and observe that in this case, $I_0(\mu) = I(\mu)$. If $I_0(\mu) < \infty$ we may reverse the above calculation to obtain the finiteness of $I(\mu)$. This completes the proof. $\square$

Chapter 4

Moderate Deviations

4.1 Introduction

In this chapter we strive to prove the moderate deviation principle for the general SPDE given in chapter three and for the two population models of our study. First we provide some background on moderate deviations. The term moderate deviation was introduced by Rubin and Sethuraman in 1965 in their work [49]. While computing results for Bayes Risk Efficiency in [50], the significance of this estimate became apparent to these authors. As described in [49], suppose $X_1, X_2, ...$ are i.i.d. random variables with common distribution function, F and $\mathbb{E}\left(X_1\right) = 0$ and $\mathbb{E}\left(X_1^2\right) = 1$. Then by the Central Limit Theorem (CLT),

$$P\left(\left|\frac{S_n}{n}\right| > \frac{a}{\sqrt{n}}\right) \to 1 - \Phi(a) \tag{4.1}$$

where $S_n = \sum_{i=1}^{n} X_i$ and $\Phi(x) \sim N(0,1)$. In words, (4.1) provides the probability of an event that deviates from the mean, $\frac{S_n}{n}$, by $\frac{a}{\sqrt{n}}$. Deviation of this amount, denoted as $\lambda_n := \frac{a}{\sqrt{n}}$, is called ordinary(or normal) deviation. When $\lambda_n := \lambda$ for all n, that is when considering $P\left(\left|\frac{S_n}{n}\right| > \lambda\right)$, then the deviation is referred to as Large deviation (as described in chapter three). The case for which $\lambda_n := c\sqrt{\log n/n}$ for all n: $P\left(\left|\frac{S_n}{n}\right| > c\sqrt{\log n/n}\right)$ is called moderate deviations. Wu [65] also classifies the three cases but with the use of Borel subsets. He explains for the above sequence that if $m = \mathbb{E}(X_1)$, then for a Borel subset of $\mathbb{R}$, A_n, the cases can be summarized as

$$P\left(\frac{S_n}{n} - m \in A_n\right)$$

for a suitable deviation, A_n. $P\left(\frac{S_n}{n} - m \in \frac{1}{\sqrt{n}}A\right)$ gives the Central Limit Theorem (henceforth CLT), $P\left(\frac{S_n}{n} - m \in A\right)$ Large Deviations and
$P\left(\frac{S_n}{n} - m \in \frac{b(n)}{\sqrt{n}}A\right)$ Moderate Deviations where $0 < b(n) \to \infty$, $\frac{b(n)}{\sqrt{n}} \to 0$.

By the conditions on $b(n)$, we deduce $\frac{b(n)}{\sqrt{n}} < 1$ for large n, therefore, $\frac{A}{\sqrt{n}} < \frac{b(n)A}{\sqrt{n}} < A$, implying that Moderate deviation is in between deviation for CLT and large deviations, hence the name Moderate is used. Analogous to Large Deviations, Moderate Deviation is defined by the moderate deviation principle.

Definition 4.1.1 (Moderate Deviation Principle)**.** *A collection $X = \{X_n\}$ of random variables satisfies the moderate deviation principle(MDP) if for some sequence $b(n)$ such that $b(n) \to 0$, the family $\{b(n)X_n\}$ satisfies the LDP with a rate of convergence slower than the one for the LDP of X and faster than the convergence rate of the central limit theorem (CLT) of X.*

Here we use the same space $\mathcal{M}_\beta(\mathbb{R})$ as in chapter three. We remark that in the case of MDP, the controlled PDE of the general SPDE has a unique solution; therefore, the quotient space and the equivalence relations $\sim_1$ and $\sim_2$ are not needed. We prove the MDP in the Polish space $\mathcal{C}([0,1];\mathbb{B}_\beta)$ and do not require the general SPDE to have the transitive property.

4.2 Moderate Deviations for the General SPDE

Our goal in this section is to establish the MDP for the general SPDE studied in chapter three. Recall this SPDE has the form,

$$u_t^\epsilon(y) = F(y) + \sqrt{\epsilon}\int_0^t \int_U G(a, y, u_s^\epsilon(y))W(dads) + \int_0^t \frac{1}{2}\Delta u_s^\epsilon(y)ds \tag{4.2}$$

having conditions,

$$\int_U |G(a,y,u_1) - G(a,y,u_2)|^2 \lambda(da) \quad \leq \quad K|u_1 - u_2| \tag{4.3}$$

$$\int_U |G(a,y,u)|^2 \lambda(da) \quad \leq \quad K(1 + |u|^2) \tag{4.4}$$

where $u_1, u_2, u, y \in \mathbb{R}$, F is a function on $\mathbb{R}$ and $G : U \times \mathbb{R}^2 \to \mathbb{R}$. To investigate the MDP for this SPDE, we consider a family $a(\epsilon)$ satisfying $0 \le a(\epsilon) \to 0$ and $\frac{a(\epsilon)}{\sqrt{\epsilon}} \to \infty$ as $\epsilon \to 0$ and establish the LDP for centered process given by,

$$v_t^\epsilon(y) := \frac{a(\epsilon)}{\sqrt{\epsilon}}(u_t^\epsilon(y) - u_t^0(y)). \tag{4.5}$$

Note that the assumption $a(\epsilon)/\sqrt{\epsilon} \to \infty$ implies that $a(\epsilon)$ converges to zero at a slower rate than $\sqrt{\epsilon}$ tending to zero. Hence, the convergence rate is slower than the rate considered in chapter three for the case of LDP, as is desired for moderate deviations. The process v_t^ϵ given by (4.5) is referred to as the general SPDE in the context of this chapter and can be explicitly written as

$$v_t^\epsilon(y) = a(\epsilon)\int_0^t \int_U G(a, y, \frac{\sqrt{\epsilon}}{a(\epsilon)} v_s^\epsilon(y) + u_s^0(y)) W(dads) + \frac{1}{2}\int_0^t \Delta v_s^\epsilon(y) ds \tag{4.6}$$

with controlled PDE,

$$v_t(y) = \int_0^t \int_U G(a, y, u_s^0(y)) h_s(a) \lambda(da) ds + \frac{1}{2}\int_0^t \Delta v_s(y) ds \tag{4.7}$$

where $h \in L^2\left([0,1] \times U, ds\lambda(da)\right)$. Notice that for fixed h, (4.7) has a unique solution. To see this, let $\tilde{v}_t$ and v_t be two solutions to (4.7) with the same $h_s(a)$, then since $\tilde{u}_s^0(y) = u_s^0(y) = F(y)$, then we can write $w := \tilde{v}_t - v_t$ in the form of homogeneous heat equation:

$$\dot{w}_t(y) - \frac{1}{2}\Delta w_t = 0$$

where $\dot{w}_t$ denotes the derivative with respect to time. Noting that $\tilde{v}_., v_. \in \mathcal{C}\left([0,1]; \mathbb{B}_\beta\right)$, then both $\tilde{v}(y)$ and $v(y)$ are bounded above by $Ke^{\beta|y|}$ for all $y \in \mathbb{R}$ and so $Ke^{\beta|y|}$ serves as a boundary condition and is zero in the case of $w_.$. Therefore, by the maximum principle we have the uniqueness of solutions.

Now we aim to prove the following theorem in this section using the same technique as was applied in chapter two.

Theorem 4.2.1. *Family $\{v_.^\epsilon\}$ given by (4.6) satisfies the LDP in $\mathcal{C}\left([0,1]; \mathbb{B}_\beta\right)$ with rate function,*

$$I(v) = \frac{1}{2}\inf\left\{\int_0^1 \int_U |h_s(a)|^2 \lambda(da) ds : v = \gamma(h)\right\} \tag{4.8}$$

where $\gamma : L^2([0,1]\times U, ds\lambda(da)) \to \mathcal{C}([0,1];\mathbb{B}_\beta)$ is a map such that for $h \in L^2([0,1]\times U, ds\lambda(da))$, $\gamma(h)$ is the unique solution to (4.7).

This result implies that family $\{u^\epsilon_\cdot\}$ obeys the MDP.

For the simplicity of presentation, we let

$$G^\epsilon_s(a,y,v) := G(a,y,\frac{\sqrt{\epsilon}}{a(\epsilon)}v + u^0_s(y)).$$

To obtain the required conditions on $G^\epsilon_s(a,y,v)$, recall that using the heat kernel, $p_t(x) = \frac{1}{\sqrt{2\pi t}}\exp(-\frac{x^2}{2t})$, SPDE (4.2) has the mild solution form,

$$u^\epsilon_t(y) = \int_{\mathbb{R}} p_t(y-x)F(x)dx + \sqrt{\epsilon}\int_0^t\int_U\int_{\mathbb{R}} p_{t-s}(y-x)G(a,x,u^\epsilon_s(x))dxW(dads) \tag{4.9}$$

By our assumption of $F \in \mathbb{B}_{\alpha,\beta_0}$, we have

$$|u^0_s(y)| \le \int_{\mathbb{R}} p_s(x-y)|F(x)|dx \le Ke^{\beta_0|y|} \tag{4.10}$$

Thus,

$$\begin{aligned}
&\int_U \left|G^\epsilon_s(a,y,v_1) - G^\epsilon_s(a,y,v_2)\right|^2 \lambda(da) \\
= &\int_U \left|G\left(a,y,\frac{\sqrt{\epsilon}}{a(\epsilon)}v_1 + u^0_s(y)\right) - G\left(a,y,\frac{\sqrt{\epsilon}}{a(\epsilon)}v_2 + u^0_s(y)\right)\right|^2 \lambda(da) \\
\le &\; K\left|\frac{\sqrt{\epsilon}}{a(\epsilon)}\right| |v_1 - v_2| \\
\le &\; K|v_1 - v_2|
\end{aligned}$$

since $a(\epsilon)/\sqrt{\epsilon} \to \infty$ so $\sqrt{\epsilon}/a(\epsilon) < 1$ for small enough $\epsilon > 0$. In addition, we have

$$\begin{aligned}
\int_U \left|G^\epsilon_s(a,y,v)\right|^2 \lambda(da) &= \int_U \left|G\left(a,y,\frac{\epsilon}{a(\epsilon)}v + u^0_s(y)\right)\right|^2 \lambda(da) \\
&\le K\left(1 + \left|\frac{\sqrt{\epsilon}}{a(\epsilon)}v + u^0_s(y)\right|^2\right) \\
&\le K\left(1 + v^2 + e^{2\beta_0|y|}\right)
\end{aligned}$$

using (4.10). Therefore, G^ϵ_s satisfies conditions,

$$\int_U \left|G^\epsilon_s(a,y,v_1) - G^\epsilon_s(a,y,v_2)\right|^2 \lambda(da) \le K|v_1 - v_2| \tag{4.11}$$

$$\int_U |G^\epsilon_s(a,y,v)|^2 \lambda(da) \le K(1 + v^2 + e^{2\beta_0|y|}) \tag{4.12}$$

for $y \in \mathbb{R}$ and $v, v_1, v_2 \in \mathbb{R}$ given by (4.5).

Since the proof of the uniqueness of a strong solution to SPDE (4.2) established in [69] only uses condition (4.3) then the same argument can be applied to SPDE (4.6) to achieve uniqueness of a strong solution. SPDE (4.6) can therefore be presented by its mild form,

$$v_t^\epsilon(y) = a(\epsilon) \int_{\mathbb{R}} \int_0^t \int_U G_s^\epsilon(a, x, v_s^\epsilon(x)) p_{t-s}(y-x) W(dads)dx \tag{4.13}$$

We show that this mild solution takes values in $\mathcal{C}([0,1]; \mathbb{B}_\beta)$. To accomplish this we need the subsequent lemma.

Lemma 4.2.1. *For every $n \geq 2$,*

$$\bar{M} = \mathbb{E} \sup_{0 \leq s \leq 1} \left(\int_{\mathbb{R}} |v_s^\epsilon(x)|^2 e^{-2\beta_1 |x|} dx \right)^n < \infty \tag{4.14}$$

Proof. This proof takes analogous steps as were taken for the proof of Theorem 3.5. Similar to the method used in [69], we consider the Hilbert space χ_i composed of all functions f on $\mathbb{R}$ satisfying

$$\|f\|_i^2 = \sum_{k=0}^{i} \int_{\mathbb{R}} f^{(k)}(x)^2 e^{-2\beta_1 |x|} dx < \infty \tag{4.15}$$

with inner product $< \cdot, \cdot >_i$ induced by this norm. As in the proof of Theorem 3.5, we replace $e^{-2\beta_1|x|}$ by $J(x) = \int_{\mathbb{R}} e^{-2\beta_1 |y|} \rho(x-y) dy$ in (4.15). Also for the simplicity of notation, since in this lemma $\epsilon > 0$ is assumed to be fixed, we may assume $\epsilon = a(\epsilon) = 1$. We write v_t^ϵ as v_t and denote $G_s^\epsilon(a, x, v_s^\epsilon(x))$ as $G_s^1(a, x, v_s(x))$. Consider for $n \geq 0$ the sequence,

$$v_t^{n+1}(y) = \int_{\mathbb{R}} \int_0^t \int_U G_s^1(a, x, v_s^{n+1}(x)) p_{t-s}(y-x) W(dads)dx$$

Smoothing out if necessary, we may assume,

$$\sum_{k=0}^{2} \int_{\mathbb{R}} \left(v_t^{n+1}\right)^{(k)} (x)^2 J(x) dx < \infty$$

Applying Itô's formula we have for every $f \in \mathcal{C}_0^\infty(\mathbb{R})$,

$$\begin{aligned} < v_t^{n+1}, f >_0 \ &= \int_0^t \int_U \int_{\mathbb{R}} G_s^1(a, y, v_s^{n+1}(y)) f(y) J(y) dy W(dads) \\ &\quad + \int_0^t < \frac{1}{2} (v_s^{n+1})'', f >_0 ds \end{aligned} \tag{4.16}$$

Itô's formula applied again this time to (4.16) gives,

$$\begin{aligned}
&< v_t^{n+1}, f >_0^2 \\
&= \int_0^t < v_s^{n+1}, f >_0 < (v_s^{n+1})'', f >_0 \, ds \\
&\quad + \int_0^t 2 < v_s^{n+1}, f >_0 \int_U \int_{\mathbb{R}} G_s^1(a, y, v_s^n(y)) f(y) J(y) dy W(dads) \\
&\quad + \int_0^t \int_U \left(\int_{\mathbb{R}} G_s^1(a, y, v_s^n(y)) f(y) J(y) dy \right)^2 \lambda(da) ds
\end{aligned} \tag{4.17}$$

Now we sum over a Complete Orthonormal System (CONS), $\{f_j\}_j$ to obtain,

$$\begin{aligned}
\|v_t^{n+1}\|_0^2 &= \int_0^t < v_s^{n+1}, (v_s^{n+1})'' >_0 \, ds \\
&\quad + \int_0^t \int_U < 2v_s^{n+1}, G_s^1(a, ., v_s^n(.)) >_0 W(dads) \\
&\quad + \int_0^t \int_U \int_{\mathbb{R}} G_s^1(a, y, v_s^n(y))^2 J(y) dy \lambda(da) ds
\end{aligned}$$

By Itô's formula,

$$\begin{aligned}
&\|v_t^{n+1}\|_0^{2p} \\
&= \int_0^t p \|v_s^{n+1}\|_0^{p-1} < v_s^{n+1}, (v_s^{n+1})'' >_0 \, ds \\
&\quad + \int_U \int_0^t p \|v_s^{n+1}\|_0^{p-1} < 2v_s^{n+1}, G_s^1(a, ., v_s^n) >_0 W(dsda) \\
&\quad + \int_0^t p \|v_s^{n+1}\|_0^{p-1} \int_U \int_{\mathbb{R}} G_s^1(a, y, v_s^n(y))^2 J(y) dy \lambda(da) ds \\
&\quad + \int_0^t \frac{1}{2} \int_U p(p-1) \|v_s^{n+1}\|_0^{2(p-2)} < 2v_s^{n+1}, G_s^1(a, ., v_s^n(.)) >_0^2 \lambda(da) ds
\end{aligned}$$

For $v \in \chi_1$,

$$\int_{\mathbb{R}} v(x) v'(x) J'(x) dx = - \int_{\mathbb{R}} v(x) \left(v'(x) J'(x) + v(x) J''(x) \right) dx$$

implying,

$$\begin{aligned}
- \int_{\mathbb{R}} v(x) v'(x) J'(x) dx &= \frac{1}{2} \int_{\mathbb{R}} v(x)^2 J''(x) dx \\
&\le K \int_{\mathbb{R}} v(x)^2 J(x) dx \\
&\le K \|v\|_0^2
\end{aligned}$$

therefore,

$$\begin{aligned}
< v, v'' >_0 &= \int_{\mathbb{R}} v(x)v''(x)J(x)dx \\
&= -\int_{\mathbb{R}} v'(x)\left(v'(x)J(x) + v(x)J'(x)\right)dx \\
&\leq K\|v\|_0^2
\end{aligned}$$

By this fact along with Doob's and Burkholder-Davis-Gundy inequalities we have,

$$\begin{aligned}
&\mathbb{E}\sup_{0\leq s\leq t}\|v_s^{n+1}\|_0^{2p} \\
&\leq K\mathbb{E}\int_0^t \|v_s^{n+1}\|_0^{2p}ds \\
&\quad + K\mathbb{E}\int_0^t\int_U \|v_s^{n+1}\|_0^{2p-2}\int_{\mathbb{R}} v_s^{n+1}(y)G_s^1(a,y,v_s^n(y))J(y)dyW(dads) \\
&\quad + K\mathbb{E}\int_0^t\int_U\int_{\mathbb{R}} \|v_s^{n+1}\|_0^{2p-2}G_s^1(a,y,v_s^n(y))^2J(y)dyds\lambda(da) \\
&\quad + K\mathbb{E}\int_0^t\int_U \|v_s^{n+1}\|_0^{2p-4}\left(\int_{\mathbb{R}} v_s^{n+1}(y)G_s^1(a,y,v_s^n(y))J(y)dy\right)^2 ds\lambda(da)
\end{aligned}$$

Now we apply Hölder's inequality and (4.12) with the observation that $J(y)e^{2\beta_0|y|} < \infty$ to arrive at

$$\begin{aligned}
\mathbb{E}\sup_{0\leq s\leq t}\|v_s^{n+1}\|_0^{2p} &\leq K\mathbb{E}\int_0^t \|v_s^{n+1}\|_0^{2p}ds \\
&\quad + K\mathbb{E}\left(\int_0^t \|v_s^{n+1}\|_0^{4p-2}\left(1+\|v_s^n\|_0^2\right)ds\right)^{1/2} \\
&\quad + K\mathbb{E}\int_0^t \|v_s^{n+1}\|_0^{2p-2}\left(1+\|v_s^n\|_0^2\right)ds
\end{aligned}$$

Letting $f_{n+1}(t) := \mathbb{E}\sup_{0\leq s\leq t}\|v_s^{n+1}\|_0^{2p}$ we obtain,

$$f_{n+1}(t) \leq K\int_0^t f_{n+1}(s)ds + K\int_0^t f_n(s)ds + \frac{1}{2}f_{n+1}(t)$$

and by Gronwall's inequality we achieve,

$$\mathbb{E}\sup_{0\leq s\leq t}\|v_s^{n+1}\|_0^{2p} \leq K\left(\int_0^t \mathbb{E}\|v_s^n\|_0^{2p}ds\right)e^{Kt}$$

So we can use induction on n to conclude,

$$\mathbb{E}\sup_{0\leq s\leq t}\|v_s^n\|_0^{2p} \leq K.$$

□

To obtain our results, we apply Lemma 3.1. For the convenience of the reader, we restate it below.

Lemma 4.2.2. *Let $\{X_t^\epsilon(y)\}$ be a family of random fields and suppose $\beta_1 \in (\beta_0, \beta)$. If there exist constants $n,\ q,\ K > 0$ such that*

$$\mathbb{E}\left|X_{t_1}^\epsilon(y_1) - X_{t_2}^\epsilon(y_2)\right|^n \leq K e^{n\beta_1(|y_1|\vee|y_2|)}\left(|y_1 - y_2| + |t_1 - t_2|\right)^{2+q}, \tag{4.18}$$

then there exists a constant $\alpha > 0$ such that

$$\sup_{\epsilon>0}\mathbb{E}\left|\sup_m \sup_{t_i\in[0,1],|y_i|\leq m, i=1,2} \frac{\left|X_{t_1}^\epsilon(y_1) - X_{t_2}^\epsilon(y_2)\right|}{\left(|y_1 - y_2| + |t_1 - t_2|\right)^\alpha} e^{-\beta m}\right|^n < \infty. \tag{4.19}$$

As a consequence $X_\cdot^\epsilon \in \mathcal{C}\left([0,1];\mathbb{B}_\beta\right)$ a.s. Furthermore, if condition (4.18) holds and $\sup_{\epsilon>0}\mathbb{E}\left|X_{t_0}^\epsilon(y_0)\right|^n < \infty$ for some $(t_0, y_0) \in [0,1]\times\mathbb{R}$, then

$$\sup_{\epsilon>0}\mathbb{E}\left|\sup_{(t,y)\in[0,1]\times\mathbb{R}} e^{-\beta|y|}|X_t^\epsilon(y)|\right|^n < \infty. \tag{4.20}$$

and the family $\{X_\cdot^\epsilon\}$ is tight in $\mathcal{C}\left([0,1];\mathbb{B}_\beta\right)$.

□

Lemma 4.2.3. *The solution to SPDE (4.6) takes values in $\mathcal{C}([0,1];\mathbb{B}_\beta)$.*

Proof. Because of the uniqueness of a strong solution, it is sufficient to prove mild solution (4.13) takes values in $\mathcal{C}([0,1];\mathbb{B}_\beta)$. First we need the following inequalities established in chapter three.

$$P_1 := p_{t-s}(y_1 - x) - p_{t-s}(y_2 - x), \tag{4.21}$$

$$P_2 := p_{t_1-s}(y - x) - p_{t_2-s}(y - x), \tag{4.22}$$

$$\int_{\mathbb{R}} |P_1|^2 e^{2\beta_1|x|}dx \leq K e^{2\beta_1(|y_1|\vee|y_2|)}(t-s)^{-(\frac{1}{2}+\alpha)}|y_1 - y_2|^\alpha, \tag{4.23}$$

$$\int_0^{t_1}\int_{\mathbb{R}}|P_2|^2 e^{2\beta_1|x|}dxds \le Ke^{2\beta_1|y|}|t_1-t_2|^{\alpha}, \tag{4.24}$$

and

$$\int_{t_1}^{t_2}\int_{\mathbb{R}} p_{t_2-s}^2(y-x)e^{2\beta_1|x|}dxds \le K|t_1-t_2|^{\alpha/2}e^{2\beta_1|y|}. \tag{4.25}$$

We proceed by demonstrating two cases. In case one, we fix $t\in[0,1]$ and let $y_1,y_2\in\mathbb{R}$ be arbitrary such that $|y_i|\le m$ for all $i=1,2$. Applying Hölder's and Burkholder-Davis-Gundy inequalities, for $n>0$ we obtain,

$$\begin{aligned}
& \mathbb{E}\left|v_t^{\epsilon}(y_1)-v_t^{\epsilon}(y_2)\right|^n \\
= \;& \mathbb{E}\left|a(\epsilon)\int_0^t\int_U\int_{\mathbb{R}} P_1 G_s^{\epsilon}(a,x,v_s^{\epsilon}(x))dxW(dads)\right|^n \\
\le \;& K\mathbb{E}\left(a(\epsilon)^2\int_0^t\int_U\left(\int_{\mathbb{R}} P_1 G_s^{\epsilon}(a,x,v_s^{\epsilon}(x))dx\right)^2 dads\right)^{n/2} \\
\le \;& K\mathbb{E}\left(a(\epsilon)^2\int_0^t\int_U\int_{\mathbb{R}} P_1^2 e^{2\beta_1|x|}dx\int_{\mathbb{R}} G_s^{\epsilon}(a,x,v_s^{\epsilon}(x))^2 e^{-2\beta_1|x|}dx\lambda(da)ds\right)^{n/2} \\
\le \;& K\mathbb{E}\left(a(\epsilon)^2\int_0^t\int_{\mathbb{R}} P_1^2 e^{2\beta_1|x|}dx\int_{\mathbb{R}}\left(1+v_s^{\epsilon}(x)^2+e^{2\beta_0|x|}\right)e^{-2\beta_1|x|}dxds\right)^{n/2}
\end{aligned}$$

By (4.23) we have,

$$\begin{aligned}
& \mathbb{E}|v_t^{\epsilon}(y_1)-v_t^{\epsilon}(y_2)|^n \\
\le \;& K\mathbb{E}\left(\int_0^t Ke^{2\beta_1(|y_1\vee y_2|)}(t-s)^{-(\frac{1}{2}+\alpha)}|y_1-y_2|^{\alpha}\left(\int_{\mathbb{R}} v_s^{\epsilon}(x)^2 e^{-2\beta_1|x|}dx\right)ds\right)^{n/2} \\
\le \;& \bar{M}Ke^{n\beta_1(|y_1|\vee|y_2|)}|y_1-y_2|^{\frac{n\alpha}{2}}
\end{aligned} \tag{4.26}$$

For the second case, we consider $y\in\mathbb{R}$ to be fixed and assume $t_1,t_2\in[0,1]$ be arbitrary, then by (4.24) and (4.25),

$$
\begin{aligned}
& \mathbb{E}\left|v_{t_1}^{\epsilon}(y)-v_{t_2}^{\epsilon}(y)\right|^n \\
\leq\ & K\mathbb{E}\left|a(\epsilon)\int_0^{t_1}\int_U\int_{\mathbb{R}} P_2 G_s^{\epsilon}(a,x,v_s(x))dxW(dads)\right|^n \\
& +K\mathbb{E}\left|a(\epsilon)\int_{t_1}^{t_2}\int_U\int_{\mathbb{R}} p_{t_2-s}(y-x)G_s^{\epsilon}(a,x,v_s(x))dxW(dads)\right|^n \\
\leq\ & K\mathbb{E}\left|\int_0^{t_1}\int_{\mathbb{R}} P_2 e^{2\beta_1|x|}dx\int_{\mathbb{R}}\left(K+v_s^{\epsilon}(x)^2\right)e^{-2\beta_1|x|}dxds\right|^{\frac{n}{2}} \\
& +K\mathbb{E}\left|\int_{t_1}^{t_2}\int_{\mathbb{R}} p_{t_2-s}(y-x)^2 e^{2\beta_1|x|}dx\int_{\mathbb{R}}\left(K+v_s^{\epsilon}(x)^2\right)e^{-2\beta_1|x|}dxds\right|^{\frac{n}{2}} \\
\leq\ & \bar{M}K\left|\int_0^{t_1}\int_{\mathbb{R}} P_2 e^{2\beta_1|x|}dx\right|^{n/2}+\bar{M}K\left|\int_{t_1}^{t_2}\int_{\mathbb{R}} p_{t_2-s}(y-x)^2 e^{2\beta_1|x|}dxds\right|^{n/2} \\
\leq\ & Ke^{n\beta_1|y|}|t_1-t_2|^{\frac{\alpha n}{2}}+Ke^{n\beta_1|y|}|t_1-t_2|^{\frac{\alpha n}{4}} \\
\leq\ & Ke^{n\beta_1|y|}|t_1-t_2|^{\frac{\alpha n}{4}}
\end{aligned}
\tag{4.27}
$$

By inequalities (4.26) and (4.27) we fulfill the assumptions for Lemma 4.2.2, and so we achieve the conclusion of the current lemma. □

We prove the MDP for the general SPDE studied in chapter three by applying the same method as in the proof of its LDP in Section 3.1. More precisely, we apply Theorem 3.3 in that section which is given by [7]. Here $\mathcal{E}_0 := \mathbb{B}_{\alpha,\beta_0}$, $\mathcal{E} := \mathcal{C}([0,1];\mathbb{B}_\beta)$, and $g^\epsilon := v_t^\epsilon$. We repeat our procedure of converting our SPDE, v_t^ϵ, to the form given as an infinite sum of independent standard Brownian motions. SPDE (4.6) can then be written as

$$v_t^\epsilon(y) = a(\epsilon)\sum_j\int_0^t G_s^{\epsilon,j}(y,v_s^\epsilon(y))dB_s^j + \frac{1}{2}\int_0^t \Delta v_s^\epsilon(y)ds \tag{4.28}$$

where

$$G_s^{\epsilon,j}(y,v) := \int_U G_s^\epsilon(a,y,v)\phi_j(a)\lambda(da) \tag{4.29}$$

Also the controlled PDE (4.7), can be written as

$$v_t(y) = \sum_j\int_0^t\int_U G(a,y,u_s^0(y))k_s^j\phi_j(a)\lambda(da)ds + \frac{1}{2}\int_0^t \Delta v_s(y)ds \tag{4.30}$$

where

$$k_s^j := \int_U h_s(a)\phi_j(a)\lambda(da).$$

We again use the inequalities below for the transformations.

$$x = \sum_j < x, \phi_j > \phi_j \tag{4.31}$$

$$< x, y > = \sum_j < x, \phi_j >< y, \phi_j > \tag{4.32}$$

Now we define

$$\mathcal{S}^N(\ell_2) := \{k \in L^2([0,1], \ell_2) : \int_0^1 \|k_s\|_{\ell_2}^2 \leq N\}$$

To verify the assumption imposed by [7] for their Theorem 6, let $\{k^\epsilon\}$ be a family of random variables taking values in $\mathcal{S}^N(\ell_2)$ such that $k^\epsilon \to k$ in distribution as $\epsilon \to 0$ and consider the SPDE,

$$\begin{aligned} v_t^{\theta,\epsilon}(y) &= \theta \sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x) G_s^{\epsilon,j}(x, v_s^\epsilon(x)) dx dB_s^j \\ &\quad + \sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x) G_s^{\epsilon,j}(x, v_s^\epsilon(x)) k_s^{\epsilon,j} dx ds \end{aligned} \tag{4.33}$$

We establish the tightness of $\{v^{\theta,\epsilon}\}$ as follows.

Lemma 4.2.4. *$v_t^{\theta,\epsilon}(y)$ is tight in $\mathcal{C}([0,1], \mathbb{B}_\beta)$.*

Proof. According to Lemma 4.2.3, we have tightness for the first term of (4.33). Therefore, it is sufficient to prove tightness for

$$w_t^{\epsilon,\theta}(y) := \sum_j \int_0^t \int_{\mathbb{R}} p_{t-s}(y-x) G_j(x, u_s(x)) k_s^{\epsilon,j} dx ds.$$

By the same method used in the proof of Theorem 3.5, we begin by fixing $t \in [0,1]$ and assuming y_1, y_2 to be any real numbers such that $|y_i| \leq m$ for $i = 1, 2$. Let

$$P_1 := p_{t-s}(y_1 - x) - p_{t-s}(y_2 - x).$$

With the help of Cauchy-Schwartz inequality, (4.32), and our result (4.23), we obtain the estimate below,

$$\begin{aligned}
& \mathbb{E}\left|w_t^{\theta,\epsilon}(y_1) - w_t^{\theta,\epsilon}(y_2)\right|^n \\
= \ & \mathbb{E}\left|\int_0^t \int_{\mathbb{R}} P_1 \sum_j G_s^{\epsilon,j}(x, v_s(x)) k_s^{\epsilon,j} dxds\right|^n \\
\leq \ & \mathbb{E}\left|\int_0^t \int_{\mathbb{R}} |P_1| \left(\sum_j G_s^{\epsilon,j}(x, v_s(x))^2\right)^{1/2} \|k_s^\epsilon\|_{\ell_2} dxds\right|^n \\
\leq \ & \mathbb{E}\left|\left(\int_0^t \left(\int_{\mathbb{R}} |P_1| \sqrt{K\left(1 + v_s(x)^2 + e^{2\beta_0|x|}\right)} dx\right)^2 ds\right)^{1/2} \left(\int_0^t \|k_s^\epsilon\|_{\ell_2}^2 ds\right)^{1/2}\right|^n \\
\leq \ & \mathbb{E}\left|\int_0^t \left(\int_{\mathbb{R}} |P_1| \sqrt{K\left(1 + v_s(x)^2 + e^{2\beta_0|x|}\right)} dx\right)^2 dx\right|^{n/2} N^{n/2} \\
\leq \ & K e^{n\beta_1(|y_1| \vee |y_2|)} |y_1 - y_2|^{\frac{\alpha n}{2}}
\end{aligned}$$

Furthermore, the case for $0 < t_1 < t_2$ arbitrary and $y \in \mathbb{R}$ fixed similarly can be given by

$$\begin{aligned}
& \mathbb{E}\left|w_{t_1}^{\theta,\epsilon}(y) - w_{t_2}^{\theta,\epsilon}(y)\right|^n \\
& \leq K\mathbb{E}\left|\int_{t_1}^{t_2} \int_{\mathbb{R}} P_2 \sum_j G_s^{\epsilon,j}(x, v_s(x)) k_s^{\epsilon,j} dxds\right|^n \\
& +K\mathbb{E}\left|\int_0^{t_1} \int_{\mathbb{R}} p_{t_2-s}^2(y-x) \sum_j G_s^{\epsilon,j}(x, v_s(x)) k_s^{\epsilon,j} dxds\right|^n \\
& \leq K e^{n\beta_1|y|} |t_1 - t_2|^{\frac{n\alpha}{2}}
\end{aligned}$$

where,

$$P_2 := p_{t_1-s}(y-x) - p_{t_2-s}(y-x).$$

Thus, $\{v_t^{\theta,\epsilon}\}$ is tight and for Assumption 3.1 to be satisfied we let $\theta = 0$ for its first part and $\theta = a(\epsilon)$ for the second part and apply the Prohorov Theorem. □

Thus, both parts of their assumption are fulfilled and so by Theorem 3.3, our Theorem can be deduced.

4.3 Moderate Deviations for SBM and FVP

Here we strive to achieve the MDP for SBM and FVP. We note that since the transitive property of solutions to SBM's controlled PDE is not needed, then we can achieve the MDP for SBM as well as for FVP. Let

$$\omega_t^\epsilon(dy) := \frac{a(\epsilon)}{\sqrt{\epsilon}} \left(\mu_t^\epsilon(dy) - \mu_t^0(dy)\right) \tag{4.34}$$

where $u_t^\epsilon(y) := \int_0^y \mu_t^\epsilon(dx)$. Hence we can write $v_t^\epsilon(y) := \int_0^y \omega_t^\epsilon(dx)$. Since transitive property is not needed in moderate deviation setting, then for SBM we can prove Lemma 3.3.2 with $\mathcal{P}_\beta(\mathbb{R})$ replaced by $\mathcal{M}_\beta(\mathbb{R})$ in map ξ and follow a similar proof. Also the continuity of $\tilde{\eta} : \mathcal{C}\left([0,1];\mathbb{B}_\beta\right) \to \mathcal{C}\left([0,1];\mathcal{P}_\beta(\mathbb{R})\right)$ can be shown with analogous reasoning to lemma 3.3.3 and for the case of $\mathcal{M}_\beta(\mathbb{R})$ as well. Therefore, noting (4.34) defined based on μ_t, we have $\{\omega_t^\epsilon\}$ is continuous in $\mathcal{C}\left([0,1];\mathbb{B}_\beta\right)$ and since the LDP was proved in previous section for $\{v_t^\epsilon\}$, the contraction principle can be applied to derive LDP for $\{\omega_t^\epsilon\}$, which implies MDP for our two population models. Our remaining task is to identify an explicit representation of their rate functions. Using the same spaces, H_{ω_0} and $\tilde{H}_{\omega_0}$ as for the LDP rate functions, we have the following.

Theorem 4.3.1. *If $\omega_0 \in \mathcal{M}_\beta(\mathbb{R})$ such that $F \in \mathbb{B}_{\alpha,\beta_0}$ then super-Brownian motion, $\{\mu_t^\epsilon\}$, obeys the MDP in $\mathcal{C}([0,1];\mathcal{M}_\beta(\mathbb{R}))$ with rate function,*

$$I(\mu) = \begin{cases} \frac{1}{2}\int_0^1 \int_{\mathbb{R}} \left| \dfrac{d\left(\dot{\omega} - \frac{1}{2}\Delta^* \omega_t\right)}{d\left(\mu_t^0\right)}(y) \right|^2 \mu_t^0(dy)dt & \text{if } \mu \in H_{\omega_0} \\ \infty & \text{otherwise.} \end{cases} \tag{4.35}$$

Proof. Recall for SBM, the general SPDE (4.2) has the following properties,

$$U = \mathbb{R}, \quad \lambda(da) = da, \quad G(a,y,u) = 1_{0\le a\le u} + 1_{u\le a\le 0}$$

then using the controlled PDE (4.7) we have,

$$
\begin{aligned}
& < \omega_t, f > = < \partial_x v_t, f > = - < v_t, f' > \\
= & -\int_0^t \int_0^\infty \int_0^{u_s^0(y)} h_s(a) f'(y) dadyds - \int_0^t \int_{-\infty}^0 \int_{u_s^0(y)}^0 h_s(a) f'(y) dadyds \\
& - \int_0^t < \frac{1}{2}\Delta v_s, f' > ds \\
= & \int_0^t \int_0^\infty h_s(a) f((u_s^0)^{-1}(a)) dads - \int_0^t \int_{-\infty}^0 h_s(a) f((u_s^0)^{-1}(a)) dads \\
& + \int_0^t < \frac{1}{2}\omega_s, \Delta f > ds \\
= & \int_0^t \int_0^\infty h_s(u_s^0(y)) f(y) du_s^0(y) ds - \int_0^t \int_{-\infty}^0 h_s(u_s^0(y)) f(y) du_s^0(y) ds \\
& + \int_0^t < \frac{1}{2}\Delta^* \omega_s, f > ds \\
= & \int_0^t < h_s(u_s^0) sgn(.) \mu_s^0, f > ds + \frac{1}{2}\int_0^t < \Delta^* \omega_s, f > ds.
\end{aligned}
$$

Next we sum over a CONS $\{f_j\}$ by using (4.31) and take the derivative with respect to t to obtain,

$$
d\dot{\omega}_t(y) = h_t(u_t^0(y)) sgn(y) d\mu_t^0(y) + \frac{1}{2} d\left(\Delta^* \omega_t\right)(y)
$$

Therefore,

$$
h_t(u_t^0(y)) = \frac{d\left(\dot{\omega}_t - \frac{1}{2}\Delta^* \omega_t\right)(y)}{sgn(y)\mu_t^0(y)}
$$

Notice that

$$
\int_{\mathbb{R}} |h_t(a)|^2 da = \int_{\mathbb{R}} |h_t(u_t^0(y))|^2 du_t^0(y) = \int_{\mathbb{R}} |h_t(u_t^0(y)) sgn(y)|^2 d\mu_t(y)
$$

Thus, if $I(v) < \infty$ then $I(v)$ given in (4.8) with $U = \mathbb{R}$ is equal to the right hand side of (4.35) and so the proof is completed. □

As noted in the introduction, Schied in [55] also established the MDP for SBM. To be complete, we provide his theorem below. First we give an overview of his notation. Schied uses X_t to denote SBM, taking values in the space of positive finite measures on $\mathbb{R}^d$ equipped with the usual weak topology denoted as $M^+(\mathbb{R}^d)$ and considers the space $\mathcal{C}\left([0,1]; M(\mathbb{R}^d)\right)$. Given $\mu \in M^+(\mathbb{R}^d)$, in order to give the rate function he defines space H_μ

as the set of all $\omega \in \mathcal{C}\left([0,1]; M(\mathbb{R}^d)\right)$ satisfying the form $\omega(t) = \int_0^t \dot{\omega}(s)ds$ for some mapping $\dot{\omega} : [0,1] \to M(\mathbb{R}^d)$ with $\dot{\omega}(t) << \mu$ for almost all t. To achieve the MDP he centers the process as follows,

$$\hat{X}_t = X_t - X_0 P_t$$

where P_t is the Brownian semigroup. His theorem is as follows.

Theorem 4.3.2 (Theorem 1.1 in [55]). *Suppose $\beta : (0,1] \to (0,1]$ is a function such that $\beta(\alpha) \to 0$ as $\alpha \to 0$. Then the distributions of the processes $\hat{X}_t^{\alpha,\beta(\alpha)} := \beta(\alpha)^{-1}\hat{X}_{\alpha\beta(\alpha)^2 t}$ for $0 \le t \le 1$ satisfy a large deviation principle on $\mathcal{C}\left([0,1]; M(\mathbb{R}^d)\right)$ with scale α and good rate function I_μ given by*

$$I_\mu(\omega) = \begin{cases} \frac{1}{4}\int_0^1 \| \frac{d\dot{\omega}(t)}{d\mu} \|^2_{L^2(\mu)} dt & \text{if } \omega \in H_\mu \\ \infty & \text{otherwise.} \end{cases} \tag{4.36}$$

i.e. if $U \subset \mathcal{C}\left([0,1]; M(\mathbb{R}^d)\right)$ is open then

$$\liminf_{\alpha \to 0} \alpha \log P_\mu \left(\hat{X}^{\alpha,\beta(\alpha)} \in U \right) \ge - \inf_{\omega \in U} I_\mu(U)$$

if $A \subset \mathcal{C}\left([0,1]; M(\mathbb{R}^d)\right)$ is closed then

$$\limsup_{\alpha \to 0} \alpha \log P_\mu \left(\hat{X}^{\alpha,\beta(\alpha)} \in A \right) \le - \inf_{\omega \in A} I_\mu(\omega)$$

The level sets $\{I_\mu \le c\}$ are compact for each $c \ge 0$.

□

Therefore, Schied proves the large deviations for the process

$$\hat{X}_t^{\alpha\beta(\alpha)} := \frac{1}{\beta(\alpha)} \left(X_{\alpha\beta(\alpha)^2 t} - X_0 P_{\alpha\beta(\alpha)^2 t} \right)$$

which implies the moderate deviations for SBM, $X_{\alpha\beta(\alpha)^2 t}$. Recall the martingale problem characterization of SBM from chapter two: for all $f \in \mathcal{C}_b^2(\mathbb{R})$,

$$M_t(f) := \langle \mu_t^\epsilon, f \rangle - \langle \mu_0^\epsilon, f \rangle - \int_0^t \left\langle \mu_s^\epsilon, \frac{1}{2}\Delta f \right\rangle ds$$

is a martingale with quadratic variation,

$$\langle M(f) \rangle_t = \epsilon \int_0^t \left\langle \mu_s^\epsilon, f^2 \right\rangle ds$$

with branching rate ϵ and μ_t^ϵ denoting the SBM. Hence if we take $\alpha = \sqrt{\epsilon}$ and $\beta(\alpha) = \sqrt[4]{\epsilon}$ we have,

$$\begin{aligned} M_{\alpha\beta(\alpha)^2}(f) = M_{\epsilon t}(f) &= \langle X_{\epsilon t}, f\rangle - \langle X_0, f\rangle - \int_0^{\epsilon t}\left\langle X_s, \frac{1}{2}\Delta f\right\rangle ds \\ &= \langle X_{\epsilon t}, f\rangle - \langle X_0, f\rangle - \epsilon\int_0^{\epsilon t}\left\langle X_{\epsilon s}, \frac{1}{2}\Delta f\right\rangle ds \end{aligned}$$

with quadratic variation,

$$\langle M(f)\rangle_{\epsilon t} = \epsilon\int_0^t \left\langle X_s, f^2\right\rangle ds \tag{4.37}$$

Therefore with the above choice of α and $\beta(\alpha)$ and observing (4.37), the SBM considered by Schied has the same branching rate, ϵ, as in our case. The difference in the rate function compared to our rate, is a consequence of the term, $\epsilon\int_0^t \left\langle X_{s\epsilon}, \frac{1}{2}\Delta f\right\rangle ds$ in his setting which vanishes when $\alpha = \sqrt{\epsilon}$ is set to go to zero. Thus, the added diffusion term does not appear in his rate function as it does for ours in the following theorem.

Similarly for FVP, since FVP satisfies the general SPDE (4.2) with

$$U = [0,1], \quad \lambda(da) = da, \quad G(a,y,u) = 1_{a<u} - u$$

then we aim to prove the following.

Theorem 4.3.3. *Suppose $\omega_0 \in \mathcal{P}_\beta(\mathbb{R})$ such that $F \in \mathbb{B}_{\alpha,\beta_0}$. Then, Fleming-Viot Process, $\{\mu^\epsilon\}$, satisfies the MDP on $\mathcal{C}([0,1];\mathcal{P}_\beta(\mathbb{R}))$ with rate function,*

$$I(\mu) = \begin{cases} \frac{1}{2}\int_0^1\int_{\mathbb{R}}\left|\frac{d\left(\dot{\omega}_t - \frac{1}{2}\Delta^*\omega_t\right)}{d\left(\mu_t^0\right)}(y)\right|^2 \mu_t^0(dy)dt & \text{if } \mu\in\tilde{H}_{\omega_0} \\ \infty & \text{otherwise.}\end{cases} \tag{4.38}$$

Proof. By the conditions on U and $G(a,y,u)$ for the case of FVP we have,

$$\begin{aligned} & < \omega_t, f >= - < v_t, f' > \\ = & -\int_0^t\int_{\mathbb{R}}\int_0^{u_s^0(y)} h_s(a)f'(y)dadyds + \int_0^t\int_{\mathbb{R}}\int_0^1 u_s^0(y)h_s(a)f'(y)dadyds \\ & -\int_0^t < \frac{1}{2}v_s''(y), f' > ds \\ = & \int_0^t < h_s(u_s^0)\mu_s^0, f > ds - \int_0^t < \int_0^1 h_s(a)da\mu_s^0, f > ds \\ & + \int_0^t < \frac{1}{2}\Delta^*\omega_s, f > ds. \end{aligned}$$

Now taking the sum and derivative we have,

$$d\dot{\omega}_t(y) - \frac{1}{2}d\omega_t(y) = h_t(u_t^0(y))d\mu_t^0(y) - \int_0^1 h_t(a)dad\mu_t^0(y)$$

Our goal is to find the infimum of $\int_0^1 |h_s(a)|^2 da$ based on $h_s(a)$. We note that if h satisfies (4.7) then $g_s(a) := h_s(a) - \int_0^1 h_s(a)da$ also satisfies the same equation. So we consider $h_s(a)$ instead of $g_s(a)$ and use the form,

$$\int_0^1 |h_s(a)|^2 da = \int_0^1 \left| \frac{d\left(\dot{\omega}_t - \frac{1}{2}\Delta^*\omega_t\right)(y)}{d\mu_t^0(y)} \right|^2 d\mu_t^0(y)$$

in (4.8) to arrive at (4.38) for the case $I(v) < \infty$. $\square$

Thus, MDP is proved for the two models.

Bibliography

[1] H. Arani and B. Zangeneh (2010). Large deviation principle for semilinear stochastic evolution equations with monotone nonlinearity and multiplicative noise. *Differential Integral Equations*. vol. 23, no. 7-8. 747-772.

[2] I. Bihari (1956). A generalization of a lemma of Bellman and its application to uniqueness problems of differential equations. *Acta Math. Acad. Sci. Hungar.*. vol. 7, 81-94.

[3] A. Brown and C. Pearcy (1995). *An Introduction to Analysis*. Graduate Texts in Mathematics, Springer.

[4] W. Bryc (1990). Large deviations by the asymptotic value method. Diffusion processes and related problems in analysis, Vol. I, 447-472, *Progr. Probab., 22*, Birkhäuser Boston,MA.

[5] A. Budhiraja and P. Dupuis (2000). A variational representation for positive functionals of infinite dimensional Brownian motion. *Probab. Math. Statist.*. vol. 20. no. 1, 39-61.

[6] A. Budhiraja, J. Chen, and P. Dupuis (2013). Large deviations for stochastic partial differential equations driven by a Poisson random measure. *Stochastic Process. Appl.*. vol. 123, no. 2, 523-560.

[7] A. Budhiraja, P. Dupuis and V. Maroulas (2008). Large Deviations for Infinite Dimensional Stochastic Dynamical Systems. *Ann Probab.*. vol. 36, no. 4, 1390-1420.

[8] R. Carmona, I. Ekeland, A. Kohatsu-Higa, J. Lasry, P. Lions, H. Pham, E. Taflin (2007). *Paris-Princeton lectures on Mathematical Finance 2004*. Lecture Notes in Mathematics, vol. 1919. Springer.

[9] A. Budhiraja, P. Dupuis and V. Maroulas (2011). Variational representations for continous time processes. *Ann. Inst. Henri Poincaré Probab. Stat.*. vol. 47, no. 3, 725-747.

[10] M. Boué and P. Dupuis (1998). A variational representation for certain functionals of Brownian motion. *Ann. Probab.*. vol. 26, no. 4, 1641-1659.

[11] D.A. Dawson (1993). Measure-valued Markov processes. cole d't de Probabilits de Saint-Flour XXI-1991, 1-260, *Lecture Notes in Math.*. 1541, Springer, Berlin.

[12] D.A. Dawson and S. Feng (1998). Large deviations for the Fleming-Viot process with neutral mutation and selection. *Stochastic Process. Appl.*. vol. 77, no. 2, 207-232.

[13] D.A. Dawson and S. Feng (2001). Large deviations for the Fleming-Viot process with neutral mutation and selection. II. *Stochastic Process. Appl.*. vol. 92, no. 1, 131-162.

[14] A. Dembo and O. Zeitouni (2010). *Large Deviations Techniques and Applications*. Springer, Berlin.

[15] J. Deuschel and D. Stroock (1989). *Large Deviations*. vol. 137. Academic Press.

[16] P. Dupuis and R. Ellis (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley Series in Probability and Statistics.

[17] E.B. Dynkin (1994). *An Introduction to Branching Measure-Valued Processes*. American Mathematical Society CRM Monograph Series, Vol. 6.

[18] R. Ellis (1985). *Entropy, large deviations, and statistical mechanics*. A series of Comprehensive studies in Mathematics, vol. 271. Springer.

[19] A. Etheridge (2000). *An Introduction to Superprocesses.* University Lecture Series, vol. 20. American Mathematical Society.

[20] S. Ethier and T. Kurtz (2005). *Markov Processes: Characterization and Convergence.* Wiley Series in Prob. and Stat.

[21] L. Evans (2010). *Partial Diefferential Equations: Second Edition.* Graduate Studies in Mathematics, vol. 19. American Mathematical Society.

[22] S.N. Ethier and T. G. Kurtz (1993). Fleming-Viot Processes in Population Genetics. *SIAM J. Control and Optimization.*.vol. 31, no. 2, 345-386.

[23] S. Feng and J. Xiong (2002). Large deviations and quasi-potential of a Fleming-Viot process. *Electron. Comm. Probab.*. vol. 7, 13-25.

[24] K. Fleischmann,J. Gärtner and I. Kaj (1996). A Schilder type theorem for super-Brownian motion. *Canad. J. Math.* vol. 48, no. 3, 542-568.

[25] K. Fleischmann and I. Kaj (1994). Large deviation probabilities for some rescaled superprocesses. *Ann. Inst. H. Poincar Probab. Statist.* vol. 30 , no. 4, 607-645.

[26] K. Fleischmann and J. Xiong (2005). Large deviation principle for the single point catalytic super-Brownian motion. *Markov Process. Related Fields.* vol 11, no. 3, 519-533.

[27] W.H. Fleming and M. Viot (1979). Some Measure-Valued Markov Processes in Population Genetics Theory. *Indiana Univ. Math.*, Vol. 28, No. 5, 817-843.

[28] M. Freidlin and A. Wentzell (1984). *Random perturbations of dynamical systems. A series of comprehensive studies in Mathematics. vol. 260. Springer.*

[29] W. Hong (1998). Central Limit Theorem for the Occupation Time of Catalytic super-Brownian Motion. Chinese Sci. Bull.. *vol 43, no. 24, 2035-2040.*

[30] W. Hong and Z. Li (1999). A Central Limit Theorem for super-Brownian Motion with super-Brownian Immigration. J. Appl. Probab.. *vol 36, no. 4, 1218-1224.*

[31] W. Hong (2002). Moderate Deviation for super-Brownian Motion with super-Brownian Immigration. J. Appl. Prob..*vol. 39, 829-838.*

[32] W. Hong (2008). Moderate Deviations for the Quenched Mean of the super-Brownian Motion with Random Immigration. Sci. in China Series. *vol 51, no. 3, 343-350.*

[33] O. Kallenberg (2010). Foundations of Modern Probability: Second Edition. *Probability and Its Applications, Springer.*

[34] G. Kallianpur and J. Xiong (1995). Stochastic Differential Equations in Infinite Dimensional Spaces. *Institute of Mathematical Statistics Lecture Notes-Monograph Series.*

[35] *F. Klebaner (2005).* Introduction to Stochastic Calculus with Applications: Second Edition. *Imperial College Press. World Scientific Publishing Co.*

[36] *M. Kimmel and D. Axelrod (2002).* Branching Processes in Biology. *Interdisciplinary Applied Mathematics (IAM) series, Springer.*

[37] *J. LeGall (1999).* Spatial Branching Processes, Random Snakes and Partial Differential Equations. *Lectures in Mathematics, ETH Zürich.*

[38] *Z. Li (2011).* Measure-valued branching Markov processes. *Probability and its Applications (New York). Springer, Heidelberg.*

[39] *W. Liu (2010). Large deviations for stochastic evolution equations with small multiplicative noise.* Appl. Math. Optim.. *vol. 61, no. 1, 27-56.*

[40] *W. Liu, M. Ròckner, and X.Zhu (2013). Large deviation principles for the stochastic quasi-geostrophic equations.* Stochastic Process. Appl.. *vol. 123, no. 8, 3299-3327.*

[41] *V. Maroulas and J. Xiong (2013). Large deviations for optimal filtering with fractional Brownian motion.* Stochastic Process. Appl.. *vol. 123, no. 6, 2340-2352.*

[42] *I. Mitoma (1985). An ∞-dimensional inhomogeneous Langevin equation.* J. Funct. Anal.. *no 61, 342-359.*

[43] *A. Pazy (2011).* Semigroups of Linear Operators and Applications to Partial Differential Equations. *Applied Mathematical Sciences, Vol. 44, Springer.*

[44] *E. Perkins (2002). Dawson-Watanabe superprocesses and measure-valued diffusions.* Lectures on probability theory and statistics (Saint-Flour, 1999). *125-324,Lecture Notes in Math., 1781, Springer, Berlin.*

[45] *S. Ramasubramanian (2008). Large Deviations: An Introduction to 2007 Abel Prize.* Proc. Indian Acad. Sci. . *Vol. 118, No.2, 161-182.*

[46] *J. Ren, S. Xu, X. Zhang (2010). Large deviations for multivalued stochastic differential equations.* J. Theoret. Probab.. *vol. 23, no. 4, 1142-1156.*

[47] *J. Ren and X. Zhang (2005). Freidlin-Wentzell's large deviations for homeomorphism flows of non-Lipschitz SDEs.* Bull. Sci. Math.. *vol. 129, no. 8, 643-655.*

[48] *M. Röckner, T. Zhang, and X. Zhang (2010). Large deviations for stochastic tamed 3D Navier-Stokes equations.* Appl. Math. Optim.. *vol. 61, no. 2, 267-285.*

[49] *H. Rubin and J. Sethuraman (1965). Probabilities of Moderate Deviations.* Sankhyā: The Indian Journal of Statistics. *Vol. 27, No. 2/4, 325-346.*

[50] *H. Rubin and J. Sethuraman (1965). Bayes Risk Efficiency.* Sankhyā: The Indian Journal of Statistics. *Vol. 27, No. 2/4, 347-356.*

[51] *W. Rudin (1991).* Functional Analysis: Second Edition. *International Series in Pure and Applied Mathematics, McGraw-Hill.*

[52] *L. Serlet (1997). A large deviation principle for the Brownian snake.* Stoch. Proc. Appl.. *Vol. 67, 101-115.*

[53] *L.Serlet (2009). New large deviation results for some super-Brownian processes.* Stoch. Proc. Appl.. *Vol. 119. No. 5, 1696-1724.*

[54] *A. Schied (1996). Sample path large deviations for super-Brownian motion.* Probab. Theory Related Fields. *vol 104, no. 3, 319-347.*

[55] *A. Schied (1997). Moderate Deviations and Functional LIL for super-Brownian Motion.* Stochastic Process. Appl. *vol. 72, no 1, 11-25.*

[56] *T. Shiga (1994). Two Contrasting Properties of Solutions for One-Dimensional Stochastic Partial Differential Equations.* Can. J. Math., *vol. 46, no 2, 415-437.*

[57] *A. Shwartz and A. Weiss (1996).* Large deviations for performance analysis: queues, communication and computing. *Stochastic Modeling Series. Chapman and Hall.*

[58] *S. Sritharan and P. Sundar (2006). Large deviations for the two-dimensional Navier-Stokes equations with multiplicative noise.* Stochastic Process. Appl.. *vol. 116, no. 11, 1636-1659.*

[59] *A. Swiech (2009). A PDE approach to large deviations in Hilbert spaces.* Stochastic Process. Appl.. *vol. 119, no. 4, 1081-1123.*

[60] *T. Taniguchi (1992). Successive approximations to solutions of stochastic differential equations.* J. Differential Equations. *vol. 96, no. 1, 152-169.*

[61] *A. Veretennikov (1999). On large deviations in the averaging principle for SDEs with a "full dependence".* Ann. Probab.. *vol. 27, no. 1, 284-296.*

[62] *Z. Wang (2008). Existence and uniqueness of solutions to stochastic Volterra equations with singular kernels and non-Lipschitz coefficients.* Statist. Probab. Lett.. *vol. 78, no. 9, 1062-1071.*

[63] *T. Yamada and S. Watanabe (1971). On the uniqueness of solutions of stochastic differential equations II.* J. Math. Kyoto Univ.. *vol. 11, 553-563.*

[64] *J.B. Walsh (1986). An Introduction to Stochastic Partial Differential Equations. École d'Été de Probabilités de Saint Flour XIV-1984.* Lecture Notes in Math.. *no. 1180, 265-439.*

[65] *L. Wu (1995). Moderate Deviations of Dependent Random Variables Related to CLT.* Annals of Probability. *vol. 23, no. 1, 420-445.*

[66] *K. Xiang (2010). An explicit Schilder-type theorem for super-Brownian motions.* Comm. Pure Appl. Math.. *vol. 63, no. 11, 1381-1431.*

[67] *K. Xiang (2012). An explicit Schilder-type theorem for super-Brownian motions: infinite initial measures.* Trans. Amer. Math. Soc.. *vol. 364, no. 12. 6513-6529.*

[68] *K. Xiang and T. Zhang (2005). Small time asymptotics for Fleming-Viot processes.* Infin. Dimens. Anal. Quantum Probab. Relat. Top.. *vol. 8, no. 4, 605-630.*

[69] *J. Xiong (2013). Super-Brownian motion as the unique strong solution to a SPDE.* Ann. Probab.. *vol. 41, no. 2, 1030-1054.*

[70] *T. Yamada (1981). On the successive approximation of solutions of stochastic differential equations.* J. Math. Kyoto Univ. *. vol. 21, no. 3, 501-515.*

[71] *X. Yang and M. Zhang (2012). Moderate Deviation for the Single Point Catalytic super-Brownian Motion.* Acta Mathematica Sinica*. vol. 28, no. 9, 1799-1808.*

[72] *K. Yosida (1995).* Functional Analysis*. Classics in Mathematics, Springer.*

[73] *M. Zhang (2008). Some Scaled Limit Theorems for an Immigration super-Brownian Motion.* Sci. in China Series*. vol. 51, no. 2, 203-214.*

[74] *X. Yang and M. Zhang (2012). Moderate Deviation for the Single Point Catalytic super-Brownian Motion.* Acta Mathematica Sinica*. vol. 28, no. 9, 1799-1808.*

[75] *M. Zhang (2005). Functional Central Limit Theorem for the super-Brownian Motion with super-Brownian Immigration.* J. Theoret. Probab.*. vol. 18, no. 3. 665-685.*

[76] *M. Zhang (2008). Some Scaled Limit Theorems for an Immigration super-Brownian Motion.* Sci. in China Series*. vol. 51, no. 2, 203-214.*

Appendix

Theorem 4.3.4 (Arzelà-Ascoli). *Let S be a compact metric space, and $\mathcal{C}(S)$ the Banach space of (real- or) complex-valued continuous functions $x(s)$ with norm $\|x\| = \sup_{s\in S}|x(s)|$. Then a sequence $\{x_n(s)\} \subset \mathcal{C}(S)$ has a compact closure in $\mathcal{C}(S)$ if the following two conditions are satisfied:*

i. $x_n(s)$ is equi-bounded in n, namely,

$$\sup_{n\geq 1}\sup_{s\in S}|x_n(S)| < \infty$$

ii. $x_n(s)$ is equi-continuous in n, that is

$$\lim_{\delta\to 0}\sup_{n\geq 1, dis(s',s'')\leq\delta}|x_n(s') - x_n(s'')| = 0$$

For a proof see [72] Section 3.3.

Theorem 4.3.5 (Burkholder-Davis-Gundy Maximal Inequality). *Let M_t be an $\mathcal{F}_t$ continuous martingale, with increasing process $\langle M\rangle_t$. Then for every $m > 0$ there exist universal constants $k_m, K_m > 0$ such that for every stopping time τ,*

$$k_m\mathbb{E}\left(< M >_\tau^m\right) \leq \mathbb{E}\left(\left(\sup_{0\leq t\leq\tau}|M_t|\right)^{2m}\right) \leq K_m\mathbb{E}\left(< M >_\tau^m\right)$$

(cf. [33] Theorem 17.7)

Theorem 4.3.6 (Contraction Principle). *Suppose $\mathcal{X}$ and $\mathcal{Y}$ are Polish spaces, I is a rate function on $\mathcal{X}$ and f is a continuous function mapping $\mathcal{X}$ to $\mathcal{Y}$ then if a sequence $\{X_n\}$ satisfies the Large Deviation Principle on $\mathcal{X}$ with rate function I, then $\{f(X_n)\}$ satisfies the Large Deviation Principle on $\mathcal{Y}$ with rate function,*

$$J(y) = \inf\left\{I(x) : x \in f^{-1}(y)\right\}$$

for $y \in \mathcal{Y}$. (cf. [16] Theorem 1.3.2)

Theorem 4.3.7 (Kolmogorov Criteria). *Let $\{X_t, t \in \mathbb{R}\}$ be a real valued stochastic process. Suppose there are constants $k > 1, K > 0,$ and $\epsilon > 0$ such that for all $s, t \in \mathbb{R}$,*

$$\mathbb{E}\left\{|X_t - X_s|^k\right\} \leq K|t-s|^{n+\epsilon}$$

then

i. X *has a continuous version,*

ii. there exist constants C, γ *depending only on* n, k, ϵ *and a random variable* Y *such that with probability one, for all* $s, t \in \mathbb{R}$,

$$|X_t - X_s| \leq Y|t-s|^{\epsilon/k} \left(\log \frac{\gamma}{|t-s|}\right)^{2/k}$$

and

$$\mathbb{E}\left(Y^k\right) \leq CK$$

iii. if $\mathbb{E}(|X_t|^k) < \infty$ *for some* t, *then*

$$\mathbb{E}\left(\sup_{t\in\mathbb{R}} |X_t|^k\right) < \infty.$$

(cf. Walsh [64] Corollary 1.2)

Theorem 4.3.8 (Levy's Characterization of Brownian motions)**.** *A process* M *with* $M(0) = 0$ *is a Brownian motion if and only if it is a continuous local martingale with quadratic variation process* $[M, M](t) = t$. *(cf. Theorem 7.36 in [35])*

Theorem 4.3.9 (Prohorov's Theorem)**.** *Suppose* $\mathcal{E}$ *is a Polish space, then a sequence* $\{\mu\} \subset \mathcal{M}(\mathcal{R})$ *is tight if and only if it has a subsequence that converges weakly.*
(cf. [33] Theorem 16.3)

Definition 4.3.1 (Quotient Space)**.** *Let* $\sim$ *be an equivalence relation on a space* X, *that partition* X *into equivalence classes defined as* $[x] = \{y \in X : y \sim x\}$ *for an element* $x \in X$. *The space* $X/\sim$ *formed by these equivalence classes is called the quotient space of* X *where by a map* $p : X \to X/\sim$, *called the natural projection map each element* $x \in X$ *is mapped to the equivalence class of relation* $\sim$ *containing it. Therefore, the elements of the quotient space* $X/\sim$ *are equivalent classes. Quotient spaces are pseudo-metric spaces. Pseudo-metric has all the properties of a metric except positive definite property, that is* $d(x, y) = 0$ *does not necessarily imply that* $x = y$. *If* X *is a Polish sapce with metric* d, *then the pseudo-metric* $\tilde{d}$ *on* $X/\sim$ *is defined as follows: for* $\tilde{x}, \tilde{y} \in X/\sim$,

$$\tilde{d}(\tilde{x}, \tilde{y}) = \inf \sum_{i=1}^{n} d(p_i, q_i)$$

where the infimum is taken over all finite sequences $(p_1, ..., p_n)$ and $(q_1, ..., q_n)$ such that $p_1 \sim \tilde{x}$, $q_n \sim \tilde{y}$ and $q_i \sim p_{i+1}$ for $i = 1, ..., n-1$.

A more general formulation of quotient space is formed by a subspace of the original space rather than by an equivalence relation. More precisely, for a subspace $N \subset X$, for every $x \in X$ let $\pi(x)$ be the coset of N that contains x. Namely, $\pi(x) = x + N$ then these cosets are elements of the quotient space X/N. The map $\pi : X \to X/N$ is linear with $\pi(0) = N$. For more information on quotient space formed by an equivalence relation, we refer the reader to [3] and for the quotient space formed by a subspace [51] is recommended.

Definition 4.3.2 (Semigroup)**.** *On a Banach space X, a one-parameter family $\{T(t)\}_{0 \le t < \infty}$ of bounded linear operator from X into X is a semigroup if*

a. $T(0) = I$ where I is the identity operator on X

b. (semigroup property) $T(t+s) = T(t)T(s)$ for every $t, s \ge 0$

a semigroup is uniformly continuous if

$$\lim_{t \searrow 0} \|T(t) - I\| = 0$$

and is a strongly continuous semigroup (called a C_0 semigroup) if

$$\lim_{t \searrow 0} T(t)x = x$$

for all $x \in X$.

For every semigroup, there is an associated infinitesimal generator denoted by A and defined as

$$Ax = \lim_{t \searrow 0} \frac{T(t)x - x}{t}$$

for $x \in D(A)$, $D(A)$ being the domain of A:

$$D(A) = \left\{ x \in X : \lim_{t \searrow 0} \frac{T(t)x - x}{t} \text{ exists} \right\}$$

For more information see [43].

Definition 4.3.3 (Strong, Weak and Mild Solution)**.** *Consider the system*

$$\begin{cases} \frac{du(t)}{dt} = Au(t) + f(t) \\ u(0) = x \end{cases} \tag{4.39}$$

where A is the infinitesimal generator of a C_0 semigroup $\{T(t)\}$ that corresponds to the homogeneous equation. A strong solution, also called a classical solution, is a function u that is differentiable a.e. on $[0,T]$, $u' \in L^1(0,T;X)$ and (4.39) is satisfied.

If $\{T(t)\}$ is not differentiable, then in general if $x \notin D(A)$, then the system does not have a solution and $t \to T(t)x$ is a generalized solution referred to as a mild solution. To determine the mild solution of (4.39), let $g(s) = T(t-s)u(s)$ then since $T(t-s) = e^{A(t-s)}$,

$$
\begin{aligned}
\frac{dg(s)}{ds} &= T(t-s)u'(s) - AT(t-s)u(s) \\
&= T(t-s)Au(s) + T(t-s)f(s) - AT(t-s)u(s) \\
&= T(t-s)Au(s)
\end{aligned}
$$

if $f \in L^1(0,T;X)$ then $T(t-s)f(s)$ is integrable so integrating from 0 to t and using the initial condition and the fact that $T(0) = I$ gives

$$u(t) = T(t)x + \int_0^t T(t-s)f(s)ds \tag{4.40}$$

which is called a mild solution of (4.39). A strong solution is also a mild solution but the converse requires the uniqueness of strong solution. More precisely, if the system (4.39) has a unique strong solution, then (4.40) is equivalent to (4.39).
See [43] for more information.

Definition 4.3.4 (Tight and Exponential Tight)**.** *For a sequence of probability measures, $\{\mu_n\}$,*
i. *$\{\mu_n\}$ is tight if there exists a compact set K such that*

$$\sup_n \mu_n(K^c) < \epsilon$$

for every $\epsilon > 0$.
ii. *$\{\mu_n\}$ is exponentially tight if for each $\alpha \in (0,\infty)$, there exists a compact set K such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(K^c) \leq -\alpha$$

Vita

The author's childhood years until middle school were spent in Iran, where she was born. In her opinion, the Iranian educational system places a great emphasis on science and mathematics. To this day, she feels comfortable with performing simple calculations such as multiplication in persian numbers in her head and quickly translating them to English. On the other hand, certain concepts, which she learned in America, such as those in trigonometry, are pronounced in English in her mind. Parisa began attending schools in America in seventh grade, taking Honors Algebra, for which she scored a 99 on its exit exam. Later she was admitted to a nationally high ranked high school, Martin Luther King Magnet High School.

Attending Martin Luther King High school proved to be a very valuable experience to the author. It offered her the opportunity to take AP classes such as AP Physics and AP Calculus. Among other things, she believes that her GPA from this high school and scoring 4 out of 5 on AP Calculus granted her admission to Belmont University, where she spent her undergraduate studies. Parisa began attending this university in the fall of 2003 with a major in Chemistry. However, in her second year after completing one semester at Belmont without a challenging mathematics course, it became apparent to her that her major should be mathematics. In that semester, she realized that her education without mathematics was not complete and that she could not leave what mathematics offers her. She later completed a Bachelor of Science in this field with a minor in Chemistry.

After graduating from Belmont University in 2007, Parisa was admitted to the University of Tennessee as a Ph.D. candidate. There she took many classes in different fields of mathematics including Analysis, Numerical Analysis, Partial Differential Equations, Probability, and Topology. The author chose to specialize in Probability theory since it acts as a bridge between pure and applied mathematics and involves concepts and ideas from both. Probability theory offers her great depth and understanding which she truly values in mathematics. Also she taught many classes in Basic Calculus, Elementary Statistics, Precalculus and had the opportunity to teach Calculus I and II in year 2012-2013. For every year she taught at the University of Tennessee, she was nominated for the Best Graduate Student Teaching Award.

Learning from all areas in knowledge is very enjoyable to the author. Mathematics has always been her main reason to attend school. It offers her great depth of understanding and helps her have a clear mind. The author wishes to continue her work and studies in Mathematics and contribute to its advancement.