1

PM-1 3" x 4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Responsible: Bibliothèque

Outlets: Bibliothèque

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Large Deviations of Particle Systems

by

Shui Feng, B.Sc., M.Sc.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of
Doctor of Philosophy

Department of Mathematics and Statistics
Ottawa-Carleton Institute for Graduate Studies and Research
in Mathematics and Statistics

Carleton University
Ottawa, Ontario
September, 1992

© Copyright
1993, Shui Feng

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-84043-9

Canada

The undersigned hereby recommend to
The Faculty of Graduate Studies and Research
acceptance of the thesis,
Large Deviations of Particle System
submitted by
Shui Feng, B.Sc., M.Sc.
in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

Dr. C. W. L. Garner, Chair,
Department of Mathematics and Statistics

Dr. D. A. Dawson, Thesis Supervisor

Dr. J. Gärtner
External Examiner

Carleton University

1992

Abstract

Systems of birth and death processes with mean field interaction have been used to model a class of chemical reactions. Many of the usual models considered in the literature such as the Schögl models have unbounded jump rates. Large deviations for such systems are investigated in the limit when the number of particles tends to infinity. The limiting behavior of the empirical process is described by the nonlinear master equation. Large deviations for the empirical distributions and the empirical processes are studied independently because of the difficulty in using the contraction principle which is applicable in the diffusion case.

The exponential tightness is necessary for getting the large deviation upper bounds when the space is Polish. What will happen if the space is not Polish? To answer this question we are led to study large deviation upper bounds for any metric space, including the non-separable metric space. Two abstract conditions for getting the large deviation upper bounds and their applications are given in chapter 6. In chapter 7 we prove that the exponential tightness is necessary for getting the large deviation principle when the space is the inductive limit of Polish spaces.

Acknowledgements

I would like to express my sincere gratitude to Prof. D.A. Dawson, my thesis supervisor, for his invaluable guidance and encouragement during the research of this thesis.

I would also like to thank Dr. J. Gärtner, Dr. M. Csörgö, Dr. D. McDonald and Dr. L. Resnick for their help and suggestions.

I'm very grateful to Carleton University for providing me with financial support during my Ph.D. studies and to the Department of Mathematics and Statistics for the strong academic atmosphere from which I benefited a lot.

At last, but not the least, I wish to thank my parents and my wife Min for the love, support and the faith they have always shown on me.

Contents

1	Introduction	1
1.1	Large Deviation Systems	1
1.2	Random Perturbations of Dynamical Systems	3
1.3	Large Deviations for Interacting Diffusions	5
1.4	Large Deviations for Processes With Jumps	9
1.5	The Large Deviations: Upper Bound	14
2	Preliminaries	18
2.1	Sanov's Theorem	18
2.2	Contraction Principle	20
2.3	Martingales and Ito's Formula	20
3	Particle System and Dynamical Model	24
3.1	Dynamical Model: Nonlinear Master Equation	24
3.2	Mean Field Interacting Particle System	26
3.3	Cameron-Martin-Girsanov Formula	31
3.4	Law of Large Numbers	34
4	Large Deviations For Empirical Distribution	36
4.1	Assumptions	36
4.2	The Main Result	38

4.3	Some Properties of the Space $\mathcal{M}_x$ and the Functional I_u	39
4.4	Large Deviation Estimates	45
5	Large Deviations For Empirical Process	55
5.1	The Main Result	55
5.2	Lower Bound	57
5.3	Upper Bound	74
6	Large Deviation: Upper Bound On Metric Space	82
6.1	Upper Bound	82
6.2	Application	89
6.2.1	Some Properties of Vector-valued L^p Space	89
6.2.2	Upper Bounds on Space of Measure-valued Cadlag Paths . . .	94
7	Upper Bound And Exponential Tightness	100
	Appendix	103
	Bibliography	105

Chapter 1

Introduction

1.1 Large Deviation Systems

The theory of large deviations is concerned with the study of the exponential decay of probabilities associated with certain random systems. An elementary example is provided by the sample means of independent, identically distributed real-valued random variables $\{X_i, i = 1, 2, \dots\}$ that have finite exponential moment. The latter holds if for all $\alpha \in \mathbf{R}$

$$E[\exp(\alpha X_i)] < \infty$$

where $E[\cdot]$ denotes the expectation. Let P_n denote the distribution of the sample mean $\frac{1}{n} \sum_{i=1}^n X_i$ and a be the first moment of X_1 . The law of large numbers states that P_n converges weakly to δ_a , the Dirac measure at a , as n goes to infinity. In 1938, H.Cramér showed that this convergence is in fact exponentially fast. This result, the first general result in the theory of large deviations, has been extensively generalized to cover many new situations. A more general framework of large deviations is given as follow.

Let X be an arbitrary topological space, $\mathcal{B}$ a σ -algebra on X . Consider a collection of probability measures $\{P_\varepsilon\}$ on $(X, \mathcal{B})$ with the positive index parameter $\varepsilon \rightarrow 0$. If there exists a function I defined on X with values in $[0, +\infty]$ such that

i). for every $\mathcal{B}$ -measurable, open subset G of X

$$\liminf_{\varepsilon \rightarrow 0} \varepsilon \log P_\varepsilon(G) \geq - \inf_{x \in G} I(x),$$

ii). for every $\mathcal{B}$ -measurable, closed subset F of X

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_\varepsilon(F) \leq - \inf_{x \in F} I(x),$$

iii). the level sets $\{x \in X : I(x) \leq s\}$ are compact for all $s \geq 0$,

then $(X, \mathcal{B}, P_\varepsilon, \varepsilon)$ is called a large deviation system with rate function (or action functional) I . Using the terminology of [13] we also say that P_ε satisfies a full large deviation principle with rate function I .

If in the above the closed subset F is replaced by a compact subset, the system will be called a weak large deviation system. For a weak large deviation system to be a large deviation system a well known sufficient condition is the following exponential tightness condition (cf. [13]): $\forall a < \infty$, there exists a $\mathcal{B}$ -measurable compact set $K_a \subset X$ such that

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_\varepsilon(K_a^c) \leq -a.$$

In the definition of large deviation system, two elements are very crucial: one is the space X , the other is the collection of probability measures $\{P_\varepsilon\}$. The space X can be a metric or nonmetric space, a separable or nonseparable space, a finite dimensional or infinite dimensional space, a real-valued or functional-valued space while the probability measures can be the empirical distributions of a sequence of i.i.d. random variables, the Markov processes, the measure-valued Markov processes, etc. Because of the richness of spaces and probability measures, the study of large deviations has already become a vast subject today.

The systematic study of large deviations began in the 1970's. These include the large deviations for Markov processes studied by Donsker and Varadhan(cf.[14], [15], [16], [17]); the small second order perturbations of first order equations and the associated exit problem studied by Freidlin and Wentzell [24] and the large deviations in the thermodynamic limits from the point of view of statistical mechanics studied by Lanford [30]. This list is in no way exhaustive.

Nowdays the study of large deviations has become extremely broad in both theory and applications. For the sake of explicitness we will only focus on the parts which are closely related to the thesis.

1.2 Random Perturbations of Dynamical Systems

Consider the following relatively simple situation. Let $b(x)$ be a uniformly Lipschitz continuous function on $\mathbf{R}^d$ (namely, $|b(x) - b(y)| \leq c|x - y|$ for some constant c). $\{x(t)\}$ be the unique solution of the ordinary differential equation

$$(2.1) \quad \frac{dx(t)}{dt} = b(x(t)), \quad x(0) = x.$$

For any $\varepsilon > 0$, let $\{x^\varepsilon(t)\}$ be the diffusion process which is the unique solution of the stochastic differential equation

$$dx^\varepsilon(t) = b(x^\varepsilon(t))dt + \sqrt{\varepsilon}dw(t) \quad t \geq 0, \quad x^\varepsilon(0) = x.$$

where $w(t)$ is a standard Brownian motion. Let P_ε denote the probability distribution of the process $\{x^\varepsilon(t)\}$. Then it is not hard to check that P_ε converges weakly to $\delta_{x(\cdot)}$ on any finite time interval as ε goes to zero. In other words the limit behavior of $\{x^\varepsilon(t)\}$ on any finite time interval is characterized by $\{x(t)\}$.

For any fixed $T > 0$, let $C([0, T])$ be the set of all real-valued continuous functions on $[0, T]$, H_x be the space of all absolute continuous functions on $[0, T]$ with value x at 0. Define

$$I_T(\psi) = \begin{cases} \frac{1}{2} \int_0^T |\dot{\psi}(t) - b(\psi(t))|^2 dt & \text{if } \psi \in H_x \\ +\infty & \text{otherwise} \end{cases}$$

Then $(C([0, T]), P_\varepsilon, \varepsilon)$ is a large deviation system with action functional $I_T(\cdot)$.

Let G be a bounded domain in $\mathbf{R}^d$ and ∂G be its boundary. If the trajectory $x(t)$ of system (2.1) issued from the point $x \in G$ leaves $G \cup \partial G$ within finite time, then for small ε the trajectories of the process $x^\varepsilon(t)$ started from x will leave G within the same time with a probability close to one and the the first exit places for the

two systems should be close with very high probability. Now consider the case in which $\bar{x}$ is the unique stable equilibrium point in G of the system

$$\frac{dx(t)}{dt} = b(x(t))$$

and the trajectory of $x(t)$ issued from any point $x \in G$ never leaves the region G . Thus the exit time and the exit place are not defined. In this situation it is still possible that the trajectories of the process $x^\epsilon(t)$ will exit from G because of the random perturbation. Indeed, suppose that the boundary of G is smooth enough such that

$$\tau_x^\epsilon = \inf\{t > 0 : x^\epsilon(t) \in \partial G\}$$

is a well defined stopping time. For any $y \in G$ we define the quasipotential of the system (2.1) with respect to the point $\bar{x}$ as

$$V(\bar{x}, y) = \inf_{t > 0} \inf_{\psi(\cdot) : \psi(0) = \bar{x}, \psi(t) = y} \{I_t(\psi)\}.$$

For any $x \in \partial G$, let $n(x)$ denote the exterior normal to the boundary of G . If we assume further that $\langle b(x), n(x) \rangle < 0$ for any $x \in \partial G$ and the quasipotential assumes its minimum $\bar{V}$ on ∂G at a unique point $y_0 \in \partial G$ then the following results can be found in [24]:

(a). For all $x \in G$ and all $\delta > 0$,

$$(2.2) \quad \lim_{\epsilon \rightarrow 0} P\left[e^{\frac{(V-\delta)}{\epsilon}} < \tau_x^\epsilon < e^{\frac{(V+\delta)}{\epsilon}}\right] = 1$$

Moreover, for all $x \in G$

$$(2.3) \quad \lim_{\epsilon \rightarrow 0} \log E[\tau_x^\epsilon] = \bar{V}$$

(b). For any $\delta > 0$ and any $x \in G$,

$$(2.4) \quad \lim_{\epsilon \rightarrow 0} P[|x^\epsilon(\tau_x^\epsilon) - y_0| < \delta] = 1$$

The proof of these results are based on the large deviation estimates.

These results can be, and were indeed, extended in various ways to cover the Gaussian perturbations of dynamical system, general Lévy process (cf. chapter 4 and 5 in [24]), non-Gaussian perturbations of partial differential equations (cf. [39]).

The exit problem is related to the long-time behavior of the diffusion process $x^\varepsilon(t)$. So is the problem of the limit behavior of the stationary distribution of $x^\varepsilon(t)$. Consider the dynamical system (2.1). A sufficient condition for the existence of finite invariant measure is the following. Let $r(x)$ be the position vector of the point x . We assume that there exists a large number N such that $\langle b(x), r(x) \rangle < -1/N$ for $|x| > N$. Let $\bar{x}$ be the unique stable equilibrium point of system (2.1) and the whole space $\mathbf{R}^d$ is attracted to $\bar{x}$. Then the process $x^\varepsilon(t)$ has a unique invariant measure μ^ε for every $\varepsilon > 0$ and

$$(2.5) \quad \lim_{\varepsilon \rightarrow 0} \varepsilon \log \mu^\varepsilon(C) = - \inf_{x \in C} V(\bar{x}, x),$$

for any domain $C \subset \mathbf{R}^d$ with compact boundary ∂C common for C and the closure of C .

The results in this section are the contents of the Freidlin and Wentzell theory. For a brief discussion (more detailed than we have given here) refer to Varadhan's book [43].

In the study of the exit problem, the assumption of nontangential drift $\langle b(x), n(x) \rangle < 0$ on the boundary can be generalized to the case of $\langle b(x), n(x) \rangle = 0$ on the boundary. Day [11] discussed this problem and obtained similar results for the exit distribution and the mean exit time.

There is also interest in the case when there are multiple stable equilibrium points of the dynamical system. These situations are more difficult to deal with and more interesting. This leads to the study of many infinite dimensional models.

1.3 Large Deviations for Interacting Diffusions

Beginning with the paper of Lanford [30] the theory of large deviations has played an important part in the study of equilibrium statistical mechanics. The objective of statistical mechanics is to explain the macroscopic behavior of matter on the basis of a large finite number of interacting particle systems which may be perturbed by noise. The usual method of study is to consider the behavior of the system as the number of particles become very large. In a series of papers, Dawson and Gärtner

made it very clear that an infinite dimensional generalization of the Freidlin and Wentzell theory is a powerful tool for the study of the dynamical aspects of phase transitions such as metastability for systems in both equilibrium and nonequilibrium statistical mechanics (cf. [5], [6], [7] and [8]).

For $k = 1, \dots, N$, consider the system of real-valued $x_k(\cdot)$ given by the stochastic differential equations

$$(3.1) \quad dx_k = -[U'_k(x_k) + \frac{\theta}{N} \sum_{l=1}^N (x_k - x_l)]dt + \sigma dw_k,$$

where $U_h(x) = \frac{x^4}{4} - \frac{x^2}{2} + hx$, $w_1, \dots, w_N$ are independent one dimensional Wiener processes and θ and σ denote positive constant. This is the continuous analog of the Curie-Weiss model of ferromagnetism. θ determines the strength of the ferromagnetic interaction, σ^2 denotes the *temperature* and h denotes an external field.

Let

$$X_N(t) = N^{-1} \sum_{k=1}^N \delta_{x_k(t)}, \quad t \geq 0,$$

denote the empirical process associated to (3.1). It is known that $X_N(\cdot)$ satisfies a law of large numbers as $N \rightarrow \infty$ and the limiting deterministic dynamic is characterized as a weak solution of the McKean-Vlasov equation

$$(3.2) \quad \dot{\mu}(t) = L(\mu(t))^* \mu(t), \quad t \geq 0,$$

where

$$L(\mu)f(x) = \frac{\sigma^2}{2} f''(x) - U'_h(x) f'(x) - \theta(x - m(\mu)) f'(x)$$

$L(\mu)^*$ denotes its adjoint and

$$m(\mu) = \int_R x \mu(dx).$$

This is a nonlinear equation. It was shown in [10] that in the case of $h = 0$, there exists a critical value σ_c such that for $\sigma \geq \sigma_c$, there is a unique fixed point μ_0 of the McKean-Vlasov dynamics and in the case of $\sigma < \sigma_c$ there exist three fixed points μ_0 and $\mu_{\pm}$, where $\mu_{\pm}$ are locally asymptotic stable and μ_0 is unstable. In the latter situation if N is large, then the empirical process $X_N(t)$ will normally follow the path of system (3.2) which is attracted by one of the equilibrium points μ_+ and μ_- . After

that X_N will perform small fluctuations around this stable equilibrium. However, because of ergodicity, from time to time a transition from one stable equilibrium to the other one will occur via large deviation. This type of dynamical phase transition due to a large deviation is called tunnelling and plays a central role in the study of metastability. To make this idea more precise we need the results of large deviations.

Let P_N^u and π_N be the distribution and the invariant distribution of the empirical process $X_N(t)$, respectively ($u^{\otimes N}$ is the initial distribution of the N -particle system). Define the free energy functional as

$$F(\nu) = \frac{\theta}{2} \langle \nu \otimes \nu, |x - y|^2 \rangle + \sigma^2 \langle \nu, \log \frac{d\nu}{d\mu_\sigma} \rangle,$$

where

$$\mu_\sigma(dx) = Z_\sigma^{-1} \exp[-\frac{2}{\sigma^2} U_h(x)] dx$$

The Gibbs free energy f is defined to be the minimum of the functional F . We have that

$$P_N^u(\mathcal{A}) = \exp\{-N \inf_{\mu(\cdot) \in \mathcal{A}} I_{0,T}(\mu(\cdot))(1 + o(1))\}$$

$$\pi_N(A) = \exp\{-N \inf_{\nu \in A} I(\nu)(1 + o(1))\}$$

as $N \rightarrow \infty$ for all appropriate subsets $\mathcal{A}$ of the space of measure-valued paths on a fixed time interval $[0, T]$ and subsets A of $M_1(\mathbf{R})$ (the space of all probability measures on space $\mathbf{R}$ equipped with the supremum norm) and where for any $s, t \in \mathbf{R}$, $s \leq t$

$$I_{s,t}(\mu(\cdot)) = \int_s^t \|\dot{\mu}(u) - L(\mu(u))^* \mu(u)\|_{\mu(u)}^2 du,$$

$$\|\nu\|_\mu^2 = \frac{1}{2\sigma^2} \sup_g \frac{|\langle \nu, g \rangle|}{\langle \mu, |g'|^2 \rangle},$$

$$I(\nu) = \sigma^{-2}(F(\nu) - f).$$

As in the finite dimensional case, we can also define the so-called quasipotential which is very important for describing transitions of the empirical process from one state to another:

$$V(\mu; \nu) = \inf\{I_{s,t}(\mu(\cdot)) : s, t \in R, \mu(s) = \mu, \mu(t) = \nu\},$$

where $\mu, \nu \in M_1(R)$. If ρ is a fixed point of the McKean-Vlasov system (3.2), then the quasipotential V_ρ with respect to ρ is defined by $V_\rho(\nu) = V(\rho; \nu)$. A very nice result obtained in [8] is that if ν lies in the domain of attraction of ρ , then

$$V_\rho(\nu) = \sigma^{-2}(F(\nu) - F(\rho))$$

In the case of $h = 0$, one of the fixed points, μ_- , is metastable. The finite particle system can escape from the domain of attraction of μ_- . It can also escape from the domain of attraction of μ_+ . However, by the metastability of μ_- , over a sufficiently large time scale the proportion of time that the empirical process $X_N(\cdot)$ spends near the "thermodynamically stable" state μ_+ approaches one as $N \rightarrow \infty$. The transition from the "thermodynamically stable" state to other states is possible, but it takes an exponentially long time for this to happen. A rigorous proof using the Freidlin and Wentzell estimates of these results is included in [5]. This gives us a much more deeper understanding about the theory of large deviations. For convenience we copy them down.

Let $G_\pm$ denote the domains of attraction of $\mu_\pm$ and let W_+ be a neighbourhood of μ_+ contained in the complement of G_- . Let

$$\tau_-(\mu(\cdot)) = \inf\{t \geq 0 : \mu(t) \notin G_-\}$$

$$\tau_+(\mu(\cdot)) = \inf\{t \geq 0 : \mu(t) \notin W_+\}$$

$$\Delta F_\pm = F(\mu_0) - F(\mu_\pm).$$

Let $\nu_N \rightarrow \nu \in G_-$ and $\delta > 0$. Then

$$\lim_{N \rightarrow \infty} P_N^{\nu_N} \left\{ \exp\left[\frac{N(\Delta F_- - \delta)}{\sigma^2}\right] \leq \tau_- \leq \tau_+ \leq \exp\left[\frac{N(\Delta F_- + \delta)}{\sigma^2}\right] \right\} = 1,$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log E_N^{\nu_N} \tau_- = \lim_{N \rightarrow \infty} \frac{1}{N} \log E_N^{\nu_N} \tau_+ = \frac{\Delta F_-}{\sigma^2}.$$

In general, if $\Delta F_+ > \Delta F_-$, then μ_+ is thermodynamically stable and μ_- is metastable.

The large deviations results described in this section are only a special case of the more general results in [6] and [8]. In fact the result in [8] hold for more general

potential $U(x)$ satisfying the conditions $U1 - U3$ in section 2 of [8]. While the results of large deviations for sample paths in [6] is valid for a much more general class of diffusions. The results and methods included in [6] are not only very useful for the study the statistical mechanics but also possess its own values. The powerful method of projective limit developed in the article has become a bridge that connects the finite dimensional case with the infinite dimensional case. It also provides a systematic treatment of large deviations for infinite dimensional diffusions. Many techniques such as a generalization of Sanov's theorem, the method of obtaining the non-variational expressions of the action functional and the estimations used to obtain the large deviations for interacting system from the independent system are very typical and can be used to solve other problems.

However in the diffusion case we can take advantage of the continuity of the sample paths. Using the contraction principle (cf. theorem 2.2.1), the large deviations for the empirical processes can be obtained from the large deviations of the empirical distributions. What will happen if there are jumps? What will the large deviations look like for pure jump processes? This is the main topic of the present thesis.

1.4 Large Deviations for Processes With Jumps

In the study of nonequilibrium thermodynamics and dissipative structures stochastic method plays an important role. Master equations (the forward Kolmogorov equations) for birth-death processes have been successfully used to describe the fluctuations in the systems.

Let $E = \{0, 1, \dots\}$, $E^{\otimes N}$ be the N -fold Cartesian product of E . The N -particle system $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$ on $E^{\otimes N}$ with mean field interaction is a Markov process generated by

$$(4.1) \quad \Omega^{(N)} \psi(x^{(N)}) = \sum_{k=1}^N Q_{x^{(N)}}^{(k)} \psi(x^{(N)})$$

where $\varepsilon_{x^{(N)}} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}}$, $x^{(N)} = (x_1^{(N)}, \dots, x_N^{(N)}) \in E^{\otimes N}$. Q_u is defined by

$$(4.2) \quad Q_u f(x) = \sum_{y \in E} q_{x,y} (f(y) - f(x)) + \|u\| (f(x+1) - f(x))$$

$Q = (q_{x,y})$ is a Q -matrix (see definition 2.3.2). $\|u\|$ is the first moment of u . $Q_u^{(k)}$ is used instead of Q_u when it acts on the k -th variable of $\psi \in C(E^{\otimes N})$ (the set of continuous functions on $E^{\otimes N}$).

Let $M_1(E)$ be the set of all probability measures on E . Given a solution $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$ associated to (4.1), let

$$\varepsilon_{x^{(N)}(t)} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}(t)}, \quad t \geq 0.$$

denote the empirical process, where δ_x denotes the Dirac measure at x . It is known that for any finite time interval the empirical process satisfies a law of large numbers as $N \rightarrow \infty$ (D.A.Dawson and X.Zheng [9]) and that the limiting deterministic dynamics is characterized as a weak solution of the nonlinear master equation

$$(4.3) \quad \frac{d\langle u(t), f \rangle}{dt} = \langle u(t), Q_{u(t)} f \rangle, \quad f \in C_k(E),$$

where $C_k(E)$ denotes the set of functions on E with compact support.

The nonlinear master equation was proposed by G.Nicolis and I.Prigogine [34] as a mean-field model of a chemical reaction with spatial diffusion. S.Feng and X.Zheng [22] established the existence and uniqueness of the solution to the martingale problem associated with the nonlinear master equation and proved the existence of at least three equilibrium states for the second Schlögl model (see [37]). Following the pattern of section 1.3, it is very natural to study the multiple equilibria transitions and the metastability for the nonlinear master equations. As a first step we discuss the problem of large deviations of the empirical process $\varepsilon_{x^{(N)}(t)}$ of the N -particle system $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$. But now we are facing a totally different situation from the diffusion case. The N -particle system $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$ is a Markov process with pure jumps. Some techniques used in the diffusion case are no longer applicable in the present situation.

The study of the large deviations for processes with jumps is not so extensive as in the diffusion case. F.Comets [4] proved a large deviation principle for the Curie-Weiss model on the torus. This may be the most systematic study of this kind of

problem in the existing literature. The newer developments include the study of large deviations for more general pure jump processes with compact state space by M.Sugiura [42] and the study of large deviations for particle systems associated with spatially homogeneous Boltzmann type equations by C.Léonard [32].

In this thesis, we will first discuss the large deviations of the empirical distribution $\eta_{x^{(N)}(\cdot)}$ of the N -particle system

$$\eta_{x^{(N)}(\cdot)} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}(\cdot)}$$

This is a random measure on the Skorohod path space $D([0, T], E)$, where $\delta_x(\cdot)$ denotes the Dirac distribution on path $x(\cdot)$. Assume that $(x_1^{(N)}(0), \dots, x_N^{(N)}(0))$ are independent and identically distributed with common distribution $u \in M_1(E)$. We can prove that $\eta_{x^{(N)}(\cdot)}$ converges weakly to a deterministic measure on $D([0, T], E)$ whose marginal distribution is a solution to the nonlinear master equation (4.3).

For fixed $T > 0$, let $D = D([0, T], E)$. $\mathcal{M} = M_1(D)$ denotes the space of all probability measures on D with weak topology. For each $R \geq 1$, define $\mathcal{M}_R = \{P \in \mathcal{M} : \sup_{0 \leq t \leq T} \int \varphi(\omega(t)) P(d\omega) \leq R\}$ equipped with the subspace topology of $\mathcal{M}$, where φ is defined as $\varphi : E \rightarrow \mathbb{R}, x \rightarrow 1 + x \log \log(x+2)$. Let $\mathcal{M}_\infty = \bigcup_{R \geq 1} \mathcal{M}_R$ be equipped with the "inductive topology".

Let $Q = (q_{x,y})_{x,y \in E}$ be a totally stable, conservative Q -matrix satisfying:

$$(4.4) \quad \inf_{x \in E} \{q_{x,x+1}\} > 0;$$

$$(4.5) \quad \exists \Lambda > 0, \text{ such that } q_{x,y} = 0 \text{ for } |x - y| \geq \Lambda;$$

$$(4.6) \quad \exists 0 < c < \infty, \text{ such that } \forall x, y \in E, y > x,$$

$$\begin{aligned} & \sum_{z \neq 0} (q_{y,y+z} - q_{x,x+z})z \\ & + 2 \sum_{z=1}^{\infty} [(q_{y,x-z} - q_{x,2x-y-z}) \vee 0 + (q_{x,y+z} + q_{y,2y-x+z}) \vee 0]z \leq c(y-x); \end{aligned}$$

$$(4.7) \quad \forall l \geq 0, \exists \lambda(l) > 0, \exists$$

$$\sup_{x \in E} \max \left\{ \sum_{y \in E} q_{x,y} (e^{y-x} - 1) + (e - 1)x, \right.$$

$$\sum_{y \in E} q_{x,y} (e^{\varphi(y) - \varphi(x)} - 1) + (e^{\varphi(x+1) - \varphi(x)} - 1)l \leq \lambda(l).$$

Remark. (4.4) is a technical condition which makes it much easier for using the Cameron-Martin-Girsanov formula. Condition (4.7) implies the following:

$$(4.8) \quad \exists \lambda > 0 \text{ such that } \sum_{y \in E} q_{x,y} (y - x) \leq \lambda(x + 1),$$

$$\sum_{y \in E} q_{x,y} (\varphi(y) - \varphi(x)) \leq \lambda \varphi(x), \quad x \in E;$$

This condition and (4.6) guarantee that the associated martingale problems are well-posed. Conditions (4.5) and (4.7) are especially needed for our large deviation results. (4.6) is also called the Lipschitz condition.

For any $u \in M_1(E)$ and $u(\cdot) \in D([0, T], M_1(E))$ satisfying

$$\langle \varphi, u \rangle = \int_E \varphi u(dx) < \infty, \quad \sup_{0 \leq t \leq T} \langle \varphi, u(t) \rangle < \infty$$

let $P_{u(\cdot), u}$ be the unique solution of the time-inhomogeneous martingale problem for $Q_{u(t)}$ with initial distribution u , $P_u^{(N)}$ be unique solution to the martingale problem for $\Omega^{(N)}$ on $D([0, T], E)^{\otimes N}$ with initial distribution $u^{\otimes N}$. For any $P \in \mathcal{M}$ let $\pi(P)(\cdot) = P \circ x_{(\cdot)}^{-1}$. Let $I_u^{u(\cdot)} : \mathcal{M}_\infty \rightarrow [0, \infty]$ be defined as

$$(4.9) \quad I_u^{u(\cdot)}(P) = \sup_{\Phi \in C_b(D)} \left\{ \int \Phi dP - \log E^{P_{u(\cdot), u}}(e^\Phi) \right\}$$

where $P \in \mathcal{M}_\infty$ and $C_b(D)$ is the set of all bounded continuous functions on D . Let $I_u = I_u^{x^{(N)}(\cdot)}(P)$, $\mathcal{P}_u^{(N)}$ be the image of $P_u^{(N)}$ under the map $\eta_{x^{(N)}(\cdot)}$. The main result of chapter 4 is the following: Let $u \in M_1(E)$ satisfy $\int_E e^{\varphi(x)} u(dx) < \infty$. Under conditions (4.4)-(4.7), $(\mathcal{M}_\infty, \mathcal{P}_u^{(N)}, N)$ is a large deviation system with action functional I_u .

The main idea of the proof is to deal with the large deviations for independent processes first, and then to derive the result through the Cameron-Martin-Girsanov formula which is obtained by T.Shiga and H.Tanaka [38] and by D.A.Dawson and X.Zheng [9]. But in our case the unboundedness of jumps makes the whole picture much more complicated.

Consider the map

$$\pi(\eta_{\varepsilon^{(N)}(\cdot)}) = (t \mapsto \varepsilon_{\varepsilon^{(N)}(t)}).$$

If π were continuous, as in the diffusion case, then the large deviation principle for the empirical process would follow directly from the large deviation principle of the empirical process by using the contraction principle. Unfortunately, this is not true. A counterexample is given in the Appendix. Thus a different method is needed when dealing with the large deviations of the empirical process. The method we use to obtain the lower bound is similar to the method used in [42]. The idea is to first obtain an lower bound by using Sanov's theorem and then to verify that it is the lower bound we want. The difficult part in proving the upper bound is usually to check the exponential tightness condition. Comets in [4] developed a method showing that some neighborhood of any level set has a exponential decay. This is just the upper bound in Freidlin-Wentzell theory. In [4] the exponential tightness condition can be obtained from the compactness of the level sets and the the exponential decay of some neighborhood of any level set because the space where the large deviations are discussed is a Polish space. While in our model the space is a non-metrizable space. Thus some extra estimates are needed for getting the upper bound.

For each $m \geq 1$, define $\mathcal{D}_m = \{\mu(\cdot) \in \mathcal{D}; \sup_{0 \leq t \leq T} \langle \nu(t), \varphi \rangle \leq m\}$ equipped with the Skorohod topology. $\mathcal{D}_\infty = \bigcup_{m \geq 1} \mathcal{D}_m$ is equipped with the "inductive topology". By definition, a set V is open in $\mathcal{D}_\infty$ if and only if $V \cap \mathcal{D}_m$ is open in $\mathcal{D}_m$ for each $m \geq 1$. Let $\hat{\mathcal{P}}_\nu^{(N)}(\cdot) = P_\nu^{(N)} \circ \varepsilon_{\varepsilon^{(N)}(\cdot)}^{-1}$. Then we have $\hat{\mathcal{P}}_\nu^{(N)} \in M_1(D([0, T], M_1))$, $\mathcal{P}_\nu^{(N)} \rightarrow \delta_{\pi(P_\nu)(\cdot)}$ and that $\hat{\mathcal{P}}_\nu^{(N)}$ is concentrated on space $\mathcal{D}_\infty$.

Under assumptions (4.4)-(4.7) we have that for any $r > 0$, there exists an $R_0 > 0$ such that for all $R \geq R_0$ and $N \geq 1$

$$(4.10) \quad \hat{\mathcal{P}}_\nu^{(N)}\{\mathcal{D}_\infty \setminus \mathcal{D}_R\} \leq \exp(-Nr).$$

Combining this estimate and the method used in [4] we get the upper bound.

For fixed $T > 0$, let $C_b(E)$ denote the set of all bounded continuous functions on E . $C_0^{1,0}([0, T] \times E)$ denotes the set of all continuous functions on $[0, T] \times E$ with compact support and first order continuous derivative with respect to t .

$\forall u, \nu \in M_1(E)$, define

$$l_u(\nu, g) = \langle \nu, g \rangle - \log \langle u, e^g \rangle, g \in C_b(E).$$

$\forall \bar{\mu}(\cdot), \mu(\cdot) \in \mathcal{D}_\infty, f \in C_0^{1,0}([0, T] \times E)$, define

$$\begin{aligned} J_{\bar{\mu}(\cdot)}(\mu(\cdot); f, T) &= \langle \mu(T), f(T) \rangle - \langle \mu(0), f(0) \rangle \\ &\quad - \int_0^T \langle \mu(s), \frac{\partial f(s)}{\partial s} + e^{-f(s)} Q_{\bar{\mu}(s)} e^{f(s)} \rangle ds, \\ J(\mu(\cdot); f, T) &= J_{\mu(\cdot)}(\mu(\cdot); f, T). \end{aligned}$$

$$(4.11) \quad \begin{aligned} S_u^{\bar{\mu}(\cdot)}(\mu(\cdot)) &= \sup \{ l_u(\mu(0), g) + J_{\bar{\mu}(\cdot)}(\mu(\cdot); f, T); \\ &\quad f \in C_0^{1,0}([0, T] \times E), g \in C_b(E) \}, \end{aligned}$$

$$(4.12) \quad S_u(\mu(\cdot)) = S_u^{\mu(\cdot)}(\mu(\cdot)).$$

Then we have the following large deviations result for the empirical process:

Let

$$u \in M_1(E), \quad \int_E e^{\varphi(x)} u(dx) < \infty.$$

Under conditions (4.4)-(4.7), $(\mathcal{D}_\infty, \hat{\mathcal{P}}_u^{(N)}, N)$ is a large deviation system with action functional S_u .

From above discussion we can see that the exponential tightness is important for proving the upper bound but very difficult to check in some cases. It is very natural to ask whether the exponential tightness is necessary for getting a full large deviation principle or not. This is the topic of next section.

1.5 The Large Deviations: Upper Bound

Recent developments of the theory of large deviations show that the topological structure, the linear structure or the convex structure of the space X plus some mild conditions on the family of the probability measures are "sufficient" for obtaining

the upper bound for compact subset. Consider a Hausdorff topological vector space X . The dual space of X , namely the space of all continuous linear functionals on X , is denoted by X^* . Let $\{P_\varepsilon\}$ be a family of probability measures on X . Define the logarithmic moment generating function $\Lambda_{P_\varepsilon}(\lambda) = \log \int_X e^{\lambda(x)} P_\varepsilon(dx)$, $\lambda \in X^*$. If the limit

$$\Lambda(\lambda) = \lim_{\varepsilon \rightarrow 0} \varepsilon \Lambda_{P_\varepsilon}(\lambda/\varepsilon)$$

exists as an extended real number, then the upper bound holds for any compact subset of X with rate function to be the Fenchel-Legendre transform of Λ . To get upper bound for any closed subset we usually add the condition of exponential tightness. When dealing with finite dimensional spaces the exponential tightness is equivalent to the following condition

$$(5.1) \quad \forall \eta > 0, \exists \text{ a bounded closed subset } B_\eta \text{ of } X \ni \limsup_{\varepsilon \rightarrow \infty} \varepsilon \log P_\varepsilon(B_\eta^c) \leq -\eta.$$

In infinite dimensional space compact sets and closed bounded sets are not equivalent in general. Thus we have two candidates for the condition to get the large deviation upper bound in infinite dimensional space: one is the "exponential tightness" and the other is (5.1). Recently, P. Baldi [1] studied the large deviation principle for a sequence of probability measures on an arbitrary normed vector spaces. Under the assumption of "exponential tightness", he obtained a large deviation principle using the same method as that in finite dimensional case. C. Léonard [31] suggested that when dealing with the case where X is the strong dual space of some separable normed vector space one could use the condition (5.1) instead of the "exponential tightness". This is not only a technically simpler condition but also a weaker condition in the case of infinite dimension.

It is well-known that when X is locally compact or Polish, the condition of exponential tightness is necessary for the existence of a full large deviation principle (cf. exercises 1.2.19 and 4.1.10 in [12] or theorem (P) of [35]). In chapter 7 we will show that this result is also true when the space is the inductive limit of a sequence of Polish spaces.

In chapter 6 we will discuss the large deviation upper bound on any metric space. Two abstract conditions are given and their relations with the exponential tightness and (5.1) are also discussed. Roughly speaking our result shows that assumption of the exponential tightness can be weakened to get a good upper bound while (5.1) may not sufficient to get even a weaker upper bound. (see theorem 6.1.4 and the remark after the proof of theorem 6.1.7)

Let d_1, d_2 be two metrics on X . The topology of the space (X, d_1) is assumed to be weaker than the topology of space (X, d_2) and any bounded closed ball in (X, d_2) is assumed to be compact in (X, d_1) . Let $\mathcal{B}_1$ be the Borel σ -algebra of the space (X, d_1) , $\{P_N\}_{N=1, \dots}$ be a sequence of probability measures on $(X, \mathcal{B})$, where $\mathcal{B}$ is a σ -algebra on X contained in the Borel σ -algebra of space (X, d_2) and containing $\mathcal{B}_1$. We assume that the balls in space (X, d_2) are $\mathcal{B}_1$ -measurable. If (X, d_2) is separable then the two spaces have the same Borel σ -algebra. The main result of this article is to obtain the upper bound for $\{P_N\}_{N=1, \dots}$ for all $\mathcal{B}$ -measurable, closed subsets of (X, d_2) under some general assumptions. The reason why we do not choose $\mathcal{B}$ to be the Borel σ -algebra of space (X, d_2) is the following. If (X, d_2) is non-separable the class of probability measures defined on its Borel σ -algebra is usually very "small". In fact from theorem 5.5 of [18] we can see that one cannot expect a probability measure spread over a non-separable metric space to be defined on all Borel sets. In other words in the case of non-separable space the result here is the "best" that one can expect.

The main idea of proof is to embed X in $M_1(X)$ (the set of all probability measures on X) by sending x to δ_x (the Dirac measure at x). This idea was proposed by C. Léonard in [31]. The upper bound on X is then obtained through the upper bound on $M_1(X)$. In the rest part of chapter 6, we first discuss some basic properties of the vector-valued L^p space which can be found in [20] and [19]. Then a detailed discussion is given about several kinds of measurability for the space $\mathcal{D}_\infty$ of measure-valued cadlag paths satisfying some moment condition. Finally, since the space $\mathcal{D}_\infty$ is contained in the the strong dual space $(L^1)^*$ of the vector-valued L^1 space, applying the abstract result to $(L^1)^*$, we obtain the upper bound for a sequence of probability measure on $\mathcal{D}_\infty$.

The outline of the development of this thesis is as follows. Some preliminary results about large deviations are given in chapter 2. In chapter 3 we give a detailed discussion about the mean field particle system and nonlinear martingale problem associated with nonlinear master equations. The law of large numbers and the Cameron-Martin-Girsanov formula are also given in chapter 3. The large deviations for the empirical distributions are the topic of chapter 4. The results in this chapter has been announced in [23]. Chapter 5 deals with the large deviations for the empirical processes. Finally in chapter 6 we get an abstract result for the large deviation upper bound on some general metric space. In chapter 7 we show that the exponential tightness is necessary for getting the large deviation principle when the space is the inductive limit of Polish spaces. This is a generalization of a result in [35].

Chapter 2

Preliminaries

In this chapter some preliminaries are given in order to make the thesis self-contained. We will not give the most general result but rather focus on the results which are closely related to the thesis.

2.1 Sanov's Theorem

Let X be a Polish space. $M_1(X)$ denote the set of all probability measures on X . $M_1(X)$ be equipped with the weak topology generated by the sets $\{U_{f,r,\delta}; f \in C_b(X), r \in R, \delta > 0\}$, where

$$U_{f,r,\delta} = \{\nu \in M_1(X) : |\langle \nu, f \rangle - r| < \delta\}, \quad \langle \nu, f \rangle = \int_X f(x) \nu(dx).$$

Let $x_1, \dots, x_n$ be a sequence of independent, X -valued random variables, identically distributed according to a $\mu \in M_1(X)$. The empirical law of $x_1, \dots, x_n$ is

$$L_n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i} \in M_1(X).$$

Let $P_\mu^{(n)}$ be the image of $\mu^{\otimes n}$ under L_n . Define the relative entropy of the probability measure ν with respect to μ as

$$H(\nu|\mu) = \begin{cases} \int_X \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} d\mu & \text{if } \nu \ll \mu \\ \infty & \text{otherwise} \end{cases}$$

where $\nu \ll \mu$ means that ν is absolute continuous with respect to μ and $\frac{d\nu}{d\mu}$ stands for the Radon-Nikodym derivative of ν with respect to μ . Another expression for the relative entropy is

$$H(\nu|\mu) = \sup_{f \in C_b(X)} [\langle \nu, f \rangle - \log \langle \mu, e^f \rangle]$$

Then we have the following Sanov's theorem

Theorem 2.1.1 $(M_1(X), P_\mu^{(n)}, n)$ is a large deviation system with rate function $H(\cdot|\mu)$.

For the proof refer to lemma 3.37 in [41].

The following is a mild generalization of Sanov's theorem which is theorem 3.5 of [6].

Let $X, Y_1, \dots, Y_r$ be Polish spaces. Given any $N \geq 1$, set

$$M^{(N)}(X) = \{N^{-1} \sum_{k=1}^N \delta_{x_k} : x_1, \dots, x_k \in X\}.$$

Let $\{P_x; x \in X\}$ be a Feller continuous family of probability measures on $Y = Y_1 \times \dots \times Y_r$. Given

$$\nu = N^{-1} \sum_{k=1}^N \delta_{x_k} \in M^{(N)}(X)$$

denote by $Q_\nu^{(N)}$ the probability law of the vector of empirical measures

$$(N^{-1} \sum_{k=1}^N \delta_{y_k^{(1)}}, \dots, N^{-1} \sum_{k=1}^N \delta_{y_k^{(r)}})$$

under $P_{x_1} \otimes \dots \otimes P_{x_N}$, where $y_k = (y_k^{(1)}, \dots, y_k^{(r)})$. We have

Theorem 2.1.2 Given $\nu_N \in M^{(N)}(X)$ and $\nu \in M_1(X)$, suppose that $\nu_N \rightarrow \nu$ in $M_1(X)$. Then $(M_1(Y_1) \times \dots \times M_1(Y_r), Q_{\nu_N}^{(N)}, N)$ is a large deviation system with action functional

$$\begin{aligned} L_\nu(\mu_1, \dots, \mu_r) = & \sup_{(f_1, \dots, f_r) \in C_b(Y_1) \times \dots \times C_b(Y_r)} \left[\sum_{i=1}^r \langle \mu_i, f_i \rangle \right. \\ & \left. - \int_Y (\log \langle P_x, \exp(\sum_{j=1}^r f_j(y^{(j)})) \rangle) \nu(dx) \right], \quad (\mu_1, \dots, \mu_r) \in M_1(Y_1) \times \dots \times M_1(Y_r). \end{aligned}$$

Obviously theorem 2.1.1 is a special case of theorem 2.1.2 when $r = 1$, $Y_1 = X$ and $P_x \equiv \mu$.

2.2 Contraction Principle

The contraction principle shows that the continuous transformations preserve the large deviation principle with a possible change in rate functions. For more precise we have

Theorem 2.2.1 *Let X and Y be any two Hausdorff topological spaces and $f : X \rightarrow Y$ a continuous function. Assume that the family of probability measures $\{P_\epsilon\}$ on X satisfies a full large deviation principle with rate function I , then the family of probability measures $\{P_\epsilon \circ f^{-1}\}$ on Y also satisfies a full large deviation principle with rate function I' defined by*

$$I'(y) = \inf\{I(x); x \in X, y = f(x)\}$$

Note: The rate functions are not unique in general. But if the space where the large deviations are considered is a regular Hausdorff topological space, then the rate function is unique (cf. lemma 4.1.4 of [12]).

The contraction principle can be extended beyond the continuous case. Roughly speaking, if the map f can be "exponentially" approximated by a sequence of continuous maps, then the contraction principle holds for f . But the "exponential" approximation is not easy to check and in some cases it is easier to prove a large deviation result directly instead of trying to use the contraction principle by checking the "exponential" approximation condition. For more detail refer to [13] or [12].

2.3 Martingales and Ito's Formula

Let $\mathbf{T}$ be the time index set: $\mathbf{T} = \{0, 1, \dots\}$ in the discrete time case or $\mathbf{T} = [0, \infty)$ in the continuous time case. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_t)_{t \in \mathbf{T}}$ be an increasing family of sub- σ -algebra of $\mathcal{F}$.

Definition 2.3.1 *A real-valued stochastic process $X = (X(t))_{t \in \mathbf{T}}$ is called a martingale (supermartingale, submartingale) with respect to $(\mathcal{F}_t)_{t \in \mathbf{T}}$ if*

- (i). $X(t)$ is integrable for each $t \in \mathbf{T}$.

(ii). X is $(\mathcal{F}_t)$ -adapted.

(iii). $E(X(t)|\mathcal{F}_s) = X(s)$, (resp. $E(X(t)|\mathcal{F}_s) \leq X(s)$, $E(X(t)|\mathcal{F}_s) \geq X(s)$) a.s. for every $t, s \in \mathbf{T}$ such that $s < t$.

Note: In the continuous case $(\mathcal{F}_t)_{t \in \mathbf{T}}$ is assumed to be right-continuous.

Definition 2.3.2 A real-valued stochastic process $X = (X(t))_{t \in \mathbf{T}}$ is called a *local martingale* (local supermartingale, local submartingale) with respect to $(\mathcal{F}_t)_{t \in \mathbf{T}}$ if it is $(\mathcal{F}_t)$ -adapted and if there exists a sequence of $(\mathcal{F}_t)$ -stopping times $\{\tau_n\}$ such that $\tau_n < \infty$, $\tau_n \uparrow \infty$ and $X_n = (X(t \wedge \tau_n))$ is an $(\mathcal{F}_t)$ -martingale (resp. supermartingale, submartingale) for each $n \geq 1$.

The following theorem is a very useful result.

Theorem 2.3.1 Let $X = (X(t))$ be a non-negative local supermartingale. If $X(0)$ is integrable, then $X = (X(t))$ is a supermartingale.

Proof: Choose a sequence of $(\mathcal{F}_t)$ -stopping time τ_n such that $\tau_n < \infty$, $\tau_n \uparrow \infty$ and $X_n = (X(t \wedge \tau_n))$ is an $(\mathcal{F}_t)$ -supermartingale. Then for any $t > 0$ and $n \geq 1$, we have

$$0 \leq E(X(t \wedge \tau_n)|\mathcal{F}_0) \leq X(0) \text{ a.s.}$$

Let $n \rightarrow \infty$, by Fatou's lemma we conclude that $X(t)$ is integrable for any $t \in \mathbf{T}$. Finally for any $s < t$, let $n \rightarrow \infty$ in the following equation

$$E(X(t \wedge \tau_n)|\mathcal{F}_s) \leq X(s \wedge \tau_n) \text{ a.s.}$$

we get the result. $\square$

Let $(X, \mathcal{B}_X)$ be a measurable space. A *point function* p on X is a mapping from a countable subset D_p of $(0, \infty)$ to X . p defines a counting measure $N_p(dt dx)$ on $(0, \infty) \times X$ by

$$N_p(t, U) = \#\{s \in D_p: s \leq t, p(s) \in U\}, \quad t > 0, \quad U \in \mathcal{B}_X.$$

Let Π_X be the set of all point functions on X and $\mathcal{B}(\Pi_X)$ be the smallest σ -algebra on Π_X with respect to which all $p \rightarrow N_p(t, U)$ are measurable. Then a *point*

process is defined to be a measurable mapping from $(\Omega, \mathcal{F}, P)$ to $(\Pi_X, \mathcal{B}(\Pi_X))$. A point process p is called $(\mathcal{F}_t)$ -adapted if for every $t > 0$ and $U \in \mathcal{B}_X$, $N_p(t, U)$ is $\mathcal{F}_t$ -measurable. p is called σ -finite if there exist $U_n \in \mathcal{B}_X, n = 1, \dots$ such that $\cup_n U_n = X$ and $E[N_p(t, U)] < \infty$ for all $t > 0$ and $n \geq 1$. For $U \in \mathcal{B}_X$ satisfying $E[N_p(t, U)] < \infty$, $N_p(t, U)$ is an adapted, integrable increasing process and hence there exists a natural integrable increasing process $\hat{N}_p(t, U)$ such that $\tilde{N}_p(t, U) = N_p(t, U) - \hat{N}_p(t, U)$ is a martingale.

Let $\Gamma_p = \{U \in \mathcal{B}_X; E[N_p(t, U)] < \infty, t > 0\}$. We have

Definition 2.3.3 An $(\mathcal{F}_t)$ -adapted point process p on $(\Omega, \mathcal{F}, P)$ is said to be of class (QL) if it is σ -finite and there exists a process $(\hat{N}_p(t, U))$ such that

- (i). for $U \in \Gamma_p$, $t \mapsto \hat{N}_p(t, U)$ is a continuous $(\mathcal{F}_t)$ -adapted increasing process,
- (ii). for each t and almost all $\omega \in \Omega$, $U \mapsto \hat{N}_p(t, U)$ is a σ -finite measure on $(X, \mathcal{B}_X)$.
- (iii). for $U \in \Gamma_p$, $\tilde{N}_p(t, U)$ is an $(\mathcal{F}_t)$ -martingale.

Definition 2.3.4 A real function $f(t, x, \omega)$ defined on $[0, \infty) \times X \times \Omega$ is called $(\mathcal{F}_t)$ -predictable if the mapping $(t, x, \omega) \mapsto f(t, x, \omega)$ is $\mathcal{L}/\mathcal{B}(\mathbb{R})$ measurable. Where $\mathcal{L}$ is the smallest σ -algebra on $[0, \infty) \times X \times \Omega$ with respect to which all g having the following properties are measurable:

- (i). for each $t > 0$, $(x, \omega) \mapsto g(t, x, \omega)$ is $\mathcal{B}_X \times \mathcal{F}_t$ measurable.
- (ii). for each (x, ω) , $t \mapsto g(t, x, \omega)$ is left continuous.

We introduce the following two classes of functions:

$$\mathbf{F}_p = \{f(t, x, \omega); f \text{ is } (\mathcal{F}_t) \text{ predictable and for each } t > 0,$$

$$\int_0^{t+} \int_X |f(t, x, \omega)| N_p(ds dx) < \infty \text{ a.s.}\}$$

$$\mathbf{F}_p^{2,loc} = \{f(t, x, \omega); f \text{ is } (\mathcal{F}_t) \text{ predictable and there exists a}$$

sequence of $(\mathcal{F}_t)$ - stopping times $\{\tau_n\}$ such that

$$\tau_n \uparrow \infty \text{ a.s. and for every } t > 0, n \geq 1,$$

$$E \int_0^{t+} \int_X |I_{[0, \tau_n]}(t) f(t, x, \omega)|^2 \hat{N}_p(ds dx) < \infty \}$$

For $f \in \mathbf{F}_p$ and $g \in \mathbf{F}_p^{2,loc}$,

$$\int_0^{t+} \int_X f(t, x, \omega) N_p(dsdx) \text{ and } \int_0^{t+} \int_X g(t, x, \omega) \tilde{N}_p(dsdx)$$

are well-defined.

Ito's formula is one of the most important tools in stochastic analysis. There are many forms of it. In this section we will write down a particular form which is closely related to our applications.

Let p be a point process of class (QL) and $f^i(t, x, \omega) \in \mathbf{F}_r$, $g^i(t, x, \omega) \in \mathbf{F}_p^{2,loc}$ ($i = 1, 2, \dots, d$) such that $f^i g^j = 0$, $i, j = 1, \dots, d$. Consider the d -dimensional stochastic process $X(t) = (X^1(t), \dots, X^d(t))$ defined by

$$\begin{aligned} X^i(t) = & X^i(0) + A^i(t) + \int_0^{t+} \int_X f^i(t, x, \omega) N_p(dsdx) \\ & + \int_0^{t+} \int_X g^i(t, x, \omega) \tilde{N}_p(dsdx), \quad i = 1, \dots, d. \end{aligned}$$

where $A^i(t)$ ($i = 1, \dots, d$) are continuous $(\mathcal{F}_t)$ -adapted process whose almost all sample functions are of bounded variation on each finite time interval and for each $1 \leq i \leq d$, $A^i(0) = 0$. Let $f = (f^1, \dots, f^d)$, $g = (g^1, \dots, g^d)$. Then we have

Theorem 2.3.2 (Ito's formula) *Let F be a continuous function with continuous first order derivative. Then the stochastic process $F(X(t))$ satisfies*

$$\begin{aligned} F(X(t)) - F(X(0)) = & \sum_{i=1}^d \int_0^t F'_i(X(s)) dA^i(s) \\ & + \int_0^{t+} \int_X \{F(X(s-) + f(s, x, \cdot)) - F(X(s-))\} N_p(dsdx) \\ & + \int_0^{t+} \int_X \{F(X(s-) + g(s, x, \cdot)) - F(X(s-))\} \tilde{N}_p(dsdx) \\ & + \int_0^{t+} \int_X \{F(X(s) + f(s, x, \cdot)) - F(X(s)) \\ & - \sum_{i=1}^d g^i(s, x, \cdot) F'_i(X(s))\} \tilde{N}_p(dsdx). \end{aligned}$$

All definitions and Ito's formula in this chapter can be found in chapter I and chapter II of [29].

Chapter 3

Particle System and Dynamical Model

3.1 Dynamical Model: Nonlinear Master Equation

Let $E = \{0, 1, \dots\}$ be equipped with the discrete topology and ρ be the associated metric. For fixed $T > 0$, let $D = D([0, T], E)$ be the space of all right continuous functions $\omega : [0, T] \rightarrow E$ which have left limits at each $t \in (0, T]$ and are left continuous at T with the Skorohod metric d on it. It is well known that (D, d) is a Polish space in which the Borel σ -algebra coincides with $\mathcal{F} = \sigma\{x(t) : 0 \leq t \leq T\}$, the smallest σ -algebra generated by $\{x(t) : 0 \leq t \leq T\}$ where $x(t, \omega) = \omega(t)$ for all $t \geq 0$ and $\omega \in D$ (cf. [21]). For each $t \geq 0$, let $\mathcal{F}_t = \sigma\{x(s) : 0 \leq s \leq t\}$. $M_1(D)$ denotes the space of all probability measures on D with weak topology. Then $M_1(D)$ is also a Polish space. Let

$$\varphi : E \rightarrow \mathbb{R}, x \rightarrow 1 + x \log \log(x + 2).$$

Let $Q = (q_{x,y})_{x,y \in E}$ be a totally stable conservative Q -matrix (see definition 3.2.1) satisfying:

$$(1.1) \quad \forall y \in E, \sup_{x \in E, x \neq y} q_{x,y} < \infty$$

$$(1.2) \quad \exists p > 1, M > 0, \lambda > 1, \text{ such that}$$

$$\sum_{y \in E} q_{x,y} (y^q - x^q) \leq M + \lambda x^q, \quad x \in E, \quad q = 1 \text{ or } p.$$

$$(1.3) \quad \exists 0 < c < \infty, \text{ such that } \forall x, y \in E, y > x,$$

$$\sum_{z \neq 0} (q_{y, y+z} - q_{x, x+z})z \\ + 2 \sum_{z=1}^{\infty} [(q_{y, x-z} - q_{x, 2x-y-z}) \vee 0 + (q_{x, y+z} - q_{y, 2y-x+z}) \vee 0] z \leq c(y - x);$$

Remark. These conditions are satisfied by many models such as the Schlögl models, the linear growth model and the autocatalytic reaction model (cf. [34]). The corresponding Kolmogorov forward equations for these models are referred to as linear master equations which describe the chemical reaction in a container. Nicolis and Prigogine also considered the systems of chemical reactions with mean-field interactions between the inside and outside of the container to model the effects of spatial diffusion. These systems can be described by nonlinear master equations. Condition (1.2) is stronger than the condition (4.8) of chapter 1. But it can be shown that the latter is enough for the existence and uniqueness of the associated martingale problem. From the point of view of applications there are no big differences between these two conditions because both conditions are satisfied by many models.

Let $M_1(E)$ denote the space of all probability measures on E . For any $u \in M_1(E)$, we define an operator Q_u on $C(E)$, the space of all continuous functions on E , as

$$Q_u f(x) = \sum_{y \in E} q_{x,y} (f(y) - f(x)) + \|u\| (f(x+1) - f(x))$$

where $f \in C(E)$, $\|u\| = \int_E x u(dx)$. The equation

$$\frac{d\langle u(t), f \rangle}{dt} = \langle u(t), Q_{u(t)} f \rangle, \quad f \in C_k(E)$$

is called the nonlinear master equation, where $u(t)$ is a $M_1(E)$ -valued function and $C_k(E)$ is the space of all continuous functions on E with compact support. In order to solve this equation we introduce

Definition 3.1.1 (a). A probability measure $P_u \in M_1(D)$ is called a solution to the nonlinear martingale problem for Q with initial distribution $u \in M_1(E)$ if

$$(1.4) \quad P_u \circ x_0^{-1} = u.$$

$$(1.5) \quad \forall f \in C_k(E), \quad f(x_t) - \int_0^t Q_{\pi(P_u)(s)} f(x_s) ds \text{ is a } P_u - \text{martingale}$$

where $\pi(P_u)(s) = P_u \circ x_s^{-1}$.

(b). For every $u(\cdot) \in D([0, T], M_1(E))$, $u \in M_1(E)$, a probability measure $P_{u(\cdot), u} \in M_1(D)$ is called a solution to the time-inhomogeneous martingale problem for $Q_{u(t)}$ with initial distribution u if (1.4) and (1.5) are satisfied with $\pi(P_u)(s)$ replaced by $u(s)$.

The next theorem gives the sufficient conditions for the above martingale problems.

Theorem 3.1.1 Assume conditions (1.1)-(1.3) are satisfied. Then for any

$u \in M_1(E)$, $u(\cdot) \in D([0, T], M_1(E))$, satisfying

$$\langle \varphi, u \rangle = \int_E \varphi u(dx) < \infty, \quad \sup_{0 \leq t \leq T} \langle \varphi, u(t) \rangle < \infty,$$

we have

(a). The nonlinear martingale problem for Q with initial distribution u is well-posed.

(b). The time-inhomogeneous martingale problem for $Q_{u(t)}$ with initial distribution u is well-posed.

Proof: (a) and (b) are the result of theorem 1.1 and theorem 1.2 of [22]. The proof there is based on the martingale approach proposed by Holley, Stroock and Varadhan [28] and [40] combining the techniques from Funaki [25] and Zheng [44]. In [9] the existence of the solution to the nonlinear martingale problem was obtained as a limit of the exchangeable N -particle system. $\square$

The solution to the nonlinear martingale problem is a weak solution to the nonlinear master equation. For some concrete models a phase transition may occur because of the nonlinearity (cf. theorem 1.3 of [22]).

3.2 Mean Field Interacting Particle System

For any $N \geq 1$, the N -particle system we consider is a Markov process $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$ with generator $\Omega^{(N)}$ defined by

$$\Omega^{(N)} \psi(x^{(N)}) = \sum_{k=1}^N Q_{x_{s(N)}^{(k)}}^{(k)} \psi(x^{(N)})$$

where $\varepsilon_{x^{(N)}} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}}$, $x^{(N)} = (x_1^{(N)}, \dots, x_N^{(N)}) \in E^{\otimes N}$. $Q_\nu^{(k)}$ is used instead of Q_ν when it acts on the k -th variable of $\psi \in C(E^{\otimes N})$. This is a particle system with mean field interaction. To get the existence and uniqueness of the related semigroup and the martingale problem some general discussions on Markov process are needed.

Definition 3.2.1 Let S be any countable set. $Q = (q_{x,y})_{x,y \in S}$ is called a Q -matrix if

$$(2.1) \quad q_{x,y} \geq 0 \text{ for } x \neq y; \quad \sum_{y \in S} q_{x,y} \leq 0 \text{ for all } x \in S$$

$Q = (q_{x,y})_{x,y \in S}$ is called to be conservative if

$$(2.2) \quad \sum_{y \in S} q_{x,y} = 0 \text{ for all } x \in S.$$

A Q -matrix is said to be totally stable if

$$(2.3) \quad |q_{x,x}| < \infty \text{ for all } x \in S.$$

Definition 3.2.2 Let $\{P(t, x, y)\}_{x,y \in S}$ be the transition function of a time homogeneous Markov process. The process is called a Q -process if

$$(2.4) \quad \lim_{t \rightarrow 0} \frac{P(t, x, y)}{t} = q_{x,y} \text{ for } x \neq y$$

$$(2.5) \quad \lim_{t \rightarrow 0} \frac{P(t, x, x) - 1}{t} = q_{x,x}.$$

For convenience we usually call the transition function $\{P(t, x, y)\}_{x,y \in S}$ a Q -process if (2.4) and (2.5) are satisfied.

Definition 3.2.3 Let $\{P(t, x, y)\}_{x,y \in S}$ be a Q -process. $\{P(t, x, y)\}_{x,y \in S}$ is said to be a minimal Q -process if it is the minimal non-negative solution of the following equations:

$$(2.6) \quad \varphi(t, x, y) = \int_0^t e^{-q(x)(t-s)} \sum_{z \in S, z \neq x} q_{x,z} \varphi(s, z, y) ds + \delta_x(y) e^{-q(x)t}$$

$$(2.7) \quad \varphi(t, x, y) = \int_0^t \sum_{z \in S, z \neq y} \varphi(s, x, z) q_{z,y} e^{-q(y)(t-s)} + \delta_x(y) e^{-q(y)t}$$

where $q(x) = -q_{x,x}$, $t \in [0, \infty)$.

Usually it's convenient to discuss the following Laplace transform of (2.6) and (2.7):

$$\varphi(\lambda, x, y) = \sum_{z \in S, z \neq x} \frac{q_{x,z}}{q(x) + \lambda} \varphi(\lambda, z, y) + \frac{\delta_x(y)}{q(x) + \lambda}, \lambda > 0.$$

$$\varphi(\lambda, x, y) = \sum_{z \in S, z \neq y} \varphi(\lambda, x, z) \frac{q_{z,y}}{q(y) + \lambda} + \frac{\delta_x(y)}{q(y) + \lambda}, \lambda > 0.$$

Theorem 3.2.1 Let $Q = (q_{x,y})_{x,y \in S}$ be a conservative and totally stable Q -matrix. Then we have: (i). the Q -process always exists.

(ii). assume that there exists a sequence of subset $\{S_n\}_{n=1}^{\infty}$ of S and a positive function h on E satisfying:

$$(2.8) \quad S_n \nearrow S, \quad \sup_{x \in S_n} q(x) < \infty.$$

$$(2.9) \quad \lim_{n \rightarrow \infty} \inf_{x \notin E_n} h(x) = \infty$$

$$(2.10) \quad \exists C > 0, \exists \sum_{y \in S} q_{x,y}(h(y) - h(x)) \leq Ch(x).$$

then the Q -process is unique.

Proof: (i) The minimal non-negative solution of (2.6) and (2.7) always exists and is a Q -process.

(ii) (2.8), (2.9) and (2.10) implies that the minimal Q -process satisfies: $\forall t \in [0, \infty)$ and $x \in S, P(t, x, S) = 1$. This implies the uniqueness of the Q -process. For more details, refer to theorem 2.2.5 and theorem 2.3.7 in M.F.Chen [3]. $\square$

Corollary 3.2.2 Let $S = E^{\otimes r}, |x| = \sum_{k=1}^r x_k$ for $x \in S$. Assume that there exists a constant $\beta > 0$ such that

$$(2.11) \quad \sum_{y \neq x} q_{x,y}(|y| - |x|) \leq \beta(1 + |x|), \quad x \in S.$$

then the Q -process is unique.

Proof: Taking $S_n = \{x \in S : |x| \leq n\}$, $h(x) = |x| + 1$, $x \in S$, then by theorem 3.2.1, the result is true. $\square$

Now we return to the N -particle system $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$. If we take $S = E^{\otimes N}$ and define for each $x \in E^{\otimes N}$

$$(2.12) \quad q_{x,y}^{(N)} = \begin{cases} & \text{If } y \neq x, \\ q_{x_k, y_k} & y = (x_1, \dots, x_{k-1}, y_k, x_{k+1}, \dots, x_N) \\ & \text{for some } k, y_k \neq x_k + 1 \\ q_{x_k, x_{k+1}} + \frac{x_1 + \dots + x_N}{N} & \text{If } y = (x_1, \dots, x_k + 1, \dots, x_N) \\ & \text{for some } 1 \leq k \leq N \\ -\sum_{y \neq x} q_{x,y}^{(N)} & \text{If } y = x \\ 0 & \text{otherwise.} \end{cases}$$

Then we have

$$\Omega^{(N)} \psi(x^{(N)}) = \sum_{y^{(N)} \neq x^{(N)}} q_{x^{(N)}, y^{(N)}}^{(N)} (\psi(y^{(N)}) - \psi(x^{(N)}))$$

Thus $\Omega^{(N)}$ can be considered as a conservative and totally stable Q -matrix.

For each fixed N and $x^{(N)} = (x_1^{(N)}, \dots, x_N^{(N)}) \in S$, applying condition (1.2), we have:

$$(2.13) \quad \begin{aligned} & \sum_{y^{(N)} \neq x^{(N)}} q_{x^{(N)}, y^{(N)}}^{(N)} (|y^{(N)}| - |x^{(N)}|) \\ &= \sum_{k=1}^N \left[\sum_{y_k^{(N)} \neq x_k^{(N)}} q_{x_k^{(N)}, y_k^{(N)}}^{(N)} (y_k^{(N)} - x_k^{(N)}) + \frac{x_1^{(N)} + \dots + x_N^{(N)}}{N} \right] \\ &\leq \sum_{k=1}^N \left(\lambda(x_k^{(N)} + 1) + |x^{(N)}|/N \right) \leq 2\lambda N(|x^{(N)}| + 1) \end{aligned}$$

Combining theorem 3.2.1 and corollary 3.2.2, we get

Theorem 3.2.3 Assume conditions (1.1)-(1.3) are satisfied. Then for each $N \geq 1$, the Q -process for $\Omega^{(N)}$ exists and is unique.

Theorem 3.2.4 Assume conditions (1.1)-(1.3) are satisfied. Let $u \in M_1(E)$ with finite $\langle \varphi, u \rangle$. Then for each $N \geq 1$, the martingale problem for $\Omega^{(N)}$ on $D([0, T], E)^{\otimes N}$ with initial distribution $u^{\otimes N}$ is well-posed.

Proof: Combining theorem 3.2.1 (i) with proposition 1.7 of chapter 4 in [21], we get the existence of the minimal $Q^{(N)}$ -process which is a solution to the martingale problem.

To consider the uniqueness, let $P_u^{(N)}$ be any solution in $D([0, T], E)^{\otimes N}$ to the martingale problem for $\Omega^{(N)}$ with initial distribution $u^{\otimes N}$. $\forall f \in C_b(E^{\otimes N})$, we have

$$(2.14) \quad E^{P_u^{(N)}} \left[f(x^{(N)}(t)) - \int_0^t \Omega^{(N)} f(x^{(N)}(s)) ds \right] = E^{P_u^{(N)}} (f(x^{(N)}(0))).$$

If $P_u^{(N)}(t, y) = P_u^{(N)}(x^{(N)}(t) = y)$, $f(x) = \chi_{\{y\}}(x)$, then

$$(2.15) \quad P_u^{(N)}(t, y) - P_u^{(N)}(0, y) = \int_0^t \left[\sum_{z \neq y} q_{z,y}^{(N)} P_u^{(N)}(s, z) - q^{(N)}(y) P_u^{(N)}(s, y) \right] ds$$

which implies

$$(2.16) \quad P_u^{(N)}(t, y) = \int_0^t \sum_{z \neq y} P_u^{(N)}(s, y) q_{z,y}^{(N)} e^{-q^{(N)}(y)(t-s)} ds + P_u^{(N)}(0, y) e^{-q^{(N)}(y)t}$$

Let $(P_N^{\min}(t, x, y))_{x, y \in S}$ denote the minimal $Q^{(N)}$ -process with initial distribution δ_x .

Define

$$(2.17) \quad P_{N,u}^{\min}(t, y) = \sum_{x \in S} u^{\otimes N}(x) P_N^{\min}(t, x, y)$$

Then $P_{N,u}^{\min}(t, y)$ is the minimal nonnegative solution of (2.16).

Since $P_u^{(N)}(t, E^{\otimes N}) = 1 = P_{N,u}^{\min}(t, E^{\otimes N})$, we get that the one dimensional distribution of any solution in $D([0, T], E)^{\otimes N}$ to the martingale problem for $\Omega^{(N)}$ with initial distribution $u^{\otimes N}$ is the same as that of $P_{N,u}^{\min}$. Applying corollary 4.3 in chapter 4 of [21], we have $P_u^{(N)} = P_{N,u}^{\min}$ which implies the uniqueness of the solution.

□

Now we introduce a continuous map T_N as follow:

$$T_N : D([0, T], E^{\otimes N}) \longrightarrow D([0, T], M_1(E))$$

$$x(t) = (x_1(t), \dots, x_N(t)) \longmapsto \varepsilon_{x(t)} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k(t)}, \quad t \in [0, T]$$

For any $u \in M_1(E)$, let ν be the image of $u^{\otimes N}$ with respect to the map T_N : $\nu = u^{\otimes N} \circ T_N^{-1}$. Let $\tilde{P}_\nu^{(N)}$ be the image of $P_\nu^{(N)}$ with respect to T_N : $\tilde{P}_\nu^{(N)} = P_\nu^{(N)} \circ T_N^{-1}$. Define $\mathcal{G}_N(E) = \{ \{a_x/N; x \in E\}; a_x \in E, \sum_{x \in E} a_x = N \}$. Then $\varepsilon_{x^{(N)}(t)}$ is a measure-valued Markov process on $D([0, +\infty), \mathcal{G}_N(E))$ with generator

$$\begin{aligned} A_N F(\mu) = & \sum_{x, y \in E} a_x q_{x, y} \left\{ f(\langle \mu, g \rangle + \frac{g(y) - g(x)}{N}) - f(\langle \mu, g \rangle) \right\} \\ & + \sum_{x \in E} x a_x \sum_{y \in E} (a_y/N) \left\{ f(\langle \mu, g \rangle + \frac{g(y) - g(x)}{N}) \right. \\ & \left. - f(\langle \mu, g \rangle) \right\} \end{aligned}$$

where $\mu \in \mathcal{G}_N(E)$, $g \in C_b(E)$, $f \in C_b^2(R^1) = \{f : R^1 \rightarrow R^1, f'' \text{ is bounded}\}$ and $F(\mu) = f(\langle \mu, g \rangle)$. We have

Corollary 3.2.5 $F(\mu(t)) - \int_0^t A_N F(\mu(s)) ds$ is a local $\tilde{P}_\nu^{(N)}$ -martingale for every $g \in C_b(E)$ and $f \in C_b^2(R^1)$.

Proof: Let

$$\begin{aligned} H^g(u(t)) &= F(\mu(t)) - \int_0^t A_N F(\mu(s)) ds \\ M_t^\psi(x^{(N)}(t)) &= \psi(x^{(N)}(t)) - \psi(x^{(N)}(0)) - \int_0^t \Omega^{(N)} \psi(x^{(N)}(s)) ds \end{aligned}$$

where $\psi(x^{(N)}) = F(\varepsilon_{x^{(N)}})$. Then M_t^ψ is a local martingale with respect to $P_\nu^{(N)}$.

Since $\tilde{P}_\nu^{(N)} = P_\nu^{(N)} \circ T_N^{-1}$, $M_t^\psi = H^g \circ T_N^{-1}$, we get the result. $\square$

3.3 Cameron-Martin-Girsanov Formula

In this section we assume that the Q -process satisfies:

$$\inf_{x \in E} q_{x, x+1} > 0.$$

The following will be used in the sequel.

For $x, x', y \in E$, define

$$(3.1) \quad Q(x, x', y) = \begin{cases} q_{x, x+1} + x' & \text{for } y = x + 1 \\ q_{x, y} & \text{for } y \neq x, y \neq x + 1 \\ 0 & \text{otherwise} \end{cases}$$

For each $u \in M_1(E)$, define

$$(3.2) \quad Q_u(x, y) = \int_E Q(x, x', y) u(dx')$$

Then we have

$$\Omega^{(N)} \psi(x^{(N)}) = \sum_{k=1}^N \int_E Q_{\epsilon_{x^{(N)}}}(x_i^{(N)}, dy) \Delta_i^y \psi(x^{(N)}) \quad \text{for } \psi \in C(E^{\otimes N}).$$

where

$$\begin{aligned} \Delta_i^y \psi(x^{(N)}) &= \psi(x_1^{(N)}, \dots, x_{i-1}^{(N)}, y, x_{i+1}^{(N)}, \dots, x^{(N)}) \\ &\quad - \psi(x_1^{(N)}, \dots, x_{i-1}^{(N)}, x_i^{(N)}, x_{i+1}^{(N)}, \dots, x^{(N)}). \end{aligned}$$

For each $\omega \in D$ and $A \subset E$, define $Q_{u(s)}(\omega(s), A) = \sum_{y \in A} Q_{u(s)}(\omega(s), y)$ and

$$(3.3) \quad N(t, A; \omega) = \#\{s : \omega(s) \in A, \omega(s) \neq \omega(s-), s \leq t\}$$

$$(3.4) \quad \hat{N}(t, A; \omega) = N(t, A; \omega) - \int_0^t Q_{u(s)}(\omega(s), A) ds$$

where $u(\cdot) \in D([0, T], M_1(E))$ satisfies $\sup_{0 \leq s \leq T} \langle \varphi, u(s) \rangle < \infty$.

For every $x, x', y \in E, u, \nu \in M_1(E)$, define

$$(3.5) \quad q_u(x, x', y) = \frac{Q(x, x', y)}{Q_u(x, y)}, \quad q_\nu^\nu(x, y) = \frac{Q_\nu(x, y)}{Q_u(x, y)}.$$

For every $u \in M_1(E)$ and $u(\cdot) \in D([0, T], M_1(E))$ which satisfies

$$\sup_{0 \leq t \leq T} \langle \varphi, u(t) \rangle < \infty \quad \text{and} \quad \langle \varphi, u \rangle < \infty,$$

let $P_{u(\cdot), u}$ be the unique solution to the time-inhomogeneous martingale problem for $Q_{u(\cdot)}$. Then we have

Lemma 3.3.1 *For every finite subset A of E*

$$(3.6) \quad \tilde{N}(t, A; \omega) \text{ is a } P_{u(\cdot), u} - \text{martingale.}$$

Proof: Noting that $\forall f \in C_k(E)$, we have

$$(3.7) \quad f(\omega(t)) - f(\omega(0)) = \int_{[0,t] \times E} (f(y) - f(\omega(s-))) \tilde{N}(ds, dy; \omega)$$

and

$$(3.8) \quad f(\omega(t)) - \int_0^t Q_{u(s)}(f(\omega(s))) ds$$

are $P_{u(\cdot), u}$ - martingales. This implies

$$(3.9) \quad \int_{[0,t] \times E} (f(y) - f(\omega(s-))) \tilde{N}(ds, dy; \omega) \text{ is a } P_{u(\cdot), u} - \text{martingale.}$$

and $\forall g \in C_b(E)$

$$(3.10) \quad \int_{[0,t] \times E} g(\omega(s-)) (f(y) - f(\omega(s-))) \tilde{N}(ds, dy; \omega) \text{ is a } P_{u(\cdot), u} - \text{martingale.}$$

In particular, if $g(x)f(x) = 0, \forall x \in E$, then

$$(3.11) \quad \int_{[0,t] \times E} g(\omega(s-)) f(y) \tilde{N}(ds, dy; \omega) \text{ is a } P_{u(\cdot), u} - \text{martingale.}$$

Since for fixed ω , $\tilde{N}(ds, dy; \omega) = \chi_{\{s: \omega(s) \in dy, \omega(s) \neq \omega(s-)\}} - Q_{u(s)}(\omega(s), dy) ds$, we have

$$(3.12) \quad \tilde{N}((s, y); \omega(s-)) = y, 0 \leq s \leq T) = 0,$$

and thus (3.11) is valid for any f and g . For every $h \in C(E^2)$ with compact support with respect to the second variable we have

$$(3.13) \quad \int_{[0,t] \times E} h(y, \omega(s-)) \tilde{N}(ds, dy; \omega) \text{ is a } P_{u(\cdot), u} - \text{martingale.}$$

If we take $h(y, \omega(s-)) = \chi_A(y)$, we get the result. $\square$

Lemma 3.3 of [9] shows that $E\tilde{N}(t, E; \omega) < \infty$. This combined with lemma 3.3.1 implies that $\tilde{N}(t, A)$ is of the class (QL) for any subset A of E .

For any integer $N \geq 1$, let $P_u^{(N)}$ be the unique solution to the martingale problem for $\Omega^{(N)}$ with initial distribution $u^{\otimes N}$. Let $P_{u(\cdot), u}^{(N)}$ be the N -fold independent products of $P_{u(\cdot), u}$. Then we have

Theorem 3.3.2 $P_{u(\cdot),u}^{(N)}$ and $P_u^{(N)}$ are mutually absolutely continuous and

$$(3.14) \quad \frac{dP_{u(\cdot),u}^{(N)}}{dP_u^{(N)}} = \exp(H_T^{(N)}(\omega)), \text{ where}$$

$$(3.15) \quad H_T^{(N)}(\omega) = \sum_{i=1}^N \int_{[0,T] \times E} \{ \log q_{u(s-)}^{\pi(\eta_{x^{(N)}})(s-)}(\omega_i(s-), y) \} N(ds, dy; \omega_i) \\ - \sum_{i=1}^N \int_0^T \{ \| \pi(\eta_{x^{(N)}})(s) \| - \| u(s) \| \} ds.$$

Proof: Lemma 2.5 of [38] proved this result in the case when the elements of the Q -matrix are bounded. For the unbounded case refer to lemma 3.7 in [9]. $\square$

For the N -particle system $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$ and any $A \in \mathcal{E}^{\otimes N}$, we can define

$$(3.16) \quad N(t, A; x^{(N)}) = \# \{ s : x^{(N)}(s) \in A, x^{(N)}(s) \neq x^{(N)}(s-), s \leq t \}$$

$$(3.17) \quad \hat{N}(t, A; x^{(N)}) = N(t, A; x^{(N)}) - \int_0^t \Omega^{(N)}(x^{(N)}(s), A) ds$$

where $\Omega^{(N)}(x^{(N)}(s), A) = \sum_{y^{(N)} \in A} q_{x^{(N)}(s), y^{(N)}}^{(N)}$.

Using an argument similar to the proof of lemma 3.3.1, we can show that $\hat{N}(t, A; x^{(N)})$ is a martingale with respect to $P_u^{(N)}$.

Since $\{x_i^{(N)}(t)\}_{i=1}^N$ have no common jumps $P_{u(\cdot),u}^{(N)}$ -almost surely, we have

$$(3.18) \quad N(dt, dy^{(N)}; x^{(N)}) = \sum_{i=1}^N N(dt, dy_i; x_i^{(N)}) \quad P_{u(\cdot),u}^{(N)} \text{ - almost surely.}$$

Theorem 3.3.2 implies that $\{x_i^{(N)}(t)\}_{i=1}^N$ have no common jumps $P_u^{(N)}$ -almost surely.

3.4 Law of Large Numbers

Consider the Q -matrix satisfying the following extra conditions:

$$(4.1) \quad \exists m \geq 1, \text{ such that } \forall \text{ integer } k, f(x) = q_{x, x+k}$$

is a polynomial of order less than m ,

$$(4.2) \quad \sum_{y \in E} q_{x,y} (y^r - x^r) \leq M(r) + \lambda(r) x^r, \quad r = 1, \dots, m;$$

where $M(r)$ and $\lambda(r)$ are positive constants related to r .

Given a solution $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$ associated to $\Omega^{(N)}$, let

$$\varepsilon_{x^{(N)}(t)} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}(t)}, \quad t \geq 0.$$

denote the empirical process. The following law of large numbers was shown in theorem 1.1 of [9].

Theorem 3.4.1 *If $\varepsilon_{x^{(N)}(0)} \rightarrow \mu(0)$ weakly,*

$$\int_E x^{m+2} \mu(0, dx) < \infty,$$

and

$$\sup_N E^{P^{(N)}} [N^{-1} \sum_{i=1}^N (x_i^{(N)}(0))] < \infty,$$

then $\varepsilon_{x^{(N)}(t)}$ converges in the sense of weak convergence of measure-valued stochastic processes to the unique solution of the nonlinear martingale problem associated with the nonlinear master equation.

Combined with the uniqueness of the solution to the nonlinear martingale problem the proof of this theorem is completed by verifying the weak compactness, the latter follows by checking the following two conditions:

$$\sup_{N, t \leq T} E^{P^{(N)}} (\langle \varepsilon_{x^{(N)}(t)}, h \rangle) < \infty$$

and $\forall k \in E$, $\langle \varepsilon_{x^{(N)}(\cdot)}, I_{E_k} \rangle$ are relative compact in $D([0, T], R)$. Where $h/x \rightarrow \infty$ as $x \rightarrow \infty$, I_{E_k} is the indicator function of $E_k = \{0, 1, \dots, k\}$.

Remark. Although the conditions in this section seem stronger than the assumptions in section 3.1. But in fact many chemical reaction models satisfy these conditions. It is expected that under some weaker assumptions on the Q -matrix, for example replacing (4.2) by (4.8) in chapter 1, the law of large numbers still holds. But some new methods are needed.

Chapter 4

Large Deviations For Empirical Distribution

For every $x^{(N)}(\cdot) \in D([0, T], E^{\otimes N})$, define

$$\eta_{x^{(N)}(\cdot)} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}(\cdot)}$$

Let $\mathcal{P}_u^{(N)}(\cdot) = P_u^{(N)}(\eta_{x^{(N)}(\cdot)} \in \cdot)$.

In this chapter, we will get the large deviation estimates for $\{\mathcal{P}_u^{(N)}\}_{N=1,2,\dots}$ as $N \rightarrow \infty$.

4.1 Assumptions

For fixed $T > 0$, let $D = D([0, T], E)$. $\mathcal{M} = M_1(D)$ denotes the space of all probability measures on D with weak topology. For each $R \geq 1$, define $\mathcal{M}_R = \{P \in \mathcal{M} : \sup_{0 \leq t \leq T} \int \varphi(\omega(t)) P(d\omega) \leq R\}$ equipped with the subspace topology of $\mathcal{M}$, where φ is defined as $\varphi : E \rightarrow \mathbf{R}, x \rightarrow 1 + x \log \log(x+2)$. Let $\mathcal{M}_\infty = \bigcup_{R \geq 1} \mathcal{M}_R$ be equipped with the "inductive topology". A set V is open in $\mathcal{M}_\infty$ if and only if for each $R \geq 1$ $V \cap \mathcal{M}_R$ is open in $\mathcal{M}_R$. The "inductive topology" is stronger than the weak topology and $\mathcal{M}_\infty$ is not a Polish space.

To get the large deviation principle some conditions are needed. Let $Q = (q_{x,y})_{x,y \in E}$ be a totally stable conservative Q -matrix satisfying:

$$(1.1) \quad \inf_{x \in E} \{q_{x,x+1}\} > 0;$$

$$(1.2) \quad \exists \Lambda > 0, \text{ such that } q_{x,y} = 0 \text{ for } |x - y| \geq \Lambda;$$

$$(1.3) \quad \exists \lambda > 0 \text{ such that } \sum_{y \in E} q_{x,y}(y - x) \leq \lambda(x + 1).$$

$$\sum_{y \in E} q_{x,y}(\varphi(y) - \varphi(x)) \leq \lambda \varphi(x), \quad x \in E;$$

$$(1.4) \quad \exists 0 < c < \infty, \text{ such that } \forall x, y \in E, y > x,$$

$$\begin{aligned} & \sum_{z \neq 0} (q_{y,y+z} - q_{x,x+z})z \\ & + 2 \sum_{z=1}^{\infty} [(q_{y,x-z} - q_{x,2x-y-z}) \vee 0 + (q_{x,y+z} + q_{y,2y-x+z}) \vee 0]z \leq c(y - x); \end{aligned}$$

$$(1.5) \quad \forall l \geq 0, \exists \lambda(l) > 0, \exists$$

$$\begin{aligned} & \sup_{x \in E} \max \left\{ \sum_{y \in E} q_{x,y}(\epsilon^{y-x} - 1) + (\epsilon - 1)x, \right. \\ & \left. \sum_{y \in E} q_{x,y}(\epsilon^{\varphi(y)-\varphi(x)} - 1) + (\epsilon^{\varphi(x+1)-\varphi(x)} - 1)l \right\} \leq \lambda(l). \end{aligned}$$

Remark. It is easy to check that (1.5) implies (1.3). We write them separately just for convenience.

Noting that with our choice of φ , for every $x \geq 1$

$$\begin{aligned} & e^{\varphi(x+1)-\varphi(x)} - 1 \\ & = \exp[(x+1) \log \log(x+3) - x \log \log(x+2)] - 1 \\ & = \left(\frac{\log(x+3)}{\log(x+2)} \right)^x \log(x+3) - 1 \\ & \leq \left(\frac{x+3}{x+2} \right)^x \log(x+3) - 1 \end{aligned}$$

we can easily check that the assumptions (1.1)-(1.5) are satisfied by the following models:

The first Schlögl model:

$$q_{x,y} = \begin{cases} \beta_0 + \beta_1 x & \text{for } y = x + 1, x \geq 0 \\ \alpha_1 x + \alpha_2 x(x-1) & \text{for } y = x - 1, x \geq 1 \\ -\sum_{y \neq x} q_{x,y} & \text{for } y = x \\ 0 & \text{otherwise} \end{cases}$$

The second Schlögl model:

$$q_{x,y} = \begin{cases} \beta_0 + \beta_2 x(x-1) & \text{for } y = x + 1, x \geq 0 \\ \alpha_1 x + \alpha_3 x(x-1)(x-2) & \text{for } y = x - 1, x \geq 1 \\ -\sum_{y \neq x} q_{x,y} & \text{for } y = x \\ 0 & \text{otherwise} \end{cases}$$

where the coefficients are all positive constants. For the background of Schlögl models refer to [37].

In general, if (1.1) and (1.2) are satisfied and the jump rates are polynomials of finite order, a sufficient condition for (1.3) and (1.4) to be true is that the order of the birth rates plus Λ is less than that of the death rates. Condition (1.4) is needed to guarantee the uniqueness of the solution to the nonlinear martingale associated with Q (cf. [22]). (1.5) is a condition needed for getting the large deviations.

4.2 The Main Result

For $u \in \mathcal{M}_1(E)$, $u(\cdot) \in D([0, T], \mathcal{M}_1(E))$ with $\langle \varphi, u \rangle < \infty$ and $\sup_{0 \leq t \leq T} \langle \varphi, u(t) \rangle < \infty$. We define a function $I_u^{u(\cdot)} : \mathcal{M}_\infty \rightarrow [0, \infty]$ as

$$(2.1) \quad I_u^{u(\cdot)}(P) = \sup_{\Phi \in C_b(D)} \left\{ \int \Phi dP - \log E^{P_{u(\cdot), u}}(e^\Phi) \right\},$$

where $P \in \mathcal{M}_\infty$, $C_b(D)$ is the set of all bounded continuous functions on D .

For $P \in \mathcal{M}_\infty$, define

$$(2.2) \quad \pi(P) : [0, T] \longrightarrow M_1(E), \quad \pi(P)(t) = P \circ x_t^{-1}$$

For any $P \in \mathcal{M}_\infty$ let I_u be defined as

$$(2.3) \quad I_u : \mathcal{M}_\infty \longrightarrow [0, \infty], \quad I_u(P) = I_u^{\pi(P)}(P).$$

Here is the main result of this chapter:

Theorem 4.2.1 *For every $u \in M_1(E)$ with $\int_E e^{\varphi(x)} u(dx) < \infty$, $(\mathcal{M}_\infty, \mathcal{P}_u^{(N)}, N)$ is a large deviation system with action functional I_u under conditions (1.1)–(1.5).*

4.3 Some Properties of the Space $\mathcal{M}_\infty$ and the Functional I_u

Lemma 4.3.1 *A sequence $\{P_n\}$ converges to P in $\mathcal{M}_\infty$ iff $\{P_n\}$ converges to P weakly in $\mathcal{M}$ and there is a $R > 0$ such that*

$$(3.1) \quad \sup_n \sup_{0 \leq t \leq T} \int \varphi(\omega(t)) P_n(d\omega) \leq R.$$

Proof: “ $\Leftarrow$ ” Assume that $\{P_n\}$ converges to P weakly in $\mathcal{M}$ and there is an $R > 0$ such that

$$\sup_n \sup_{0 \leq t \leq T} \int \varphi(\omega(t)) P_n(d\omega) \leq R.$$

Then $\{P_n\}$ converges to P weakly in $\mathcal{M}_R$. By the definition of $\mathcal{M}_\infty$ we get the result.

“ $\Rightarrow$ ” $\{P_n\}$ converges to P in $\mathcal{M}_\infty$ implies that $\{P_n\}$ converges to P weakly in $\mathcal{M}$ since the inductive topology is stronger than the weak topology.

For the proof of (3.1), let's assume the contrary. Then without loss of generality, we can assume that

$$(3.2) \quad \sup_{0 \leq t \leq T} \int \varphi(\omega(t)) P_n(d\omega) \nearrow \infty.$$

Let $S_n = \{s \in [0, T] : P_n(\omega(s) = \omega(s-)) = 1\}^c$. By lemma 7.7 of chapter 3 in [21], S_n is at most countable. Let $S = (\cup S_n)^c$. Then S is dense in $[0, T]$ and its complement in $[0, T]$ is at most countable.

(3.2) implies that there exists a sequence $\{t_n\} \subset S$ such that

$$(3.3) \quad \int \varphi(\omega(t_n)) P_n(d\omega) \nearrow \infty.$$

Since $\omega(t) \wedge M \rightarrow \omega(t)$ as $M \rightarrow \infty$, we can find a sequence $\{M_n\}$ tending to ∞ such that

$$(3.4) \quad \int \varphi(\omega(t_n) \wedge M_n) P_n(d\omega) \nearrow \infty.$$

Let $R > 0$ be such that $\int \varphi(\omega(t)) P(d\omega) < R$. Take

$$m = \inf\{n : \int \varphi(\omega(t_n) \wedge M_n) P_n(d\omega) > R\}.$$

$$V = \cap_{n \geq m} \{P \in \mathcal{M}_\infty : \int \varphi(\omega(t_n) \wedge M_n) dP < \int \varphi(\omega(t_n) \wedge M_n) dP_n\}.$$

For any $R_0 > 0$, we have for some $k > m$

$$(3.5) \quad V \cap \mathcal{M}_{R_0} = \mathcal{M}_{R_0} \cap \cap_{n=m}^k \{P \in \mathcal{M}_\infty : \int \varphi(\omega(t_n) \wedge M_n) dP < \int \varphi(\omega(t_n) \wedge M_n) dP_n\}.$$

Noting that the map $\omega \mapsto \omega(t_n) \wedge M_n$ is continuous for any fixed n , we have that $V \cap \mathcal{M}_{R_0}$ is open in $\mathcal{M}_{R_0}$ and thus V is open in $\mathcal{M}_\infty$. Since $\{P_n\}_{n \geq m} \cap V = \emptyset$, we get that $\{P_n\}$ does not converge to P weakly in $\mathcal{M}$. This is a contradiction. Hence (3.1) is true. $\square$

Lemma 4.3.2 *A real-valued function f on $\mathcal{M}_\infty$ is continuous iff it is sequentially continuous.*

Proof: It suffices to verify that the sequential continuity of f implies it's continuity.

Since $\mathcal{M}_R$ is metrizable, the restriction of f on $\mathcal{M}_R$ is continuous and thus f is continuous. For more detail refer to the Appendix in [26]. $\square$

Lemma 4.3.3 *Under (1.2), (1.3) and (1.5), the functional I_u is lower semicontinuous on $\mathcal{M}_\infty$ for every $u \in M_1(E)$ with $\langle \varphi, u \rangle < \infty$.*

Proof: Since the supremum of continuous functions is lower semicontinuous, it suffices to verify that the map

$$\mathcal{M}_\infty \longrightarrow \mathcal{M}, \quad P \longmapsto P_{\pi(P)(\cdot), u}$$

is sequentially continuous. The map is well defined because the martingale problem for $Q_{\pi(P)(\cdot)}$ with initial distribution u is well-posed.

Let P_n be any convergent sequence in $\mathcal{M}_\infty$ with P as its limit point. Then we have

$$(3.6) \quad P_n \longrightarrow P \text{ weakly.}$$

$$(3.7) \quad \sup_{n \geq 1} \sup_{0 \leq t \leq T} \int \varphi(\omega(t)) P_n(d\omega) < \infty.$$

This combined with lemma 7.7 of chapter 3 in [21] implies

$$\pi(P_n)(t) \text{ converges weakly to } \pi(P)(t) \text{ almost surely on } [0, T].$$

$$\|\pi(P_n)(t)\| \text{ converges to } \|\pi(P)(t)\| \text{ almost surely on } [0, T].$$

In next lemma we will prove that $\{P_{\pi(P_n), u}\}$ is relatively compact.

It is easy to see that every limit point of $\{P_{\pi(P_n), u}\}$ is a solution to the time inhomogeneous martingale problem for $Q_{\pi(P)(\cdot)}$ with initial distribution u . But the latter is well-posed, we have that the map is continuous. Hence I_u as the supremum of continuous functions is lower semicontinuous. $\square$

In order to verify the relative compactness of $\{P_{\pi(P_n), u}\}$, we introduce the following notations:

$$\tau_0(\omega) = 0, \tau_n(\omega) = \inf\{t \geq \tau_{n-1} : \rho(x_t, x_{\tau_{n-1}}) \geq \epsilon/2\}.$$

$$N_T(\omega) = \min\{n : \tau_{n+1}(\omega) > T\}.$$

$$\delta_\omega^T(\epsilon) = \min\{\tau_n(\omega) - \tau_{n-1}(\omega) : 1 \leq n \leq N_T(\omega),$$

where $\omega \in D([0, T], E)$, ρ is the discrete metric on E .

Let $\sigma_m = \inf\{t \geq 0 : x_t \geq m\}$, $t \in [0, T]$. Combining (1.5) and (3.i) one gets that there exists an $M(T) > 0$ such that

$$(3.8) \quad Q_{\pi(P_n)(t)} f(x) \leq M(T) f(x), \text{ for } f(x) = x + 1.$$

Take $f_R(x) = x \wedge R + 1$. Then $f_R(x_t) - \int_0^t Q_{\pi(P_n)(s)} f_R(x_s) ds$ is a $P_{\pi(P_n)(\cdot), u}$ -martingale, i.e., $\forall 0 \leq t_1 \leq t_2$ and $A \in \mathcal{F}_{t_1}$, we have

$$\begin{aligned}
 (3.9) \quad & E^{P_{\pi(P_n)(\cdot), u}}[f_R(x_{t_2}); A] \\
 &= E^{P_{\pi(P_n)(\cdot), u}}[f_R(x_{t_1}); A] + E^{P_{\pi(P_n)(\cdot), u}}\left[\int_{t_1}^{t_2} Q_{\pi(P_n)(s)} f_R(x_s) ds; A\right] \\
 &\leq E^{P_{\pi(P_n)(\cdot), u}}[f(x_{t_1}); A] + E^{P_{\pi(P_n)(\cdot), u}}\left[\int_{t_1}^{t_2} Q_{\pi(P_n)(s)} M(T) f(x_s) ds; A\right].
 \end{aligned}$$

This implies

$$\begin{aligned}
 (3.10) \quad & E^{P_{\pi(P_n)(\cdot), u}}[f(x_{t_2}); A] \\
 &\leq E^{P_{\pi(P_n)(\cdot), u}}[f(x_{t_1}); A] + M(T) \int_{t_1}^{t_2} E^{P_{\pi(P_n)(\cdot), u}}[f(x_s); A] ds.
 \end{aligned}$$

By Gronwall's lemma we have

$$(3.11) \quad E^{P_{\pi(P_n)(\cdot), u}}[f(x_{t_2}); A] \leq E^{P_{\pi(P_n)(\cdot), u}}[f(x_{t_1}); A] e^{M(T)(t_2 - t_1)}.$$

This means that $Y_t = f(x_t) e^{-M(T)t}$ is a $P_{\pi(P_n)(\cdot), u}$ -supermartingale. Hence

$$\begin{aligned}
 (3.12) \quad & e^{-M(T)T} m P_{\pi(P_n)(\cdot), u}\{\sigma_m < T\} \leq E^{P_{\pi(P_n)(\cdot), u}}[f(x_{T \wedge \sigma_m}) e^{-M(T)(T \wedge \sigma_m)}] \\
 &\leq E^{P_{\pi(P_n)(\cdot), u}}(Y_0) = \|u\| + 1.
 \end{aligned}$$

Thus

$$(3.13) \quad P_{\pi(P_n)(\cdot), u}\{\sigma_m < T\} \leq c(T)/m,$$

where $c(T) = (\|u\| + 1) e^{M(T)T}$.

Let $B_m = \{0, 1, \dots, m\}$. Then B_m is compact in E and by (3.13) we have as $m \rightarrow \infty$

$$(3.14) \quad \inf_n P_{\pi(P_n)(\cdot), u}\{x_t \in B_m : 0 \leq t \leq T\} \geq \inf_n P_{\pi(P_n)(\cdot), u}\{\sigma_m \geq T\} \rightarrow 1.$$

Lemma 4.3.4 Under assumptions (1.2), (1.3) and (1.5), $\{P_{\pi(P_n)(\cdot), u}\}$ is a relatively compact subset of $\mathcal{M}$.

Proof: Applying theorem 14.3 in [2] and theorem 1.1.4 in [40], we have that a sequence P_n in $M_1(D)$ is relatively compact if and only if these two conditions hold:

(i) For each positive η , there exists an u such that

$$(3.15) \quad \sup_n P_n \{ \omega : \sup_{0 \leq t \leq T} \omega_t > a \} < \eta,$$

(ii) For each positive ϵ and η , there exists a $\delta, 0 < \delta < 1$, and an integer n_0 such that

$$(3.16) \quad \sup_n P_n \{ \omega : w'_\omega(\delta) \geq \epsilon \} < \eta,$$

where $w'_\omega(\delta) = \inf_{\{t_i\}} \max_{0 < i \leq r} \sup_{s, t \in [t_{i-1}, t_i]} \rho(\omega_s, \omega_t)$. $\{t_i\}$ satisfies $0 = t_0 < t_1 < \dots < t_{r-1} < T = t_r$ and $\min_{0 < i \leq r} (t_i - t_{i-1}) \geq \delta > 0$.

It's not difficult to check that

$$(3.17) \quad \{ \delta_\omega^T(\epsilon) > \delta \} \subset \{ w'_\omega(\delta) \geq \epsilon \}.$$

Since (3.14) implies (3.15), for the proof of lemma 4.3.4 we need only to prove that

$$(3.18) \quad \sup_n P_{\pi(P_n)(\cdot), u} \{ \delta_\omega^T(\epsilon) < \delta \} \longrightarrow 0 \text{ as } \delta \longrightarrow 0.$$

For any $L > 0$, let $A_L = \{ \omega : \sup_{0 \leq t \leq T} \omega_t \leq L \}$. Then

$$(3.19) \quad P_{\pi(P_n)(\cdot), u} \{ \delta_\omega^T(\epsilon) < \delta \} \leq P_{\pi(P_n)(\cdot), u} \{ \delta_\omega^T(\epsilon) < \delta, A_L \} + P_{\pi(P_n)(\cdot), u} \{ A_L^c \}.$$

$$(3.20) \quad \begin{aligned} P_{\pi(P_n)(\cdot), u} \{ \delta_\omega^T(\epsilon) < \delta, A_L \} &\leq P_{\pi(P_n)(\cdot), u} \{ \delta_\omega^T(\epsilon) < \delta, A_L, N_T(\omega) \leq k \} \\ &\quad + P_{\pi(P_n)(\cdot), u} \{ A_L, N_T(\omega) > k \}. \end{aligned}$$

$$(3.21) \quad \begin{aligned} &P_{\pi(P_n)(\cdot), u} \{ \delta_\omega^T(\epsilon) < \delta, A_L, N_T(\omega) \leq k \} \\ &= \sum_{i=1}^k P_{\pi(P_n)(\cdot), u} \{ \delta_\omega^T(\epsilon) < \delta, A_L, N_T(\omega) = i \} \\ &\leq \sum_{i=1}^k \sum_{j=1}^i P_{\pi(P_n)(\cdot), u} \{ \tau_j - \tau_{j-1} < \delta, A_L, N_T(\omega) = i \} \\ &\leq \sum_{i=1}^k \sum_{j=1}^i P_{\pi(P_n)(\cdot), u} \{ \tau_j - \tau_{j-1} < \delta, \forall_{l=1}^{j-1} x_{\tau_l} \leq L, \tau_{j-1} \leq T \} \\ &\leq \sum_{i=1}^k \sum_{j=1}^i E^{P_{\pi(P_n)(\cdot), u}} \{ P_{\pi(P_n)(\cdot), u}^{j-1} [I_{\{x_{\tau_{j-1}}\}^c}(x_{\tau_j \wedge (\tau_{j-1} + \delta)}) = 1, \end{aligned}$$

$$\begin{aligned}
 & \{ \bigvee_{i=1}^{j-1} x_{\tau_i} \leq L, \tau_{j-1} \leq T \} \\
 & \leq \sum_{i=1}^k \sum_{j=1}^i E^{P_{\pi(P_n)(\cdot),u}} \{ E^{P_{\pi(P_n)(\cdot),u}^{j-1}} [\int_{\tau_{j-1}}^{\tau_j \wedge (\tau_{j-1} + \delta)} \Omega_{\|\pi(P_n)(s)\|} I_{\{x_{\tau_{j-1}}\}^c}(x_s) ds] ; \\
 & \quad \{ \bigvee_{i=1}^{j-1} x_{\tau_i} \leq L, \tau_{j-1} \leq T \} \\
 & \leq \sum_{i=1}^k \sum_{j=1}^i E^{P_{\pi(P_n)(\cdot),u}} \{ E^{P_{\pi(P_n)(\cdot),u}^{j-1}} [- \int_{\tau_{j-1}}^{\tau_j \wedge (\tau_{j-1} + \delta)} \Omega_{\|\pi(P_n)(s)\|} I_{\{x_{\tau_{j-1}}\}}(x_s) ds] ; \\
 & \quad \{ \bigvee_{i=1}^{j-1} x_{\tau_i} \leq L, \tau_{j-1} \leq T \} \\
 & \leq \sum_{i=1}^k \sum_{j=1}^i C(L, T) \delta = \frac{k(k+1)}{2} C(L, T) \delta
 \end{aligned}$$

where $P_{\pi(P_n)(\cdot),u}^{j-1}$ is the r.c.p.d. of $P_{\pi(P_n)(\cdot),u}$ given $\mathcal{F}_{\tau_{j-1}}$, $C(L, T) = \sup_{0 \leq y \leq L} (|q_{y,y}| + \sup_{z \in E, z \neq y} |q_{z,y}|) + \sup_n \sup_{0 \leq s \leq T} \|\pi(P_n)(s)\| < \infty$.

We also have

$$\begin{aligned}
 (3.22) \quad & P_{\pi(P_n)(\cdot),u} \{ A_L, N_T(\omega) > k \} = P_{\pi(P_n)(\cdot),u} \{ \tau_k \leq T, A_L \} \\
 & \leq E^{P_{\pi(P_n)(\cdot),u}} [e^{T-\tau_k}; \tau_k \leq T, A_L] \\
 & \leq e^T E^{P_{\pi(P_n)(\cdot),u}} \{ E^{P_{\pi(P_n)(\cdot),u}^{k-1}} [e^{\tau_k - \tau_{k-1}} ; \\
 & \quad \tau_k \leq T, A_L] e^{-\tau_{k-1}} \} \\
 & \leq e^T E^{P_{\pi(P_n)(\cdot),u}} \{ e^{-\tau_{k-1}} [e^{-t_0} P_{\pi(P_n)(\cdot),u}^{k-1}(\tau_k - \tau_{k-1} \leq t_0, \tau_k \leq T, A_L) \\
 & \quad + P_{\pi(P_n)(\cdot),u}^{k-1}(\tau_k - \tau_{k-1} > t_0, \tau_k \leq T, A_L)] \} \\
 & \leq e^T E^{P_{\pi(P_n)(\cdot),u}} \{ [e^{-\tau_{k-1}} e^{-t_0} + (1 - e^{-t_0}) P_{\pi(P_n)(\cdot),u}^{k-1}(\tau_k - \tau_{k-1} \leq t_0)] ; \\
 & \quad \tau_{k-1} \leq T, \bigvee_{i=1}^{k-1} x_{\tau_i} \leq L \} \\
 & \leq e^T E^{P_{\pi(P_n)(\cdot),u}} \{ e^{-\tau_{k-1}} (e^{-t_0} + (1 - e^{-t_0}) C(L, T) t_0) ; \\
 & \quad \tau_{k-1} \leq T, \bigvee_{i=1}^{k-1} x_{\tau_i} \leq L \}
 \end{aligned}$$

Let $g(t) = e^{-t} + (1 - e^{-t})C(L, T)t$, we have

$$g(0) = 1, g'(0) = -1.$$

Thus for fixed $L > 0$ there is a $t_0 \in (0, 1)$ such that $0 < g'(t_0) < 1$.

Denote this $g(t_0)$ by $\Gamma(L, T)$. By induction we have

$$(3.23) \quad P_{\pi(P_n)(\cdot), u} \{A_L, N_T(\omega) > k\} \leq \Gamma(L, T)^k e^T.$$

Combining (3.19)-(3.23) and (3.14), we get

$$(3.24) \quad \sup_n P_{\pi(P_n)(\cdot), u} \{\delta_\omega^T(\epsilon) < \delta\} \leq \frac{k(k+1)}{2} C(L, T) \delta + \Gamma(L, T)^k e^T + \gamma_L,$$

where $\gamma_L = \sup_n P_{\pi(P_n)(\cdot), u} \{A_L^c\}$.

Let $\delta \rightarrow 0$, then $k \rightarrow \infty$. Finally let $L \rightarrow \infty$, we get (3.18). $\square$

4.4 Large Deviation Estimates

For $u \in M_1(E)$, $u(\cdot) \in D([0, T], M_1(E))$ with $\langle \varphi, u \rangle < \infty$ and $\sup_{0 \leq t \leq T} \langle \varphi, u(t) \rangle < \infty$, we have

Theorem 4.4.1 $(\mathcal{M}, \mathcal{P}_{u(\cdot), u}^{(N)}, N)$ is a large deviation system with action functional $I_u^{u(\cdot)}(\cdot)$.

The proof follows immediately from Sanov's theorem (cf. section 2.1 of chapter 2).

From now on, we always assume

$$(4.1) \quad \alpha = \int_E e^{\varphi(x)} u(dx) < \infty.$$

Lemma 4.4.2 For any $r > 0$, there exists an $R_0 > 0$ such that for all $R \geq R_0$ and $N \geq 1$, we have

$$(4.2) \quad \sup_{\nu \in \mathcal{C}} \mathcal{P}_\nu^{(N)} \{\mathcal{M}_\infty \setminus \mathcal{M}_R\} \leq 2\alpha^N \exp(-Nr).$$

where $\mathcal{C} = \{\nu : \nu \in M_1(E), \int_E e^{\varphi(x)} \nu(dx) \leq \alpha\}$.

Proof: For every $\nu \in \mathcal{C}$, we have

$$\begin{aligned} H &= \mathcal{P}_\nu^{(N)} \{\mathcal{M}_\infty \setminus \mathcal{M}_R\} = P_\nu^{(N)} \{(\eta_{x^{(N)}(\cdot)} \in \mathcal{M}_\infty \setminus \mathcal{M}_R)\} \\ &= P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} N^{-1} \sum_{k=1}^N \varphi(x_k^{(N)}(t)) > R \right\} \\ &= P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} \exp \left\{ \sum_{k=1}^N \varphi(x_k^{(N)}(t)) \right\} > \exp(NR) \right\}. \end{aligned}$$

Let $f(x) = x$. $A_i(t) = \int_0^t \int_E \{\exp[f(y) - f(x_i^{(N)}(s-))] - 1\} Q_{\epsilon_{x_i^{(N)}(s-), s-}}(x_i^{(N)}(s-), dy) ds$.

$$A(t) = \sum_{i=1}^N A_i(t).$$

Then by (1.4), we have

$$(4.3) \quad |A(t)| = \left| \int_0^t \sum_{i=1}^N \sum_{y \in E} [\exp((y - x_i^{(N)}(s-))) - 1] q_{x_i^{(N)}(s-), y} ds + \int_0^t (\epsilon - 1) \sum_{i=1}^N x_i^{(N)}(s-) ds \right| \leq N\lambda(1)T,$$

where $\lambda(1)$ be the constant in (1.4). For any $t > 0$, we have

$$(4.4) \quad P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} \exp\left(\sum_{k=1}^N (x_k^{(N)}(t))\right) > \exp(Nl) \right\} \leq P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} \exp\left(\sum_{k=1}^N (x_k^{(N)}(t)) - A(t)\right) > \exp(Nl - N\lambda(1)T) \right\}.$$

Let $F(x^{(N)}(t)) = \sum_{k=1}^N x_k^{(N)}(t)$. By Ito's formula (cf. chapter 2), we have

$$(4.5) \quad \begin{aligned} Z_t &= \exp\left(\sum_{k=1}^N (x_k^{(N)}(t)) - A(t)\right) = \exp(F(x^{(N)}(t)) - A(t)) \\ &= \exp\{F(x^{(N)}(0))\} - \int_0^t Z_{s-} dA(s) \\ &\quad + \int_0^t \sum_{y \in E^{\otimes N}} Z_{s-} [\exp(F(y) - F(x^{(N)}(s-))) - 1] N(ds, dy; x^{(N)}) \\ &= \exp\{F(x^{(N)}(0))\} - \int_0^t Z_{s-} dA(s) \\ &\quad + \int_0^t \sum_{y \in E^{\otimes N}} Z_{s-} [\exp(F(y) - F(x^{(N)}(s-))) - 1] \tilde{N}(ds, dy; x^{(N)}) \\ &\quad + \int_0^t \sum_{i=1}^N \sum_{z \in E} Z_{s-} [\exp(f(z) - f(x_i^{(N)}(s-))) - 1] Q_{\epsilon_{x_i^{(N)}(s-), s-}}(x_i^{(N)}(s-), dz) ds \\ &= \exp\{F(x^{(N)}(0))\} - \int_0^t Z_{s-} dA(s) \end{aligned}$$

$$\begin{aligned}
 & + \int_0^t \sum_{y \in E^{\otimes N}} Z_{s-} [\exp(F(y) - F(x^{(N)}(s-))) - 1] \tilde{N}(ds, dy; x^{(N)}) \\
 & + \int_0^t \sum_{i=1}^N Z_{s-} dA_i(s) \\
 & = \exp\{F(x^{(N)}(0))\} + \int_0^t \sum_{y \in E^{\otimes N}} Z_{s-} [\exp(F(y) - F(x^{(N)}(s-))) - 1] \tilde{N}(ds, dy; x^{(N)})
 \end{aligned}$$

Here we use the fact that $\{x_i^{(N)}(t)\}_{i=1}^N$ have no common jumps $P_\nu^{(N)}$ -almost surely. Thus Z_t is a local $P_\nu^{(N)}$ -martingale. Since Z_t is non-negative, it is also a supermartingale.

By Doob's inequality for supermartingale and (4.4), we have

$$\begin{aligned}
 (4.6) \quad & P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} \exp\left(\sum_{k=1}^N (x_k^{(N)}(t))\right) > \exp(Nl) \right\} \\
 & \leq \exp(-Nl + N\lambda(1)T) E^{P_\nu^{(N)}} \left(\exp\left(\sum_{k=1}^N x_k^{(N)}(0)\right) \right)
 \end{aligned}$$

We now return to the original estimate. It's easy to see that

$$\begin{aligned}
 (4.7) \quad & P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} \exp\left(\sum_{k=1}^N \varphi(x_k^{(N)}(t))\right) > \exp(NR) \right\} \\
 & \leq P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} \exp\left(\sum_{k=1}^N (x_k^{(N)}(t))\right) > \exp(Nl) \right\} \\
 & + P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} \exp\left(\sum_{k=1}^N \varphi(x_k^{(N)}(t))\right) > \exp(NR), \right. \\
 & \quad \left. \sup_{0 \leq t \leq T} \exp\left(\sum_{k=1}^N (x_k^{(N)}(t))\right) \leq \exp(Nl) \right\} \\
 & \leq H_1 + H_2.
 \end{aligned}$$

(4.6) implies

$$H_1 \leq \exp(-Nl + N\lambda(1)T) \alpha^N.$$

For H_2 , we have

$$A_i^\varphi(t) = \int_0^t \int_{y \in E} [\exp(\varphi(y) - \varphi(x_i^{(N)}(s-))) - 1] Q_{t, x^{(N)}(s-)}(x_i^{(N)}(s-), dy) ds.$$

Define

$$\begin{aligned}
 (4.8) \quad A^\varepsilon(t) &= \sum_{i=1}^N A_i^\varepsilon(t) \\
 &= \sum_{i=1}^N \int_0^t \left[\sum_{y \in E} q_{x_i^{(N)}(s-), y} (\exp(\varphi(y) - \varphi(x_i^{(N)}(s-))) - 1) \right] \\
 &\quad + \left(\frac{1}{N} \sum_{k=1}^N x_i^{(N)}(s-) \right) \{ \exp[\varphi(x_i^{(N)}(s-) + 1) - \varphi(x_i^{(N)}(s-))] - 1 \} ds
 \end{aligned}$$

By condition (1.4), this implies

$$\begin{aligned}
 (4.9) \quad H_2 &= P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} \exp \left(\sum_{k=1}^N \varphi(x_k^{(N)}(t)) - A^\varepsilon(t) \right) > \exp(NR - \lambda(l)NT) \right\} \\
 &\quad \sup_{0 \leq t \leq T} \exp \left(\sum_{k=1}^N (x_k^{(N)}(t)) \right) \leq \exp(Nl) \} \\
 &\leq P_\nu^{(N)} \left\{ \sup_{0 \leq t \leq T} \exp \left(\sum_{k=1}^N \varphi(x_k^{(N)}(t)) - A^\varepsilon(t) \right) > \exp(NR - \lambda(l)NT) \right\}
 \end{aligned}$$

Similarly we can prove that $Z_t^\varepsilon = \exp(\sum_{k=1}^N \varphi(x_k^{(N)}(t)) - A^\varepsilon(t))$ is also a $P_\nu^{(N)}$ -supermartingale. Thus

$$(4.10) \quad H_2 \leq \exp(-NR + \lambda(l)NT) \alpha^N.$$

Combining (4.6), (4.7) and (4.10) we have

$$(4.11) \quad H \leq \alpha^N [\exp(-Nl + \lambda(1)NT) + \exp(-NR + \lambda(l)NT)].$$

Taking l, R_0 such that $l - \lambda(1)T > r, R_0 = l + \lambda(l)T - \lambda(1)T$. Then for any $N \geq 1, R \geq R_0$ we have

$$\begin{aligned}
 (4.12) \quad H &\leq 2\alpha^N \exp(-Nl + \lambda(1)NT) \\
 &\leq 2\alpha^N \exp(-rN). \quad \square
 \end{aligned}$$

Lemma 4.4.3 For every $P \in \mathcal{M}_\infty$ and any open neighborhood V of $\bar{P}$, we have

$$(4.13) \quad \liminf_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(V) \geq -I_u(\bar{P}).$$

Proof: If $I_u(\bar{P}) = \infty$, the result is obvious. Now we assume $I_u(\bar{P}) < \infty$. Let

$$\begin{aligned} K_T^{(N)}(\omega) &= \sum_{i=1}^N \int_{[0,T] \times E} \log q_{\pi(\bar{P})(s-)}^{\pi(\eta^{(N)})(s-)}(\omega_i(s-), y) N(ds, dy; \omega_i) \\ G_T^{(N)}(\omega) &= - \sum_{i=1}^N \int_0^T \left\{ \left(\frac{\sum_{j=1}^N \omega_j(s)}{N} \right) - \|\pi(\bar{P})(s)\| \right\} ds \\ &= - \sum_{i=1}^N \int_0^T \left\{ \|\pi(\eta^{(N)})(\cdot)(s)\| - \|\pi(\bar{P})(s)\| \right\} ds \\ L_{p',p}^{(N)}(\omega) &= \sum_{i=1}^N \int_{[0,T] \times E} \left[(q_{\pi(\bar{P})(s-)}^{\pi(\eta^{(N)})(s-)}(\omega_i(s-), y))^{-\frac{p'}{p}} - 1 \right] Q_{\pi(\bar{P})(s-)}(\omega_i(s-), dy) ds. \end{aligned}$$

where $p > 1, p' > 1, 1/p + 1/p' = 1$.

Take R large enough such that $\bar{P} \in \mathcal{M}_R$. Then for any $\delta > 0$, we can find a smaller open neighborhood W of $\bar{P} \in \mathcal{M}$ such that

$$(4.14) \quad |G_T^{(N)}(\omega)| \leq \delta N \quad \text{uniformly on } W \cap \mathcal{M}_R.$$

Noting that

$$\begin{aligned} (4.15) \quad & [(q_{\pi(\bar{P})(t)}^{\pi(\eta^{(N)})(t)}(x, y))^{-\frac{p'}{p}} - 1] Q_{\pi(\bar{P})(t)}(x, y) \\ &= \frac{(q_{x,x+1} + \|\pi(\bar{P})(t)\|)^{p'}}{(q_{x,x+1} + \|\pi(\eta^{(N)})(\cdot)(t)\|)^{p'-1}} \\ &\quad - \frac{(q_{x,x+1} + \|\pi(\bar{P})(t)\|)(q_{x,x+1} + \|\pi(\eta^{(N)})(\cdot)(t)\|)^{p'-1}}{(q_{x,x+1} + \|\pi(\eta^{(N)})(\cdot)(t)\|)^{p'-1}} \\ &= \frac{(q_{x,x+1} + \|\pi(\bar{P})(t)\|)^{p'}}{(q_{x,x+1} + \|\pi(\eta^{(N)})(\cdot)(t)\|)^{p'-1}} \\ &\quad - \frac{(q_{x,x+1} + \|\pi(\bar{P})(t)\|)}{(q_{x,x+1} + \|\pi(\eta^{(N)})(\cdot)(t)\|)^{p'-1}} \\ &\quad \times (q_{x,x+1} + \|\pi(\bar{P})(t)\| + \|\pi(\eta^{(N)})(\cdot)(t)\| - \|\pi(\bar{P})(t)\|)^{p'-1}. \end{aligned}$$

which is a continuous function of $\{\eta^{(N)}\}$ for almost all $t \in [0, T]$ on $\mathcal{M}_\infty$. Thus we conclude that

$$(4.16) \quad |L_T^{(N)}(\omega)| \leq N \delta r(p') \quad \text{uniformly on } W \cap \mathcal{M}_R.$$

where $r(p')$ is a constant associated with p' .

All these combined with theorem 3.3.2 of chapter 3 imply

$$\begin{aligned}
 (4.17) \quad \mathcal{P}_u^{(N)}(V) &\geq \mathcal{P}_u^{(N)}(W \cap \mathcal{M}_R) \\
 &\geq E^{P_{\pi(\rho), u}^{(N)}}[\exp(K_T^{(N)}(\omega) - G_T^{(N)}(\omega)); \eta^{(N)} \in W \cap \mathcal{M}_R] \\
 &\geq e^{(-Ns)} E^{P_{\pi(\rho), u}^{(N)}}[\exp(K_T^{(N)}(\omega)); \eta^{(N)} \in W \cap \mathcal{M}_R]
 \end{aligned}$$

Applying Hölder's inequality we get

$$\begin{aligned}
 (4.18) \quad &E^{P_{\pi(\rho), u}^{(N)}}[\exp(K_T^{(N)}(\omega)); \eta^{(N)} \in W \cap \mathcal{M}_R] \\
 &\geq E^{P_{\pi(\rho), u}^{(N)}}[\exp(-\frac{p'}{p}K_T^{(N)}(\omega) - L_T^{(N)}(\omega) + L_T^{(N)}(\omega))]^{-p/p'} \\
 &\quad \times [P_{\pi(\rho), u}^{(N)}(\eta^{(N)} \in W \cap \mathcal{M}_R)]^p
 \end{aligned}$$

Combining (4.17) and (4.18), we get

$$\begin{aligned}
 (4.19) \quad \mathcal{P}_u^{(N)}(V) &\geq \mathcal{P}_u^{(N)}(W \cap \mathcal{M}_R) \\
 &\geq e^{-Ns} e^{-Nr(p')\delta p/p'} [P_{\pi(\rho), u}^{(N)}(\eta^{(N)} \in W \cap \mathcal{M}_R)]^p \\
 &\quad \times E^{P_{\pi(\rho), u}^{(N)}}[\exp(-p'/p K_T^{(N)}(\omega) - L_T^{(N)}(\omega))]^{-p/p'}
 \end{aligned}$$

Let $Y_t = \exp(-p'/p K_t^{(N)}(x^{(N)}) - L_t^{(N)}(x^{(N)}))$. Since $\{x_i^{(N)}(t)\}_{i=1}^N$ have no common jumps $P_u^{(N)}$ -almost surely, by Ito's formula we have

$$\begin{aligned}
 Y_t &= 1 - \int_0^t Y_{s-} dL_t^{(N)}(x^{(N)}) \\
 &\quad + \sum_{i=1}^N \int_{[0, t] \times E} Y_{s-} [(q_{\pi(\rho)(s-)}^{\tau(\eta^{(N)})}(x_i^{(N)}(s-), y))^{-\frac{p'}{p}} - 1] \cdot \tilde{N}(ds, dy; x_i^{(N)}) \\
 &= 1 - \int_0^t Y_{s-} dL_t^{(N)}(x^{(N)}) \\
 &\quad + \sum_{i=1}^N \int_{[0, t] \times E} Y_{s-} [(q_{\pi(\rho)(s-)}^{\tau(\eta^{(N)})}(x_i^{(N)}(s-), y))^{-\frac{p'}{p}} - 1] \cdot \tilde{N}(ds, dy; x_i^{(N)}) \\
 &\quad + \sum_{i=1}^N \int_{[0, t] \times E} Y_{s-} [(q_{\pi(\rho)(s-)}^{\tau(\eta^{(N)})}(x_i^{(N)}(s-), y))^{-\frac{p'}{p}} - 1] Q_{\pi(\rho)(s-)}(x_i^{(N)}(s-), dy) ds \\
 &= 1 + \sum_{i=1}^N \int_{[0, t] \times E} Y_{s-} [(q_{\pi(\rho)(s-)}^{\tau(\eta^{(N)})}(x_i^{(N)}(s-), y))^{-\frac{p'}{p}} - 1] \cdot \tilde{N}(ds, dy; x_i^{(N)}).
 \end{aligned}$$

Thus $\exp(-p'/pK_t^{(N)}(x^{(N)}) - L_t^{(N)}(x^{(N)}))$ is a $P_{\pi(\bar{p}),u}^{(N)}$ -supermartingale. This implies

$$(4.20) \quad \frac{1}{N} \log \mathcal{P}_u^{(N)}(V) \geq -(1 + r(p')p/p')\delta + p/N \log P_{\pi(\bar{p}),u}^{(N)}(\eta^{(N)} \in W \cap \mathcal{M}_R).$$

Since $(\mathcal{M}, P_{\pi(\bar{p}),u}^{(N)}, N)$ is a large deviation system, we have

$$(4.21) \quad \liminf_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_{\pi(\bar{p}),u}^{(N)}(W) \geq -I_u(\bar{P}).$$

It's easy to see that $\mathcal{P}_{\pi(\bar{p}),u}^{(N)}$ is concentrated on $\mathcal{M}_\infty$ and the estimate (4.2) is also true for it. Combining (4.20) and (4.21), let $\delta \rightarrow 0$, then let $p \rightarrow 1$, we finally get (4.13) which is just the lower bound. $\square$

Now we turn to the proof of the upper bound and the compactness of the level sets of I_u .

The next lemma verifies the "exponential tightness" of $\{\mathcal{P}_u^{(N)}\}$.

Lemma 4.4.4 *For any $R > 0, C > 0$, there exists a compact set $K_R \subset \mathcal{M}_\infty$ such that*

$$(4.22) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(\mathcal{M}_R \setminus K_R) < -C.$$

Proof: For any $R > 0$, we have

$$(4.23) \quad |G_T^{(N)}(\omega)| \leq 2NRT \text{ on } \mathcal{M}_R.$$

and

$$(4.24) \quad |W_T^{(N)}(\omega)| \leq N(2R + R^2)T \text{ on } \mathcal{M}_R.$$

where

$$|W_T^{(N)}(\omega)| = \sum_{i=1}^N \int_{[0,T] \times E} [(q_{\pi(\bar{p})(s-)}^{\pi(\eta^{(N)}(s-))}(\omega_i(s-), y))^2 - 1] Q_{\pi(\bar{p})(s-)}(\omega_i(s-), y) ds.$$

By lemma 3.32 in [41], there exists a compact set K in $\mathcal{M}$ such that

$$(4.25) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_{\pi(\bar{p}),u}^{(N)}(\mathcal{M} \setminus K) < -2(C + 4RT + R^2T).$$

Let $K_R = K \cap \mathcal{M}_R$, then K_R is compact in $\mathcal{M}_R$ and so is in $\mathcal{M}_\infty$.

By Cameron-Martin-Girsanov formula and Schwartz's inequality, we have

$$\begin{aligned}
 (4.26) \quad \mathcal{P}_u^{(N)}(\mathcal{M}_R \setminus K_R) &= E_{\pi(\bar{P}), u}^{P^{(N)}}[\exp(H_T^{(N)}(\omega)); \eta^{(N)} \in \mathcal{M}_R \setminus K_R] \\
 &\leq \epsilon^{2NRT} E_{\pi(\bar{P}), u}^{P^{(N)}}[\exp(K_T^{(N)}(\omega)); \eta^{(N)} \in \mathcal{M}_R \setminus K_R] \\
 &\leq \epsilon^{(4R+R^2)NT} \mathcal{P}_{\pi(\bar{P}), u}^{(N)}(\mathcal{M}_R \setminus K_R)^{1/2} \\
 &\leq \epsilon^{(4R+R^2)NT} \mathcal{P}_{\pi(\bar{P}), u}^{(N)}(\mathcal{M} \setminus K)^{1/2}
 \end{aligned}$$

which implies (4.22). $\square$

Lemma 4.4.5 *For every $\bar{P} \in \mathcal{M}_\infty$ and $\gamma > 0$, there exists an open neighborhood V of $\bar{P}$ in $\mathcal{M}$ such that*

$$(4.27) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(V) \leq -I_u(\bar{P}) + \gamma.$$

Proof: Because of (4.2), it suffices to verify that for any $\gamma > 0$ and $R > 0$ with $\bar{P} \in \mathcal{M}_R$ there exists an open neighborhood V of $\bar{P}$ such that

$$(4.28) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(V \cap \mathcal{M}_R) \leq -I_u(\bar{P}) + \gamma.$$

To this end, we assume that $I_u(\bar{P}) < \infty$. The proof also works with standard modifications in the case $I_u(\bar{P}) = \infty$.

We next choose $p, p' > 1$ with $1/p + 1/p' = 1$ and $\delta > 0$ so that

$$(4.29) \quad 1/p'(I_u(\bar{P}) - \gamma/2) - (1 + 1/p)\delta T \geq I_u(\bar{P}) - \gamma.$$

Let $l_p^{(N)}(\omega) = L_{p,-1}^{(N)}(\omega)$.

As in the proof of lemma 4.4.3, we can choose V in $\mathcal{M}$ so small such that

$$(4.30) \quad |G_T^{(N)}(\omega)| \leq \delta NT \text{ on } V \cap \mathcal{M}_R.$$

$$(4.31) \quad |l_T^{(N)}(\omega)| \leq \delta NT \text{ on } V \cap \mathcal{M}_R.$$

for all $N \geq 1$ and

$$(4.32) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_{\pi(\bar{P}), u}^{(N)}(V \cap \mathcal{M}_R) \leq -I_u(\bar{P}) + \gamma/2.$$

That the assertion (4.32) holds follows from the fact that $(\mathcal{M}, \mathcal{P}_{\pi(\bar{P}),u}^{(N)}, N)$ is a large deviation system with action functional $I_u^{\pi(\bar{P})}(\cdot)$ and $I_u^{\pi(\bar{P})}(\bar{P}) = I_u(\bar{P})$. Applying (4.30), (4.31), (4.32) and Hölder's inequality we obtain for all N

$$\begin{aligned} \mathcal{P}_u^{(N)}(\mathcal{M}_R \cap V) &= E^{\mathcal{P}_{\pi(\bar{P}),u}^{(N)}}[\exp(H_T^{(N)}(\omega)); \eta^{(N)} \in \mathcal{M}_R \cap V] \\ &\leq e^{NT\delta} E^{\mathcal{P}_{\pi(\bar{P}),u}^{(N)}}[\exp(pK_T^{(N)}(\omega))]^{1/p} \mathcal{P}_{\pi(\bar{P}),u}^{(N)}(\eta^{(N)} \in \mathcal{M}_R \cap V)^{1/p'} \\ &\leq e^{(1+1/p)NT\delta} \mathcal{P}_{\pi(\bar{P}),u}^{(N)}(\mathcal{M}_R \cap V)^{1/p'} \\ &\quad \times E^{\mathcal{P}_{\pi(\bar{P}),u}^{(N)}}[\exp(pK_T^{(N)}(\omega) - U_{T,p}^{(N)}(\omega))]^{1/p}. \end{aligned}$$

Similarly we can prove that $\exp(pK_t^{(N)}(\omega) - U_{t,p}^{(N)}(\omega))$ is a $\mathcal{P}_{\pi(\bar{P}),u}^{(N)}$ -supermartingale. This implies

$$\mathcal{P}_u^{(N)}(\mathcal{M}_R \cap V) \leq e^{(1+1/p)NT\delta} \mathcal{P}_{\pi(\bar{P}),u}^{(N)}(\mathcal{M}_R \cap V)^{1/p'}.$$

Thus, applying (4.29), we have

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(V \cap \mathcal{M}_R) &\leq (1 + 1/p)\delta T + 1/p'(-I_u(\bar{P}) + \gamma/2) \\ &\leq -I_u(\bar{P}) + \gamma. \end{aligned}$$

which implies (4.28) and thus (4.27) is true. $\square$

Lemma 4.4.6 *Assume conditions (1.1)-(1.5) and (4.1) are satisfied. Then for any closed subset F of $\mathcal{M}_\infty$, we have*

$$(4.33) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(F) \leq - \inf_{\bar{P} \in F} I_u(\bar{P}).$$

Proof: We assume that $\inf_{\bar{P} \in F} I_u(\bar{P}) < \infty$. The case of $\inf_{\bar{P} \in F} I_u(\bar{P}) = \infty$ can be handled similarly.

For any $s > \inf_{\bar{P} \in F} I_u(\bar{P})$, applying lemma 4.4.2 and lemma 4.4.4, we can take R large enough such that

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(\mathcal{M}_\infty \setminus \mathcal{M}_R) < -s.$$

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(\mathcal{M}_R \setminus K_R) < -s.$$

Now fix γ arbitrarily. By lemma 4.4.5, we find for any $\bar{P} \in F \cap K_R$ an open neighborhood V of $\bar{P}$ such that

$$(4.34) \quad \begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(V \cap K_R) &\leq -I_u(\bar{P}) + \gamma \\ &\leq -\inf_{P \in F} I_u(P) + \gamma. \end{aligned}$$

Thus we have

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(F) &= \limsup_{N \rightarrow \infty} \frac{1}{N} \log [\mathcal{P}_u^{(N)}(F \cap \mathcal{M}_R^c) \\ &\quad + \mathcal{P}_u^{(N)}(F \cap K_R) + \mathcal{P}_u^{(N)}(F \cap (\mathcal{M}_R \setminus K_R))] \end{aligned}$$

Since $F \cap K_R$ can be covered by finite number of open set V in $\mathcal{M}$ satisfying (4.34), we have

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(F) \leq -\inf_{P \in F} I_u(P) + \gamma.$$

Since γ is arbitrary, we get the estimate (4.33). $\square$

In order to finish the proof of theorem 4.1, it suffices to verify that the level sets are compact. To this end, we have

Lemma 4.4.7 *Under assumptions (1.1)-(1.5) and (4.1), we have that for every $r > 0$, $\Phi_u(r) = \{P \in \mathcal{M}_\infty : I_u(P) \leq r\}$ is compact.*

Proof: Lemma 4.3.3 implies that for any $r > 0$, $\Phi_u(r)$ is a closed subset of $\mathcal{M}_\infty$. Applying lemma 4.4.2 and lemma 4.4.4, for any $r > 0$, we can find a compact subset K of $\mathcal{M}_\infty$ such that

$$(4.35) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(\mathcal{M}_\infty \setminus K) \leq -r.$$

For any $\bar{P} \notin K$, there exists an open neighborhood U of $\bar{P}$ such that $U \cap K = \emptyset$.

By lemma 4.4.3, we have

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(\mathcal{M}_\infty \setminus K) &\leq -r. \\ \liminf_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(U) &\geq -I_u(\bar{P}). \end{aligned}$$

Thus $I_u(\bar{P}) > r$. This means that $\bar{P} \notin \Phi_u(r)$ which implies the compactness of the level sets.

Combining lemma 4.4.3, lemma 4.4.6 and lemma 4.4.7, we get theorem 4.2.1. $\square$

Chapter 5

Large Deviations For Empirical Process

In this chapter we will study the large deviations for the empirical process

$$\varepsilon_{x^{(N)}(t)} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}(t)}, \quad t \geq 0.$$

as $N \rightarrow \infty$. We also assume that the totally stable and conservative Q - matrix satisfies the same conditions as the conditions (1.1)-(1.5) of chapter 4.

5.1 The Main Result

Let $M_1^w(E)$ denote the space of all probability measures on E with the usual weak topology. r denotes the vague metric defined by $\forall u, v \in M_1(E), r(u, v) = \sum_{n=0}^{\infty} 2^{-n}(|u(n) - v(n)|)$. $M_1^r(E)$ denotes the space $M_1(E)$ endowed with the vague topology. Let $\mathcal{D} = D([0, T], M_1^w(E))$.

For each $m \geq 1$, define $\mathcal{D}_m = \{\mu(\cdot) \in \mathcal{D}; \sup_{0 \leq t \leq T} \langle \nu(t), \varphi \rangle \leq m\}$ equipped with the Skorohod topology. $\mathcal{D}_\infty = \bigcup_{m \geq 1} \mathcal{D}_m$ is equipped with the "inductive topology". By definition, a set V is open in $\mathcal{D}_\infty$ if and only if $V \cap \mathcal{D}_m$ is open in $\mathcal{D}_m$ for each $m \geq 1$. Another useful property is that a function is continuous on $\mathcal{D}_\infty$ if and only if it is sequential continuous. $\forall \mu(\cdot), \nu(\cdot) \in \mathcal{D}_\infty$ let $r_{0T}(\mu(\cdot), \nu(\cdot)) = \sup_{0 \leq t \leq T} r(\mu(t), \nu(t))$. Then r_{0T} induces the uniform topology on $\mathcal{D}_\infty$. From the definition of r and the right continuity we can prove that for any fixed $\nu_0(\cdot) \in \mathcal{D}_\infty$, $F(\mu(\cdot)) = r_{0T}(\mu(\cdot), \nu_0(\cdot))$ is measurable with respect to the Borel σ -algebra of space $\mathcal{D}_\infty$ with the "inductive topology". This will be used in proving the upper bound. For a detailed discussion

on the relation of the measurability for different topology refer to chapter 6.

For every $x^{(N)}(\cdot) \in D([0, T], E^{\otimes N})$, define

$$\varepsilon_{x^{(N)}(\cdot)} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}(\cdot)}$$

It is proved in [9] that $\varepsilon_{x^{(N)}(\cdot)}$ converges weakly to a measure-valued function $u(\cdot)$ which equals to $\pi(P_u)(\cdot) = P_u \circ x(\cdot)^{-1}$, where P_u is the unique solution to the nonlinear martingale problem for Q with initial distribution u .

Let $P_u^{(N)}$ be the unique solution to the martingale problem for $\Omega^{(N)}$ with initial distribution $u^{\otimes N}$ and $\hat{P}_u^{(N)}(\cdot) = P_u^{(N)} \circ \varepsilon_{x^{(N)}(\cdot)}^{-1}$. Then we have $\hat{P}_u^{(N)} \in M_1(D([0, T], M_1))$ and $\hat{P}_u^{(N)} \longrightarrow \delta_{\pi(P_u)(\cdot)}$. In fact $\hat{P}_u^{(N)}$ is concentrated on space $\mathcal{D}_\infty$.

For fixed $T > 0$, let $C_b(E)$ denote the set of all bounded continuous functions on E . $C_0^{1,0}([0, T] \times E)$ denotes the set of all continuous functions on $[0, T] \times E$ with compact support and first order continuous derivative with respect to t .

$\forall u, \nu \in M_1(E)$, define

$$l_u(\nu, g) = \langle \nu, g \rangle - \log \langle u, e^g \rangle, g \in C_b(E).$$

$\forall \bar{\mu}, \mu(\cdot) \in \mathcal{D}_\infty, f \in C_0^{1,0}([0, T] \times E)$, define

$$\begin{aligned} J_{\bar{\mu}(\cdot)}(\mu(\cdot); f, T) &= \langle \mu(T), f(T) \rangle - \langle \mu(0), f(0) \rangle \\ &\quad - \int_0^T \langle \mu(s), \frac{\partial f(s)}{\partial s} + e^{-f(s)} Q_{\bar{\mu}(s)} e^{f(s)} \rangle ds, \\ J(\mu(\cdot); f, T) &= J_{\mu(\cdot)}(\mu(\cdot); f, T). \end{aligned}$$

$$\begin{aligned} (1.1) \quad S_u^{\mu(\cdot)}(\mu(\cdot)) &= \sup \{ l_u(\mu(0), g) + J_{\bar{\mu}(\cdot)}(\mu(\cdot); f, T); \\ &\quad f \in C_0^{1,0}([0, T] \times E), g \in C_b(E) \}, \end{aligned}$$

$$(1.2) \quad S_u(\mu(\cdot)) = S_u^{\mu(\cdot)}(\mu(\cdot)).$$

It's not hard to show that $\forall g \in C_b(E), f \in C_0^{1,0}([0, T] \times E), l_u(\mu(0), g) + J(\mu(\cdot); f, T)$ is continuous on space $\mathcal{D}_\infty$.

The main result of this paper is:

Theorem 5.1.1 *Let u satisfy*

$$(1.3) \quad \eta = \int_E e^{\varphi(x)} u(dx) < \infty.$$

Then under assumptions (1.1)-(1.5) of chapter 4, $(\mathcal{D}_\infty, \hat{\mathcal{P}}_u^{(N)}, N)$ is a large deviation system with action functional S_u .

The following lemma is exactly lemma 4.4.2 of chapter 4 and will be frequently used in the sequel.

Lemma 5.1.2 *Assume conditions (1.1)-(1.5) of chapter 4 and (1.3) are satisfied. Then for any $r > 0$, there exists an $R_0 > 0$ such that for all $R \geq R_0$ and $N \geq 1$, we have*

$$(1.4) \quad \mathcal{P}_u^{(N)}\{\mathcal{D}_\infty \setminus \mathcal{D}_R\} \leq \exp(-Nr).$$

5.2 Lower Bound

In this section we'll get the lower bound through a series of lemmas. The method we use is similar to that used in [42]. But because of the unboundedness of jumps some new methods and techniques are needed.

The next lemma is a very standard result. We omit the proof which can be found in any book about measure theory.

Lemma 5.2.1 *Let $\{u_n\}$ be any sequence in $M_\varphi(E)$ and $u \in M_\varphi(E)$. If $u_n \rightarrow u$ vague topology and $\sup_n \langle u_n, \varphi \rangle < \infty$, then $|\langle u_n, x \rangle - \langle u, x \rangle| \rightarrow 0$. $M_\varphi(E) = \{\nu \in M_1(E); \langle \nu, \varphi(x) \rangle < \infty\}$.*

This lemma also implies that the vague topology and the weak topology are the same on $\{u; \langle u, \varphi \rangle \leq m\}$ for any m . Thus they will generate the same inductive topology on $\mathcal{D}_\infty$.

For $\bar{\mu}(\cdot) \in \mathcal{D}_\infty$, let $P_{\bar{\mu}(\cdot), (s, x)}$ be the unique solution of the martingale problem for $\{Q_{\bar{\mu}(t)}; t \in [0, T]\}$ with initial distribution δ_x at time s . For $0 \leq s < t \leq T, u, \nu \in M_\varphi(E), f \in C_b(E)$, we define

$$I_{\bar{\mu}(\cdot)}^{s, t} : M_1(E) \times M_1(E) \times C_b(E) \longrightarrow R^1,$$

$$I_{\bar{\mu}(\cdot)}^{s, t}(u, \nu; f) = \langle \nu, f(x(t)) \rangle - \langle u, \log E^{P_{\bar{\mu}(\cdot), (s, x)}}[e^{f(x(t))}] \rangle.$$

For $0 = t_0 < t_1 < \cdots < t_M = T$, we denote by $\pi^{(N)}(t_0, t_1, \dots, t_M)$ the joint distribution of $(\varepsilon_{x^{(N)}(t_0)}, \dots, \varepsilon_{x^{(N)}(t_M)})$ on $M_1(E)^{\otimes(M+1)}$ under $P_{\bar{\mu}(\cdot), u}^{(N)}$, $p(t_0, t_1, \dots, t_M)$ the joint distribution of $(x(t_0), \dots, x(t_M))$ on $E^{\otimes(M+1)}$ under $P_{\bar{\mu}(\cdot), u}$. Then we have

Lemma 5.2.2 *If $M_1(E)$ is endowed with the weak topology, then for any $M \geq 1, 0 = t_0 < t_1 < \cdots < t_M = T, (M_1(E)^{\otimes(M+1)}, \pi^{(N)}(t_0, t_1, \dots, t_M), N)$ is a large deviation system with action functional $L_{\bar{\mu}(\cdot), u}^{t_0, \dots, t_M}$ defined by*

$$\begin{aligned}
 (2.1) \quad L_{\bar{\mu}(\cdot), u}^{t_0, \dots, t_M} &= \sup_{g \in C_b(E)} l_u(\mu_0; g) + \sum_{k=0}^{M-1} \sup_{f \in C_k(E)} l_{\bar{\mu}(\cdot)}^{t_k, t_{k+1}}(\mu_k, \mu_{k+1}; f) \\
 &= \sup_{g \in C_b(E)} l_u(\mu_0; g) + \sum_{k=0}^{M-1} \sup_{f \in C_k(E)} l_{\bar{\mu}(\cdot)}^{t_k, t_{k+1}}(\mu_k, \mu_{k+1}; f).
 \end{aligned}$$

The second equality holds because $C_k(E)$ is pointwise dense in $C_b(E)$.

Proof: Theorem 3.5 in [6] implies that $(M_1(E)^{\otimes(M+1)}, \pi^{(N)}(t_0, t_1, \dots, t_M), N)$ is a large deviation system with action functional

$$\begin{aligned}
 \tilde{L}(\mu_0, \mu_1, \dots, \mu_M) &= \sup_{f_0, \dots, f_M \in C_k(E)} \left[\sum_{i=0}^M \langle \mu_i, f_i \rangle - H(f_0, \dots, f_M) \right]
 \end{aligned}$$

where

$$\begin{aligned}
 &\mu_0, \dots, \mu_M \in M_1(E), \\
 &H(f_0, \dots, f_M) \\
 &= \log \int_{E^{\otimes(M+1)}} p(t_0, t_1, \dots, t_M)(dx_0, \dots, dx_M) \exp\left(\sum_{i=1}^M f_i(x_i)\right).
 \end{aligned}$$

By the Markov property of $P_{\bar{\mu}(\cdot), u}$ we can rewrite $H(f_0, \dots, f_M)$ as

$$\log E^{P_{\bar{\mu}(\cdot), u}}[\exp\{f_0(x(t_0)) + h(f_1, \dots, f_M)(x(t_0))\}],$$

where

$$h(f_1, \dots, f_M)(x(t_0)) = \log E^{P_{\bar{\mu}(\cdot), u}(t_0, x(t_0))}[\exp\{\sum_{i=1}^M f_i(x(t_i))\}].$$

Since $h(f_1, \dots, f_M) \in C_b(E)$ we conclude that

$$\begin{aligned}
 & \tilde{L}(\mu_0, \mu_1, \dots, \mu_M) \\
 &= \sup_{f_0, \dots, f_M \in C_b(E)} [l_u(\mu_0; f_0 + h(f_1, \dots, f_M)) \\
 &+ \sum_{i=1}^M \langle \mu_i, f_i \rangle - \langle \mu_0, h(f_1, \dots, f_M) \rangle] \\
 &= \sup_{g \in C_b(E)} l_u(\mu_0; g) \\
 &+ \sup_{f_1, \dots, f_M \in C_b(E)} \sum_{i=1}^M [\langle \mu_i, f_i \rangle - \langle \mu_0, h(f_1, \dots, f_M) \rangle].
 \end{aligned}$$

By induction we get the result. $\square$

Let $\hat{\mathcal{P}}_{\tilde{\mu}(\cdot), u}^{(N)}(\cdot) = P_{\tilde{\mu}(\cdot), u}^{(N)} \circ \varepsilon_{x^{(N)}(\cdot)}^{-1}$. We have

Lemma 5.2.3 *For every $\gamma > 0$, $\tilde{\mu}(\cdot) \in \mathcal{D}_\infty \cap C([0, T], M_1(E))$ and open neighborhood V of $\tilde{\mu}(\cdot)$ in $\mathcal{D}_\infty$ with the inductive topology, there exist finitely many $0 = t_0 < t_1 < \dots < t_M = T$ and open neighborhood $\tilde{V}_i$ of $\tilde{\mu}(t_i)$, $i = 0, \dots, M-1$ in $M_1(E)$ with vague topology such that*

$$(2.2) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_{\tilde{\mu}(\cdot), u}^{(N)}(\{\varepsilon \notin V\} \cap \bigcap_{i=0}^{M-1} \{\varepsilon(t_i) \in \tilde{V}_i\}) \leq -\gamma.$$

Proof: Lemma 5.1.2 implies that $\forall \gamma > 0, \exists m > 0$, such that

$$(2.3) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_{\tilde{\mu}(\cdot), u}^{(N)}(\mathcal{D}_m^c) \leq -\gamma.$$

The continuity of $\tilde{\mu}(\cdot)$ implies the uniform continuity of it. Thus

$$(2.4) \quad \exists \varepsilon, \delta > 0 \ni \sup_{0 \leq s < t \leq T, t-s < \delta} r(\tilde{\mu}(t), \tilde{\mu}(s)) < \frac{\varepsilon}{4}.$$

$$(2.5) \quad \{\mu(\cdot) \in \mathcal{D}_m : \sup_{t \in [0, T]} r(\mu(t), \tilde{\mu}(t)) < \varepsilon\} \subset V.$$

If we take $\tilde{V}_i = \{\nu; r(\nu, \tilde{\mu}(t_i)) < \frac{\varepsilon}{4}\}$, then we have

$$\begin{aligned}
 (2.6) \quad & P_{\tilde{\mu}(\cdot), u}^{(N)}(\{\varepsilon \notin V\} \cap \bigcap_{i=0}^{M-1} \{\varepsilon(t_i) \in \tilde{V}_i\}) \\
 & \leq P_{\tilde{\mu}(\cdot), u}^{(N)}(\mathcal{D}_m^c) + \sum_{i=1}^{M-1} P_{\tilde{\mu}(\cdot), u}^{(N)}\left(\sup_{t \in [t_i, t_{i+1})} r(\varepsilon(t), \varepsilon(t_i)) \geq \frac{\varepsilon}{2}\right).
 \end{aligned}$$

By definition we have

$$\begin{aligned}
 (2.7) \quad & P_{\bar{\mu}(\cdot), u}^{(N)} \left\{ \sup_{t \in [t_i, t_{i+1})} r(\varepsilon(t), \varepsilon(t_i)) \geq \frac{\varepsilon}{2} \right\} \\
 & \leq P_{\bar{\mu}(\cdot), u}^{(N)} \left\{ \sup_{t \in [t_i, t_{i+1})} \sum_{k=0}^{m_\varepsilon} \frac{1}{N} \left| \sum_{j=1}^N (I_k(x_j^{(N)}(t_i)) - I_k(x_j^{(N)}(t))) \right| \geq \frac{\varepsilon}{4} \right\} \\
 & \leq \sum_{k=0}^{m_\varepsilon} P_{\bar{\mu}(\cdot), u}^{(N)} \left\{ \sup_{t \in [t_i, t_{i+1})} \frac{1}{N} \left| \sum_{j=1}^N (I_k(x_j^{(N)}(t_i)) - I_k(x_j^{(N)}(t))) \right| \geq \frac{\varepsilon}{4m_\varepsilon} \right\}.
 \end{aligned}$$

and

$$\begin{aligned}
 & P_{\bar{\mu}(\cdot), u}^{(N)} \left\{ \sup_{t \in [t_i, t_{i+1})} |(I_k(x_1^{(N)}(t_i)) - I_k(x_1^{(N)}(t)))| \geq \frac{\varepsilon}{4m_\varepsilon} \right\} \\
 & \leq P_{\bar{\mu}(\cdot), u}^{(N)} \left\{ \sup_{t \in [t_i, t_{i+1})} |(I_k(x_1^{(N)}(t_i)) - I_k(x_1^{(N)}(t)))| \geq \frac{\varepsilon}{4m_\varepsilon}, \right. \\
 & \quad \left. \exists t \in [t_i, t_{i+1}], x_1^{(N)}(t) \neq x_1^{(N)}(t_i) = k \right\} \\
 & + P_{\bar{\mu}(\cdot), u}^{(N)} \left\{ \sup_{t \in [t_i, t_{i+1})} |(I_k(x_1^{(N)}(t_i)) - I_k(x_1^{(N)}(t)))| \geq \frac{\varepsilon}{4m_\varepsilon}, \right. \\
 & \quad \left. \exists t \in [t_i, t_{i+1}], x_1^{(N)}(t_i) \neq x_1^{(N)}(t) = k \right\} \\
 & \leq 2P_{\bar{\mu}(\cdot), u}^{(N)} \{N([t_i, t_{i+1}], \{0, 1, \dots, m_\varepsilon + \Lambda\}; x_1^{(N)}) \geq 1\} \\
 & \leq 2E^{P_{\bar{\mu}(\cdot), u}^{(N)}} \{N([t_i, t_{i+1}], \{0, 1, \dots, m_\varepsilon + \Lambda\}; x_1^{(N)})\} \\
 & \leq 2E^{P_{\bar{\mu}(\cdot), u}^{(N)}} \left\{ \int_{t_i}^{t_{i+1}} Q_{\bar{\mu}(s)}(x_1^{(N)}(\cdot); \{0, 1, \dots, m_\varepsilon + \Lambda\}) ds \right\} \\
 & \leq q_{m_\varepsilon}(t_{i+1} - t_i).
 \end{aligned}$$

where m_ε is chosen to be an integer such that $\sum_{n=m_\varepsilon+1}^{\infty} 2^{-n} < \frac{\varepsilon}{2}$ and

$$q_{m_\varepsilon} = \sum_{x=0}^{m_\varepsilon+\Lambda} \sum_{y \in E} q_{x,y} + \sup_{0 \leq t \leq T} \{ \|\bar{\mu}(t)\| \},$$

$$N([t', t], A; \omega) = \# \{s : \omega(s) \in A, \omega(s) \neq \omega(s-), s \in [t', t]\}.$$

This implies

$$P_{\bar{\mu}(\cdot), u}^{(N)} \left\{ \sup_{t \in [t_i, t_{i+1})} \frac{1}{N} \left| \sum_{j=1}^N (I_k(x_j^{(N)}(t_i)) - I_k(x_j^{(N)}(t))) \right| \geq \frac{\varepsilon}{4m_\varepsilon} \right\}$$

$$\begin{aligned}
&\leq \sum_{n=N_\varepsilon}^N \binom{N}{n} (2r_{\mu(\cdot),u}^{(N)} \{N([t_i, t_{i+1}]), \{0, 1, \dots, m_\varepsilon + \Lambda\}; x_1^{(N)}\} \geq 1\}) \\
&\leq 4^N [q_{m_\varepsilon}(t_{i+1} - t_i)]^{N_\varepsilon},
\end{aligned}$$

where $N_\varepsilon = \lfloor \frac{N\varepsilon}{4m_\varepsilon} \rfloor - 1$.

Thus we have

$$(2.8) \quad \sum_{i=1}^{M-1} P_{\mu(\cdot),u}^{(N)} \left(\sup_{t \in [t_i, t_{i+1}]} r(\varepsilon(t), \varepsilon(t_i)) \geq \frac{\varepsilon}{2} \right) \leq M m_\varepsilon 4^N [q_{m_\varepsilon} \max(t_{i+1} - t_i)]^{N_\varepsilon}.$$

If we take

$$(2.9) \quad \max_i \{t_{i+1} - t_i\} \leq \left(\exp \left[\frac{4m_\varepsilon(-\gamma - \log 4)}{\varepsilon} - \log q_{m_\varepsilon} \right] \right) \wedge \delta.$$

then we obtain the inequality. $\square$

Lemma 5.2.4 *If $\mu(\cdot) \in \mathcal{D}_\infty$, $S_u^{\mu(\cdot)}(\mu(\cdot)) < \infty$, then there exists an integrable function $h_{0,T}$ such that*

$$(2.10) \quad \langle \mu(t), g \rangle - \langle \mu(\cdot), g \rangle = \int_s^t h_{0,T}(r) dr + \int_s^t \langle \mu(r), Q_{\tilde{\mu}(r)} g(r) \rangle dr.$$

This also implies that $\mu(\cdot)$ is continuous with respect to t and $\forall g \in C_k(E)$, $\langle \mu(\cdot), g \rangle$ is absolutely continuous.

Proof: Because $\mu(\cdot) \in \mathcal{D}_\infty$, to prove the continuity of $\mu(\cdot)$ it suffices to prove that the map $t \longrightarrow \langle \mu(t), g \rangle$ is continuous $\forall g \in C_k(E)$.

For any continuity points $0 \leq s < t \leq T$, we can find a sequence of smooth functions $\{h_n\}$ with $h_n \rightarrow \chi_{[s,t]}$ such that for any $f \in C_0^{1,0}([0, T] \times E)$

$$\begin{aligned}
J_{\tilde{\mu}(\cdot)}(\mu(\cdot); h_n f) &= \langle \mu(T), h_n f(T) \rangle - \langle \mu(0), h_n f(0) \rangle \\
&\quad - \int_0^T \langle \mu(r), \frac{\partial h_n f(r)}{\partial r} + e^{-h_n f(r)} Q_{\tilde{\mu}(r)} e^{h_n f(r)} \rangle dr \\
&= \langle \mu(T), h_n f(T) \rangle - \langle \mu(0), h_n f(0) \rangle \\
&\quad - \int_0^T \langle \mu(r), f(r) \frac{\partial h_n(r)}{\partial r} + h_n(r) \frac{\partial f(r)}{\partial r} + e^{-h_n f(r)} Q_{\tilde{\mu}(r)} e^{h_n f(r)} \rangle dr
\end{aligned}$$

$$\begin{aligned}
&= \langle \mu(T), h_n f(T) \rangle - \langle \mu(0), h_n f(0) \rangle \\
&- \int_0^T \langle \mu(r), f(r) \rangle dh_n(r) + \int_0^T \langle \mu(r), h_n(r) \frac{\partial f(r)}{\partial r} + e^{-h_n f(r)} Q_{\bar{\mu}(r)} e^{h_n f(r)} \rangle dr \\
&\longrightarrow \langle \mu(t), f(t) \rangle - \langle \mu(s), f(s) \rangle - \\
&\int_s^t \langle \mu(r), \frac{\partial f(r)}{\partial r} + e^{-f(r)} Q_{\bar{\mu}(r)} e^{f(r)} \rangle dr \text{ as } n \rightarrow 0.
\end{aligned}$$

Thus $S_{\bar{\mu}}^{\bar{\mu}(\cdot)} < \infty$ implies that there exists $R > 0$ such that for $0 \leq s < t \leq T$

$$\begin{aligned}
(2.11) \quad &\sup_{f \in C_0^{1,0}(E)} \{ \langle \mu(t), f(t) \rangle - \langle \mu(s), f(s) \rangle \\
&- \int_s^t \langle \mu(r), \frac{\partial f(r)}{\partial r} + e^{-f(r)} Q_{\bar{\mu}(r)} e^{f(r)} \rangle dr \} < R.
\end{aligned}$$

In general (2.11) is true for arbitrary $0 \leq s < t \leq T$ by the right continuity of $\bar{\mu}(\cdot)$. Fix $g \in C_k(E)$, $g \neq 0$, $0 \leq s < t \leq T$ arbitrarily. For every $h \in C^1([s, t])$, define

$$\begin{aligned}
I^{s,t}(h) &= h(t) \langle \mu(t), g \rangle - h(s) \langle \mu(s), g \rangle \\
&- \int_s^t \frac{\partial h(r)}{\partial r} \langle \mu(r), g \rangle dr - \int_s^t h(r) \langle \mu(r), Q_{\bar{\mu}(r)} g \rangle dr.
\end{aligned}$$

Let $\|h\| = \sup_{r \in [s, t]} |h(r)|$, then for any h with $\|h\| \leq 1$ we have

$$\begin{aligned}
|I^{s,t}(h)| &= \max\{I^{s,t}(h), I^{s,t}(-h)\} \\
&\leq \max\{I^{s,t}(h) - \int_s^t \langle \mu(r), e^{-g(r)h(r)} Q_{\bar{\mu}(r)} e^{g(r)h(r)} - Q_{\bar{\mu}(r)} h(r)g(r) \rangle dr \\
&+ \int_s^t \langle \mu(r), e^{-g(r)h(r)} Q_{\bar{\mu}(r)} e^{g(r)h(r)} - Q_{\bar{\mu}(r)} h(r)g(r) \rangle dr, \\
&I^{s,t}(-h) - \int_s^t \langle \mu(r), e^{-g(r)(-h(r))} Q_{\bar{\mu}(r)} e^{g(r)(-h(r))} - Q_{\bar{\mu}(r)} (-h(r))g(r) \rangle dr \\
&+ \int_s^t \langle \mu(r), e^{-g(r)(-h(r))} Q_{\bar{\mu}(r)} e^{g(r)(-h(r))} - Q_{\bar{\mu}(r)} (-h(r))g(r) \rangle dr \} \\
&\leq R + M_g(e^{2\|g\|} + 2\|g\| + 1).
\end{aligned}$$

Here $M_g = \sup_{r \in [s, t]} \langle \mu(r), x \rangle + \sum_{x \in su(g), x \neq y} q_{x,y} + \sum_{y \in su(g), x \neq y} q_{x,y} < \infty$ by the finite range property. Also $su(g)$ denotes the support of g . Since

$$|I^{s,t}(h)| = \|h\| I^{s,t}\left(\frac{h}{\|h\|}\right).$$

and $C^1([s, t])$ is dense in $C([s, t])$, we conclude that $I^{s,t}$ belongs to the dual space $C^*([s, t])$ of $C([s, t])$. By Riesz-Markov- Kakutani(cf. [36]), there exists a unique Baire measure $m_{s,t}$ (not necessarily non-negative) on $[s, t]$ such that $I^{s,t}(h) = \int_s^t h(r)m_{s,t}(dr)$. We also have from (2.11)

$$|\int_s^t h(r)m_{s,t}(dr)| \leq R + \int_s^t |\langle \mu(r), e^{-g(r)h(r)} Q_{\tilde{\mu}(r)} e^{g(r)h(r)} - Q_{\tilde{\mu}(r)} h(r)g(r) \rangle| dr.$$

This implies $\forall A \subset [s, t]$ with Lebesgue measure being zero, we have

$$|\int_A h(r)m_{s,t}(dr)| \leq R, \forall h \in C([s, t]).$$

This means that $m_{s,t}$ is absolutely continuous with respect to Lebesgue measure. Thus there exists an integrable function $h_{s,t}$ such that

$$I^{s,t}(h) = \int_s^t h(r)h_{s,t}(r)dr.$$

Since

$$\begin{aligned} I^{0,t}(h) &= I^{0,T}(h) - I^{t,T}(h) \\ &= \int_0^T h(r)h_{0,T}(r)dr - \int_t^T h(r)h_{t,T}(r)dr \\ &= \int_0^t h(r)h_{0,T}(r)dr + \int_t^T h(r)(h_{0,T}(r) - h_{t,T}(r))dr. \end{aligned}$$

and the lefthand side does not depend on the value of h on $[t, T]$, we get

$$h_{0,T}(r) = h_{0,t}(r), \text{ for } r \in [0, t].$$

Thus taking $h = \chi_{[s,t]}$ we finally get (2.10) which implies the continuity of $\mu(\cdot)$ and absolute continuity of $\langle \mu(\cdot), g \rangle \forall g \in C_k(E)$. $\square$

Note. If we adopt the definition of absolute continuity in section 4.1 of [6], then $\mu(t)$ in lemma 5.2.4 is actually absolutely continuous.

Lemma 5.2.5 For every $\hat{\mu}(\cdot) \in \mathcal{D}_\infty$ and open neighborhood V of $\hat{\mu}(\cdot)$ in $\mathcal{D}_\infty$,

$$(2.12) \quad \liminf_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(V) \geq -\hat{S}_u^{\hat{\mu}(\cdot)}(\hat{\mu}),$$

where

$$(2.13) \quad \hat{S}_u^{\hat{\mu}(\cdot)}(\mu(\cdot)) = \sup_{0=t_0 < \dots < t_M, M \geq 1} L_{\hat{\mu}(\cdot), u}^{t_0, \dots, t_M}(\mu(t_0), \dots, \mu(t_M))$$

Proof: For every open set $V \subset \mathcal{D}_\infty$, $m > 0$, we have $V \cap \mathcal{D}_m$ is an open subset of $\mathcal{D}_m$. If $\hat{S}_u^{\hat{\mu}(\cdot)}(\hat{\mu}(\cdot)) = \infty$, then (2.12) is obvious. Now we assume $c = \hat{S}_u^{\hat{\mu}(\cdot)}(\hat{\mu}(\cdot)) < \infty$. Then by lemma 5.2.4, $\hat{\mu}(\cdot)$ is continuous. Thus we have

$$\begin{aligned} \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(V) &\geq \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(V \cap \cap_{i=1}^{M-1} V_i) \\ &= \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(\cap_{i=1}^{M-1} V_i) - \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(V^c \cap \cap_{i=1}^{M-1} V_i) \\ &= \pi^{(N)}(t_0, \dots, t_M)(\hat{V}_0 \times \dots \times \hat{V}_M) - \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(V^c \cap \cap_{i=1}^{M-1} V_i), \end{aligned}$$

where $V_i = \{\mu(\cdot); \mu(t_i) \in \hat{V}_i\}$.

Since $\hat{V}_0 \times \dots \times \hat{V}_M$ is an open subset of $M_1(E)^{\otimes M+1}$ in the weak topology, if we take the γ in (2.2) be bigger than c , then combining (2.2) and lemma 5.2.2 we get

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(V) &\geq \limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(V \cap \cap_{i=1}^{M-1} V_i) \\ &= \max \left\{ \limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(V \cap \cap_{i=1}^{M-1} V_i), \limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_{\hat{\mu}(\cdot), u}^{(N)}(V^c \cap \cap_{i=1}^{M-1} V_i) \right\} \\ &\geq \limsup_{N \rightarrow \infty} \frac{1}{N} \log \pi^{(N)}(t_0, \dots, t_M)(\hat{V}_0 \times \dots \times \hat{V}_M) \geq -\hat{S}_u^{\hat{\mu}(\cdot)}(\hat{\mu}). \end{aligned}$$

which gives (2.12). $\square$

Lemma 5.2.6 For every $\bar{\mu}(\cdot), \mu(\cdot) \in \mathcal{D}_\infty$ we have

$$(2.14) \quad \hat{S}_u^{\bar{\mu}(\cdot)}(\mu(\cdot)) \leq S_u^{\bar{\mu}(\cdot)}(\mu(\cdot)).$$

Proof: $\forall n > 0$, $E_n = \{0, 1, \dots, n\}$, we define

$$(2.15) \quad q_{x,y}^{(n)}(t) = \begin{cases} q_{x,y}(t) & \text{for } y \neq x, x \in E_n \\ -\sum_{y \neq x} q_{x,y}^{(n)}(t) & \text{for } y = x \\ 0 & \text{otherwise} \end{cases}$$

Let $P_{\bar{\mu}(t), u}^n$ be the unique solution to the nonlinear martingale problem for $Q_{\bar{\mu}(t)}^{(n)}$ with initial distribution u . Then $P_{\bar{\mu}(t), u}^n$ converges weakly to $P_{\bar{\mu}(t), u}$, the unique solution to the nonlinear martingale problem for $Q_{\bar{\mu}(t)}$ with initial distribution u .

Let's define

$$h^{(n)}(s, x) = E^{P_{\bar{\mu}(s), \Lambda(s)}^{(n)}}[e^{f(x(t))} - 1],$$

$$g^{(n)}(s, x) = \log(1 + h^{(n)}(s, x)).$$

Then by the backward equation we get

$$(2.16) \quad \frac{\partial h^{(n)}(s, x)}{\partial s} + Q_{\bar{\mu}(s)}^{(n)} h^{(n)}(s, x) = 0,$$

$$h^{(n)}(t, x) = e^{f(x)} - 1.$$

This implies

$$(2.17) \quad \frac{\partial g^{(n)}(s, x)}{\partial s} + e^{-g^{(n)}(s, x)} Q_{\bar{\mu}(s)}^{(n)} e^{g^{(n)}(s, x)} = 0,$$

$$g^{(n)}(t, x) = f(x).$$

If the support of f is contained in E_n , then by definition the supports of $h^{(n)}$ and $g^{(n)}$ are both contained in E_n .

$\forall t_{i+1} > t_i$, let

$$\begin{aligned} J_{\bar{\mu}(\cdot)}^{t_i, t_{i+1}, n}(\mu(\cdot); g) &= \langle \mu(t_{i+1}), g(t_{i+1}) \rangle - \langle \mu(t_i), g(t_i) \rangle \\ &\quad - \int_{t_i}^{t_{i+1}} \langle \mu(r), \frac{\partial g(r)}{\partial r} + e^{-g(r)} Q_{\bar{\mu}(r)}^{(n)} e^{g(r)} \rangle dr \}, \\ J_{\bar{\mu}(\cdot)}^{t_i, t_{i+1}}(\mu(\cdot); g) &= \langle \mu(t_{i+1}), g(t_{i+1}) \rangle - \langle \mu(t_i), g(t_i) \rangle \\ &\quad - \int_{t_i}^{t_{i+1}} \langle \mu(r), \frac{\partial g(r)}{\partial r} + e^{-g(r)} Q_{\bar{\mu}(r)} e^{g(r)} \rangle dr \}, \\ l_{\bar{\mu}(\cdot)}^{t_i, t_{i+1}, n}(\mu(\cdot), f) &= \langle \mu(t_{i+1}), f(x) \rangle - \langle \mu(t_i), g^{(n)}(t_i, x) \rangle. \end{aligned}$$

Take n big enough such that the support of f is contained in E_n . If we take $f \leq 0$, then we have $g^{(n)} \leq 0$ and

$$Q_{\bar{\mu}(s)} g^{(n)}(s, x) - Q_{\bar{\mu}(s)}^{(n)} g^{(n)}(s, x) = \begin{cases} 0 & \text{for } x \leq n \text{ or } x > n + \Lambda \\ \sum_{y \in E_n} q_{x,y} g^{(n)}(s, y) & \text{for } x \in [n + 1, n + \Lambda] \end{cases}$$

Thus we have

$$Q_{\bar{\mu}(s)}^{(n)} g^{(n)}(s, x) \geq Q_{\bar{\mu}(s)} g^{(n)}(s, x)$$

$$l_{\bar{\mu}(\cdot)}^{t_i, t_{i+1}, n}(\mu(\cdot), f) \leq J_{\bar{\mu}(\cdot)}^{t_i, t_{i+1}}(\mu(\cdot); g^{(n)}).$$

Thus combining with (2.17) we get

$$(2.18) \quad l_{\bar{\mu}(\cdot)}^{t_i, t_{i+1}, n}(\mu(\cdot); f) \leq \sup_{g \in C_0^{1,0}([t_i, t_{i+1}] \times E)} J_{\bar{\mu}(\cdot)}^{t_i, t_{i+1}}(\mu(\cdot); g).$$

For $0 = t_0 < t_1 < \dots < t_M = T$, let

$$\varepsilon < \frac{1}{4} \min_{0 \leq i \leq M-1} \{|t_{i+1} - t_i|\}.$$

$$\sup_{g \in C_0^{1,0}([t_i, t_{i+1}] \times E)} J_{\bar{\mu}(\cdot)}^{t_i, t_{i+1}}(\mu(\cdot); g) \leq J_{\bar{\mu}(\cdot)}^{t_i, t_{i+1}}(\mu(\cdot); g_i) + \varepsilon/M.$$

$$\text{for some } g_i \in C_0^{1,0}([t_i, t_{i+1}] \times E).$$

Since we can take $\tilde{f} = f - \|f\|$, $\forall f \in C_k(E)$, then it can be seen that (2.18) is still true for general f . Then each g_i has an extension on $C_0^{1,0}([0, T] \times E)$ which will be also denoted by g_i . For $0 < i < M - 1$, we define

$$h_i^\varepsilon(t) = \begin{cases} 1 & \text{for } t \in [t_i, t_{i+1}] \\ 0 & \text{for } t \leq t_i - \varepsilon/2 \text{ or } t \geq t_{i+1} + \varepsilon/2 \\ \text{smooth} & \text{otherwise} \end{cases}$$

$$h_0^\varepsilon(t) = \begin{cases} 1 & \text{for } t \in [t_0, t_1] \\ 0 & \text{for } t \geq t_1 + \varepsilon/2 \\ \text{smooth} & \text{otherwise} \end{cases}$$

$$h_{M-1}^\varepsilon(t) = \begin{cases} 1 & \text{for } t \in [t_{M-1}, t_M] \\ 0 & \text{for } t \leq t_{M-1} - \varepsilon/2 \\ \text{smooth} & \text{otherwise} \end{cases}$$

Let

$$\tilde{g}^\varepsilon(t, x) = \sum_{i=0}^{M-1} h_i^\varepsilon(t) g_i(t, x).$$

Then we have

$$\begin{aligned} J_{\tilde{\mu}(\cdot)}(\mu(\cdot); \tilde{g}^\varepsilon) &= \langle \mu(T), \tilde{g}^\varepsilon(T) \rangle - \langle \mu(0), \tilde{g}^\varepsilon(0) \rangle \\ &\quad - \int_0^T \langle \mu(r), \frac{\partial \tilde{g}^\varepsilon(r)}{\partial r} + e^{-\tilde{g}^\varepsilon(r)} Q_{\tilde{\mu}(r)} e^{\tilde{g}^\varepsilon(r)} \rangle dr \\ &= \sum_{i=0}^{M-1} \{ \langle \mu(T), h_i^\varepsilon(T) g_i(T) \rangle - \langle \mu(0), h_i^\varepsilon(0) g_i(0) \rangle \} \\ &\quad - \int_0^T \langle \mu(r), \frac{\partial \tilde{g}^\varepsilon(r)}{\partial r} + e^{-\tilde{g}^\varepsilon(r)} Q_{\tilde{\mu}(r)} e^{\tilde{g}^\varepsilon(r)} \rangle dr \\ &= \sum_{i=0}^{M-1} \{ \langle \mu(T), h_i^\varepsilon(T) g_i(T) \rangle - \langle \mu(0), h_i^\varepsilon(0) g_i(0) \rangle \\ &\quad - \int_0^T \langle \mu(r), \frac{\partial h_i^\varepsilon(r) g_i(r)}{\partial r} + e^{-h_i^\varepsilon(r) g_i(r)} Q_{\tilde{\mu}(r)} e^{h_i^\varepsilon(r) g_i(r)} \rangle dr \} \\ &\quad + \int_0^T \{ \langle \mu(r), e^{-\tilde{g}^\varepsilon(r)} Q_{\tilde{\mu}(r)} e^{\tilde{g}^\varepsilon(r)} \rangle - \sum_{i=0}^{M-1} e^{-h_i^\varepsilon(r) g_i(r)} Q_{\tilde{\mu}(r)} e^{h_i^\varepsilon(r) g_i(r)} \rangle \} dr \\ &= \sum_{i=0}^{M-1} \{ J_{\tilde{\mu}(\cdot)}(\mu(\cdot); h_i^\varepsilon g_i) \} \\ &\quad + \int_0^T \{ \langle \mu(r), e^{-\tilde{g}^\varepsilon(r)} Q_{\tilde{\mu}(r)} e^{\tilde{g}^\varepsilon(r)} \rangle - \sum_{i=0}^{M-1} e^{-h_i^\varepsilon(r) g_i(r)} Q_{\tilde{\mu}(r)} e^{h_i^\varepsilon(r) g_i(r)} \rangle \} dr. \end{aligned}$$

If $\varepsilon \rightarrow 0$, then the second term will go to zero. Thus we get

$$J_{\tilde{\mu}(\cdot)}(\mu(\cdot); \tilde{g}^\varepsilon) \rightarrow \sum_{i=0}^{M-1} J_{\tilde{\mu}(\cdot)}^{t_i, t_{i+1}}(\mu(\cdot); g_i)$$

This implies that $\forall \eta > 0$, there exists an ε such that

$$\sum_{i=0}^{M-1} \sup_{g \in C_0^{1,0}([t_i, t_{i+1}] \times E)} \{ J_{\tilde{\mu}(\cdot)}^{t_i, t_{i+1}}(\mu(\cdot); g) \} \leq J_{\tilde{\mu}(\cdot)}(\mu(\cdot); \tilde{g}^\varepsilon) + \eta$$

$$\leq \sup_{g \in C_0^{1,0}([0,T] \times E)} \{J_{\bar{\mu}(\cdot)}(\mu(\cdot); g)\} + \eta.$$

Combining this with (2.1) and (2.18) we finally get (2.14). $\square$

Now we turn to prove the lower bound for the non-independent case. We have

Lemma 5.2.7 *For every $\bar{\mu}(\cdot) \in \mathcal{D}_\infty$ and any open neighborhood V of $\bar{\mu}(\cdot)$, we have*

$$(2.19) \quad \liminf_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_u^{(N)}(V) \geq -S_u(\bar{\mu}(\cdot)).$$

Proof: If $S_u(\bar{\mu}(\cdot)) = \infty$, the result is obvious. Now we assume $S_u(\bar{\mu}(\cdot)) < \infty$. Let

$$\begin{aligned} K_T^{(N)}(\omega) &= \sum_{i=1}^N \int_{[0,T] \times E} \log q_{\bar{\mu}(s-)}^{\varepsilon_{x^{(N)}(s-)}(\omega_i(s-); y)} N(ds, dy; \omega_i) \\ G_T^{(N)}(\omega) &= - \sum_{i=1}^N \int_0^T \left\{ \left(\frac{\sum_{j=1}^N \omega_j(s)}{N} \right) - \|\bar{\mu}(s)\| \right\} ds \\ &= - \sum_{i=1}^N \int_0^T \left\{ \|\varepsilon_{x^{(N)}(s)}\| - \|\bar{\mu}(s)\| \right\} ds \\ L_T^{(N)}(\omega) &= \sum_{i=1}^N \int_{[0,T] \times E} [(q_{\bar{\mu}(s-)}^{\varepsilon_{x^{(N)}(s-)}(\omega_i(s-); y)})^{-\frac{p'}{p}} - 1] Q_{\bar{\mu}(s-)}(\omega_i(s-), dy) ds. \end{aligned}$$

where $p > 1, p' > 1, 1/p + 1/p' = 1$.

Take m large enough such that $\bar{\mu} \in \mathcal{D}_m$. Then for any $\delta > 0$, we can find a smaller open neighborhood W of $\bar{\mu} \in \mathcal{D}_\infty$ such that

$$(2.20) \quad |G_T^{(N)}(\omega)| \leq \delta N \quad \text{uniformly on } W \cap \mathcal{D}_m.$$

Noting that

$$\begin{aligned} (2.21) \quad & [(q_{\bar{\mu}(t)}^{\varepsilon_{x^{(N)}(t)}(x, y)})^{-\frac{p'}{p}} - 1] Q_{\bar{\mu}(t)}(x, y) \\ &= \frac{(q_{x, x+1} + \|\bar{\mu}(t)\|)^{p'}}{(q_{x, x+1} + \|\varepsilon_{x^{(N)}(t)}\|)^{p'-1}} \\ &\quad - \frac{(q_{x, x+1} + \|\bar{\mu}(t)\|)(q_{x, x+1} + \|\varepsilon_{x^{(N)}(t)}\|)^{p'-1}}{(q_{x, x+1} + \|\varepsilon_{x^{(N)}(t)}\|)^{p'-1}} \\ &= \frac{(q_{x, x+1} + \|\bar{\mu}(t)\|)^{p'}}{(q_{x, x+1} + \|\varepsilon_{x^{(N)}(t)}\|)^{p'-1}} \end{aligned}$$

$$\begin{aligned}
& - \frac{(q_{x,x+1} + \|\bar{\mu}(t)\|)}{(q_{x,x+1} + \|\varepsilon_{x^{(N)}(t)}\|)^{p'-1}} \\
& \times (q_{x,x+1} + \|\bar{\mu}(t)\| + \|\varepsilon_{x^{(N)}(t)}\| - \|\bar{\mu}(t)\|)^{p'-1},
\end{aligned}$$

which is a continuous function of $\{\varepsilon_{x^{(N)}(t)}\}$ on $\mathcal{D}_\infty$. Thus we conclude that

$$(2.22) \quad |L_T^{(N)}(\omega)| \leq N\delta r(p') \text{ uniformly on } W \cap \mathcal{D}_m.$$

where $r(p')$ is a constant associated with p' .

All these combined with the Cameron-Martin-Girsanov formula imply

$$\begin{aligned}
(2.23) \quad \mathcal{P}_u^{(N)}(V) & \geq \mathcal{P}_u^{(N)}(W \cap \mathcal{D}_m) \\
& \geq E^{P_{\bar{\mu},u}^{(N)}}[\exp(K_T^{(N)}(\omega) - G_T^{(N)}(\omega)); \varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m] \\
& \geq e^{(-N\delta)} E^{P_{\bar{\mu},u}^{(N)}}[\exp(K_T^{(N)}(\omega)); \varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m]
\end{aligned}$$

Applying Hölder's inequality we get

$$\begin{aligned}
(2.24) \quad & E^{P_{\bar{\mu},u}^{(N)}}[\exp(K_T^{(N)}(\omega)); \varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m] \\
& \geq E^{P_{\bar{\mu},u}^{(N)}}[\exp(-\frac{p'}{p}K_T^{(N)}(\omega) - L_T^{(N)}(\omega) + L_T^{(N)}(\omega))]^{-p/p'} \\
& \quad \times [P_{\bar{\mu},u}^{(N)}(\varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m)]^p
\end{aligned}$$

Combining (2.23) and (2.24), we get

$$\begin{aligned}
(2.25) \quad \mathcal{P}_u^{(N)}(V) & \geq \mathcal{P}_u^{(N)}(W \cap \mathcal{D}_m) \\
& \geq e^{-N\delta} e^{-Nr(p')\delta p/p'} [P_{\bar{\mu},u}^{(N)}(\varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m)]^p \\
& \quad \times E^{P_{\bar{\mu},u}^{(N)}}[\exp(-p'/p K_T^{(N)}(\omega) - L_T^{(N)}(\omega))]^{-p/p'}
\end{aligned}$$

Let $Y_t = \exp(-p'/p K_t^{(N)}(x^{(N)}) - L_t^{(N)}(x^{(N)}))$. Since $\{x_i^{(N)}(t)\}_{i=1}^N$ have no common jumps $P_u^{(N)}$ -almost surely, by Ito's formula (cf. [29], p.66), we have

$$\begin{aligned}
Y_t &= 1 - \int_0^t Y_{s-} dL_s^{(N)}(x^{(N)}) \\
& \quad + \sum_{i=1}^N \int_{[0,t] \times E} Y_{s-} [(q_{\bar{\mu}(s-)}^{x_i^{(N)}(s-)}(x_i^{(N)}(s-), y))^{-\frac{p'}{p}} - 1] N(ds, dy; x_i^{(N)})
\end{aligned}$$

$$\begin{aligned}
&= 1 - \int_0^t Y_{s-} dL_t^{(N)}(x^{(N)}) \\
&\quad + \sum_{i=1}^N \int_{[0,t] \times E} Y_{s-} [(q_{\bar{\mu}(s-)}^{x^{(N)}(s-)}(x_i^{(N)}(s-), y))^{-\frac{p'}{p}} - 1] \tilde{N}(ds, dy; x_i^{(N)}) \\
&\quad + \sum_{i=1}^N \int_{[0,t] \times E} Y_{s-} [(q_{\bar{\mu}(s-)}^{x^{(N)}(s-)}(x_i^{(N)}(s-), y))^{-\frac{p'}{p}} - 1] Q_{\bar{\mu}(s-)}(x_i^{(N)}(s-), dy) ds \\
&= 1 + \sum_{i=1}^N \int_{[0,t] \times E} Y_{s-} [(q_{\bar{\mu}(s-)}^{x^{(N)}(s-)}(x_i^{(N)}(s-), y))^{-\frac{p'}{p}} - 1] \tilde{N}(ds, dy; x_i^{(N)}).
\end{aligned}$$

Thus $\exp(-p'/p K_t^{(N)}(x^{(N)}) - L_t^{(N)}(x^{(N)}))$ is a $P_{\bar{\mu}, u}^{(N)}$ -supermartingale. This implies

$$(2.26) \quad \frac{1}{N} \log P_u^{(N)}(V) \geq -(1 + r(p')p/p')\delta + p/N \log P_{\bar{\mu}, u}^{(N)}(\varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m).$$

By lemma 5.2.5 and lemma 5.2.6, we get

$$(2.27) \quad \frac{1}{N} \log P_{\bar{\mu}, u}^{(N)}(W) \geq -S_u(\bar{\mu}).$$

Let $\delta \rightarrow 0$, then let $p \rightarrow 1$, we finally get (2.19) which is just the lower bound.

□

The compactness of the level sets are proved in the next lemma.

Lemma 5.2.8 *Let us assume (1.1)-(1.5) of chapter 4 and (1.3). Then for any $\gamma > 0$, the level set $\Phi_u(\gamma) = \{\mu(\cdot) \in \mathcal{D}_\infty; S_u(\mu(\cdot)) \leq \gamma\}$ is compact in $\mathcal{D}_\infty$.*

Proof: Since $C([0, T], M_1(E)) \cap \mathcal{D}_\infty$ is a closed subset of $\mathcal{D}_\infty$, to prove the lemma it suffices to prove the compactness of $\Phi_u(\gamma)$ in $C([0, T], M_1(E)) \cap \mathcal{D}_\infty$. To do this it suffices to prove that

$$(2.28) \quad \sup_{\mu(\cdot) \in \Phi_u(\gamma)} \sup_{0 \leq t \leq T} \langle \varphi, \mu(t) \rangle < \infty.$$

and $\Phi_u(\gamma)$ is compact in $C([0, T], M_1(E))$ in the uniform topology.

For every $n \geq 1$, let $\varphi_n(x) = \varphi(x \wedge n)$. Then $\{\varphi_n\}$ is a sequence of bounded continuous functions on E . By the definition of the rate function, we conclude

$$(2.29) \quad \sup_{\mu(\cdot) \in \Phi_u(\gamma), n \geq 1} \{\langle \mu(0), \varphi_n \rangle - \log \langle u, e^{\varphi_n} \rangle\} \leq \gamma.$$

This combined with (1.3) implies

$$(2.30) \quad \sup_{\mu(\cdot) \in \Phi_u(\gamma), n \geq 1} \{ \langle \mu(0), \varphi_n \rangle \} \leq \eta + \gamma.$$

By Fatou's lemma we get

$$(2.31) \quad \sup_{\mu(\cdot) \in \Phi_u(\gamma)} \{ \langle \mu(0), \varphi \rangle \} \leq \eta + \gamma.$$

For every $n \geq 1$, let $f_n(x) = \varphi(x) \chi_{\{0,1,\dots,n\}}$. Then $\{f_n\}$ is in $C_k(E)$. By (1.4) of chapter 4, it's easy to check that

$$(2.32) \quad e^{-f_n(x)} Q_{\mu(t)} e^{f_n(x)} \leq \lambda(0) + \|\mu(t)\|(\varepsilon - 1).$$

Thus

$$(2.33) \quad \begin{aligned} \sup_{\mu(\cdot) \in \Phi_u(\lambda), n \geq 1} \{ J^{0,t}(\mu(\cdot); f_n) \} &= \sup_{\mu(\cdot) \in \Phi_u(\lambda), n \geq 1} \{ J_{\mu(\cdot)}^{0,t}(\mu(\cdot); f_n) \} \\ &= \sup_{\mu(\cdot) \in \Phi_u(\lambda), n \geq 1} \{ \langle \mu(t), f_n \rangle - \langle \mu(0), f_n \rangle \\ &\quad - \int_0^t \langle \mu(s), e^{-f_n(x)} Q_{\mu(s)} e^{f_n(x)} \rangle ds \} \leq \gamma. \end{aligned}$$

Combining (2.31)-(2.33) and Fatou's lemma, we get

$$(2.34) \quad \sup_{\mu(\cdot) \in \Phi_u(\lambda)} \{ \langle \mu(t), \varphi \rangle \} \leq \lambda(0) + \eta + \gamma + \int_0^t \sup_{\mu(\cdot) \in \Phi_u(\lambda)} \langle \mu(s), \varphi \rangle ds.$$

Let $R = (\lambda(0) + \eta + \gamma)e^T$. Then by Gronwall's lemma we conclude that

$$(2.35) \quad \Phi_u(\gamma) \subset \mathcal{D}_R.$$

By lemma 5.2.4, $\Phi_u(\gamma) \subset C([0, T], M_1(E))$. To finish the proof, it suffices to show the relative compactness of $\Phi_u(\gamma)$ in $C([0, T], M_1(E))$. By [26], it suffices to show the compactness of $\{ \langle \mu(\cdot), f \rangle; \mu(\cdot) \in \Phi_u(\gamma) \}$ in $C([0, T], \mathbb{R}^1) \quad \forall f \in C_k(E)$. Clearly, $\sup_{0 \leq t \leq T} | \langle \mu(t), f \rangle | < \infty, \forall f \in C_k(E)$. By Ascoli's theorem, it remains to prove the

equicontinuity of $\{t \mapsto \langle \mu(t), f \rangle\}$ for any f in $C_k(E)$. For any $\mu(\cdot) \in \mathcal{D}_\infty$, $f \in C_k(E)$, we define

$$(2.36) \quad D_{\delta, \rho}^f = \{\mu(\cdot) \in \mathcal{D}_\infty; \sup_{s, t \in [0, T], t-s < \delta} |\langle \mu(t), f \rangle - \langle \mu(s), f \rangle| \leq \rho\}, \rho > 0.$$

For fixed γ we can find an R by lemma 5.1.2 such that

$$(2.37) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{P}_u^{(N)}(\mathcal{D}_R^c) \leq -2\gamma.$$

Fix $R > 0$, $s > 0$ and $f \in C_k(E)$ arbitrarily. Applying the Markov property, we derive for any $\delta \in [0, T/2]$ and $\rho > 0$ the estimate (cf. [2], p.56)

$$\begin{aligned} (2.38) \quad & \hat{P}_u^{(N)}(\mathcal{D}_R \cap (D_{\delta, \rho}^f)^c) \\ & \leq \hat{P}_u^{(N)}(\sup_{s, t \in [0, T], t-s < \delta} |\langle \mu(t), f \rangle - \langle \mu(s), f \rangle| > \rho, \mu(\cdot) \in \mathcal{D}_R) \\ & \leq \sum_{i=0}^{\lfloor T/\delta \rfloor - 1} P_u^{(N)}(\sup_{1\delta \leq t \leq (i+2)\delta \wedge T} |\frac{1}{N} \sum_{i=1}^N (f(x_i^{(N)}(t)) - f(x_i^{(N)}(1\delta)))| > \rho/2, \\ & \quad \varepsilon_{X^{(N)}(\cdot)} \in \mathcal{D}_R) \\ & \leq \frac{T}{\delta} \sup_{\nu \in M_\nu(E)} P_\nu^{(N)}(\sup_{0 \leq t \leq 2\delta} |\frac{1}{N} \sum_{i=1}^N (f(x_i^{(N)}(t)) - f(x_i^{(N)}(0)))| > \rho/2, \\ & \quad \varepsilon_{X^{(N)}(\cdot)} \in \mathcal{D}_R) \end{aligned}$$

For any $f \in C_k(E)$, $\omega(\cdot) \in D([0, T], E)^{\otimes N}$, we set

$$(2.39) \quad K^N(f, t; \omega) = \sum_{i=1}^N \{f(\omega_i(t)) - f(\omega_i(0)) - \int_0^t e^{-f(\omega_i(s))} Q_{\varepsilon_{\omega(s)}} e^{f(\omega_i(s))} dr\}$$

Then for every $\beta > 0$, $e^{K^N(\beta f, t; \omega)}$ is a $P_u^{(N)}$ -martingale. Combining (2.38), (2.39) and Doob's inequality we derive the estimate

$$\begin{aligned} (2.40) \quad & P_u^{(N)}(\sup_{0 \leq t \leq 2\delta} |\frac{1}{N} \sum_{i=1}^N (f(x_i^{(N)}(t)) - f(x_i^{(N)}(0)))| > \rho/2, \varepsilon_{X^{(N)}(\cdot)} \in \mathcal{D}_R) \\ & \leq \sup_{\nu \in M_\nu(E)} P_\nu^{(N)}\{\sup_{0 \leq t \leq 2\delta} K^N(\beta f, t; \omega) > N(\beta\rho/2 - 2c(f, R)\delta(e^{2\beta\|f\|} - 1)), \\ & \quad \varepsilon_{X^{(N)}(\cdot)} \in \mathcal{D}_R\} \\ & \leq \exp[-N\{\beta\rho/2 - 2c(f, R)\delta(e^{2\beta\|f\|} - 1)\}]. \end{aligned}$$

where $c(f, R) = \sup\{\sum_{y \in E} q_{x,y}; x \in su(f)\} + R$. This implies

$$(2.41) \quad \hat{\mathcal{P}}_u^{(N)}(\mathcal{D}_R \cap (D_{\delta, \rho}^f)^c) \leq \frac{2T}{\delta} \exp[-N\{\beta\rho/2 - 2c(f, R)\delta(e^{2\beta\|f\|} - 1)\}].$$

Minimizing the expression on the right of (2.41) with respect to $\beta > 0$, we have

$$(2.42) \quad \hat{\mathcal{P}}_u^{(N)}(\mathcal{D}_R \cap (D_{\delta, \rho}^f)^c) \leq \frac{2T}{\delta} \exp[-N\{\frac{\rho}{4\|f\|}(\log(\frac{\rho}{8c(f, R)\|f\|\delta}) - 1) + 2c(f, R)\delta\}].$$

For the given γ , let $n_0 = \lfloor \frac{1}{4\gamma\|f\|} \rfloor$. If we choose

$$\rho_n = \frac{1}{n}, \delta_n = \frac{1}{8\|f\|c(f, R)n} e^{-12\|f\|\gamma n},$$

then we have

$$(2.43) \quad \begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_u^{(N)}(\mathcal{D}_R \cap (D_{\delta_n, \rho_n}^f)^c) \\ \leq \limsup_{N \rightarrow \infty} \frac{1}{N} \log \frac{2T}{\delta_n} \exp[-N\{\frac{\rho_n}{4\|f\|}(\log(\frac{\rho_n}{8c(f, R)\|f\|\delta_n}) - 1) + 2c(f, R)\delta_n\}] \\ \leq -3\gamma + \frac{1}{4\|f\|n} \leq -2\gamma, \forall n > n_0. \end{aligned}$$

Since D_{δ_n, ρ_n}^f is a closed subset of $\mathcal{D}_\infty$, we have

$$(2.44) \quad \begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_u^{(N)}\{(D_{\delta_n, \rho_n}^f)^c\} \\ \leq \max\{\limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_u^{(N)}\{\mathcal{D}_R^c\}, \limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_u^{(N)}\{\mathcal{D}_R \cap (D_{\delta_n, \rho_n}^f)^c\}\} \\ \leq -2\gamma. \end{aligned}$$

By (2.19), we have

$$(2.45) \quad \inf_{\mu(\cdot) \in (D_{\delta_n, \rho_n}^f)^c} S_u(\mu(\cdot)) > \gamma.$$

Thus for any $n > n_0$, we have

$$(2.46) \quad \Phi_u(\gamma) \subset D_{\delta_n, 1/n}^f.$$

This means that

$$(2.47) \quad \Phi_u(\gamma) \subset C([0, T], M_1(E)) \cap_{n > n_0} D_{\delta_n, 1/n}^f.$$

Thus we obtain the relative compactness of $\Phi_u(\gamma), \forall \gamma > 0$. Since I_u is lower semi-continuous we in fact obtain the compactness of $\Phi_u(\gamma), \forall \gamma > 0$. $\square$

5.3 Upper Bound

In this section we obtain the large deviation upper bound. The method we use is a generalization of the method used by F.Comets [4] in proving the large deviation principle for the Curie-Weiss model on the torus.

Let

$$H_u(\mu(\cdot); f, g, T) = I_u(\mu(0), g) + J(\mu(\cdot); f, T)$$

$$S_u(\mu(\cdot)) = \sup\{H_u(\mu(\cdot); f, g, T);$$

$$f \in C_0^{1,0}([0, T] \times E), g \in C_b(E)\},$$

$$F(t, x^{(N)}(t)) = \exp\left[\sum_{i=1}^N (f(t, x_i^{(N)}(t)) - f(0, x_i^{(N)}(0)) - B(t))\right].$$

$$B(t) = \sum_{i=1}^N \int_0^t \left(\frac{\partial f(s, x_i^{(N)}(s))}{\partial s} + e^{-f(s, x_i^{(N)}(s))} Q_{\varepsilon_{x_i^{(N)}(s)}} e^{f(s, x_i^{(N)}(s))} \right) ds.$$

By Ito's formula we have

$$\begin{aligned} & F(t, x^{(N)}(t)) - 1 \\ &= \int_0^t F'_s(s, x^{(N)}(s)) ds - \int_0^t F(s, x^{(N)}(s)) dB(s) \\ &+ \int_0^{t+} \int_E F(s-, x^{(N)}(s-)) \sum_{i=1}^N [e^{(f(s-, y) - f(s-, x_i^{(N)}(s-)))} - 1] \cdot N(ds, dy; x_i^{(N)}) \\ &= \int_0^t F(s, x^{(N)}(s)) \sum_{i=1}^N \frac{\partial f(s, x_i^{(N)}(s))}{\partial s} ds - \int_0^t F(s, x^{(N)}(s)) dB(s) \\ &+ \int_0^{t+} \int_E F(s-, x^{(N)}(s-)) \sum_{i=1}^N [e^{(f(s-, y) - f(s-, x_i^{(N)}(s-)))} - 1] \cdot \tilde{N}(ds, dy; x_i^{(N)}) \\ &+ \int_0^t \sum_{i=1}^N e^{-f(s-, x_i^{(N)}(s-))} Q_{\varepsilon_{x_i^{(N)}(s)}} e^{f(s-, x_i^{(N)}(s-))} ds \\ &= \int_0^{t+} \int_E F(s-, x^{(N)}(s-)) \sum_{i=1}^N [e^{(f(s-, y) - f(s-, x_i^{(N)}(s-)))} - 1] \tilde{N}(ds, dy; x_i^{(N)}). \end{aligned}$$

Thus $F(t, x^{(N)}(t))$ is a martingale with respect to $F_u^{(N)}$.

Let

$$Y_u(\mu(\cdot); f, g, T) = \exp[N H_u(\mu(\cdot); f, g, T)]$$

then we have

Lemma 5.3.1 $E^{\hat{P}_u^{(N)}}[Y_u(\mu(\cdot); f, g, T)] = E^{\hat{P}_u^{(N)}}[Y_u(\varepsilon_{x^{(N)}(\cdot)}; f, g, T)] = 1.$

Proof: By direct calculation we have $E^{\hat{P}_u^{(N)}} \exp[Nl_u(\varepsilon_{x^{(N)}(0)}; g)] = 1.$ This combined with the fact that $F(t, x^{(N)}(t))$ is a martingale with respect to $P_u^{(N)}$ we get the result. $\square$

The next theorem is crucial for obtaining the upper bound.

Theorem 5.3.2 Assume (1.1)-(1.5) of chapter 4 and (1.3) are satisfied. $\forall \gamma > 0, \varepsilon > 0, s \geq 0,$ there exists an integer N_0 such that $\forall N \geq N_0,$

$$(3.1) \quad \hat{P}_u^{(N)}\{r_{0T}(\mu(\cdot), \Phi_u(s)) \geq \varepsilon\} \leq \exp[-N(s - \gamma)].$$

where $r_{0T}(\mu(\cdot), \Phi_u(s)) = \inf_{\nu(\cdot) \in \Phi_u(s)} \sup_{0 \leq t \leq T} r(\mu(t), \nu(t)).$

In order to prove this theorem we need the following two results:

Lemma 5.3.3 For any $\bar{\mu}(\cdot) \in \mathcal{D}_\infty$ and $I < S_u(\bar{\mu}(\cdot))$ there exists $\delta' > 0$ and an integer N_1 such that $\forall N \geq N_1,$

$$(3.2) \quad \hat{P}_u^{(N)}\{r_{0T}(\mu(\cdot), \bar{\mu}(\cdot)) < \delta'\} \leq \exp[-NI].$$

Proof: For any $\bar{\mu}(\cdot) \in \mathcal{D}_\infty$ and $I < S_u(\bar{\mu}(\cdot))$, there exists $\gamma > 0$ such that $I + \gamma < S_u(\bar{\mu}(\cdot)).$ By lemma 5.1.2, there exists $R_0 \geq 1$ such that

$$(3.3) \quad \hat{P}_u^{(N)}\{(\mathcal{D}_{R_0})^c\} \leq \exp[-N(I + \gamma)].$$

For fixed $\bar{\mu}(\cdot) \in \mathcal{D}_\infty$, we have that

$$\exists f \in C_0^{1,0}([0, T] \times E), g \in C_b(E), \text{ such that}$$

$$S_u(\bar{\mu}(\cdot)) \leq H_u(\bar{\mu}(\cdot); f, g, T) + \gamma/3.$$

Choose $R \geq R_0$ such that $\bar{\mu}(\cdot) \in \mathcal{D}_R.$

Since $H_u(\mu(\cdot); f, g, T)$ is continuous at $\bar{\mu}(\cdot)$ on $\mathcal{D}_R$ in the uniform topology there exists an $\delta' > 0$ such that

$$H_u(\bar{\mu}(\cdot); f, g, T) \leq \inf_{\mu(\cdot) \in V} H_u(\mu(\cdot); f, g, T) + \gamma/3.$$

where $V = \{\mu(\cdot) \in \mathcal{D}_R; r_{0T}(\mu(\cdot), \bar{\mu}(\cdot)) < \delta'\}$.

By Chebyshev's inequality we get

$$\begin{aligned} \hat{\mathcal{P}}_u^{(N)}(V) &\leq \hat{\mathcal{P}}_u^{(N)}\{\mu(\cdot); Y_u(\mu(\cdot); f, g, T) \in Y_u(V; f, g, T)\} \\ &\leq \frac{\int_{\mathcal{D}_R} Y_u(\mu(\cdot); f, g, T) \hat{\mathcal{P}}_u^{(N)}(d\mu(\cdot))}{\inf_{\mu(\cdot) \in V} Y_u(\mu(\cdot); f, g, T)} \\ &= \frac{1}{\inf_{\mu(\cdot) \in V} Y_u(\mu(\cdot); f, g, T)}. \end{aligned}$$

Combining all these we obtain

$$\begin{aligned} (3.1) \quad \hat{\mathcal{P}}_u^{(N)}(V) &\leq \exp[-N \inf_{\mu(\cdot) \in V} H_u(\mu(\cdot); f, g, T)] \\ &\leq \exp[-N(H_u(\bar{\mu}(\cdot); f, g, T) - \gamma/3)] \\ &\leq \exp[-N(S_u(\bar{\mu}(\cdot)) - 2\gamma/3)] \\ &\leq \exp[-N(I + \gamma/3)]. \end{aligned}$$

This combined with (3.3) implies (3.2). $\square$

Lemma 5.3.4 For all $a \geq 0$, there exists an $R \geq 1$ and a compact subset $K_a \subset \mathcal{D}_R \cap C([0, T], M_1(E))$ such that the following is true: $\forall \delta > 0, \exists N_2 \geq 1$ such that $\forall N \geq N_2$

$$(3.5) \quad \hat{\mathcal{P}}_u^{(N)}\{r_{0T}(\mu(\cdot), K_a) \geq \delta\} \leq \exp[-Na].$$

Proof: If $a = 0$ the result is trivial. Suppose $a > 0$. By lemma 5.1.2, there exists $R \geq 1$ such that

$$(3.6) \quad \hat{\mathcal{P}}_u^{(N)}\{(\mathcal{D}_R)\} \leq \exp[-N(a+1)].$$

For fixed a and R we construct a sequence $\{\Delta_j\}_{j=1, \dots}$ satisfying:

$$T/\Delta_j \in \mathbb{N}, \quad \Delta_j/\Delta_{j+1} \in \mathbb{N} \setminus \{0, 1\},$$

$$\Delta_j \leq (a+1)/2\hat{c}(j, R)(\exp[16(a+1)j] - 1),$$

$$j\Delta_j \text{ is decreasing.}$$

where $\mathbf{N}$ denotes the set of integers, $\hat{c}(j, R)$ is defined later and is a positive increasing function of j .

Let K'_a be defined as

$$K'_a = \{\mu(\cdot) \in \mathcal{D}_R; \sup_{t,s \in [0,T], |t-s| \leq \Delta_j} r(\mu(t), \mu(s)) < \frac{1}{j}, j = 1, 2, \dots\}.$$

The Ascoli-Arzelà theorem proves that K'_a is a relatively compact subset of the space $\mathcal{D}_R$ in the uniform topology. Let K_a be the closure of K'_a in $\mathcal{D}_R$ in the uniform topology. Then K_a is a compact subset of $\mathcal{D}_R$ in both the uniform and Skorohod topology. For any $\delta > 0$ choose j_0 such that $\frac{1}{j_0} < \delta$. Let

$$D(j_0) = \{\mu(\cdot) \in \mathcal{D}_R; \sup_{t,s \in [0,T], |t-s| \leq \Delta_{j_0}} r(\mu(t), \mu(s)) < \frac{1}{j_0}\}.$$

Choose $n_0 \in \mathbf{N}$ such that $\sum_{n > n_0} \frac{1}{2^n} < \frac{1}{2j_0}$. For $k = 1, \dots, n_0$ we define $f_k(x) = \delta_k(x)$. Then f_k is continuous function on E with compact support. Similar to the estimates (2.38) and (2.41) we have for any $\beta > 0$

$$\begin{aligned} (3.7) \quad & \hat{P}_u^{(N)}(\mathcal{D}_R \cap D(j_0)^c) \\ & \leq \sum_{k=0}^{n_0} \sum_{l=0}^{[T/\Delta_{j_0}]-1} P_u^{(N)}\{\varepsilon_{X^{(N)}(\cdot)} \in \mathcal{D}_R, \\ & \sup_{l\Delta_{j_0} \leq t \leq ((l+2)\Delta_{j_0}) \wedge T} \left| \frac{1}{N} \sum_{i=1}^N (f_k(x_i^{(N)}(t)) - f_k(x_i^{(N)}(l\Delta_{j_0})) \right| \geq 1/4j_0\} \\ & \leq \frac{T}{\Delta_{j_0}} \sum_{k=0}^{n_0} \exp[-N(\beta/4j_0 - 2c(f_k, R)\Delta_{j_0}(\epsilon^{2\beta} - 1))] \\ & \leq \frac{T}{\Delta_{j_0}} n_0 \exp[-N(\frac{\beta}{4j_0} - 2\hat{c}(j_0, R)\Delta_{j_0}(\epsilon^{2\beta} - 1))] \end{aligned}$$

where $\hat{c}(j_0, R) = \sup_{k \leq j_0} c(\hat{f}_k, R)$. Let $\beta = 8j_0(a+1)$ then by the construction of Δ_j we conclude that

$$(3.8) \quad \hat{P}_u^{(N)}(\mathcal{D}_R \cap D(j_0)^c) \leq \frac{T}{\Delta_{j_0}} n_0 \epsilon^{[-N(a+1)]}.$$

Now let $m_0 = \frac{T}{\Delta_{j_0}}$, $t_k = k\Delta_{j_0}$, $k = 1, \dots, m_0$. For any $\mu(\cdot) \in \mathcal{D}_\infty$ we define $\mu^0(\cdot)$

as follow:

$$(3.9) \quad \mu^0(t) = \begin{cases} \mu(t_k) & \text{if } t = t_k, \text{ for } 1 \leq k \leq m_0 \\ \text{linear} & \text{otherwise} \end{cases}$$

Obviously $r_{0T}(\mu(\cdot), \mu^0(\cdot)) < \delta$ on $D(j_0)$. Thus

$$(3.10) \quad \hat{\mathcal{P}}_u^{(N)}\{\mu(\cdot) \in D(j_0); r_{0T}(\mu(\cdot), K_a) \geq \delta\} \leq \hat{\mathcal{P}}_u^{(N)}\{D(j_0) \cap \{\mu^0(\cdot) \notin K'_a\}\}.$$

Using an argument similar to that in F.Comets [4] on page 28-29 we can verify that on $D(j_0)$, the slope of $\mu^0(\cdot)$ satisfies

$$\begin{aligned} \frac{r(\mu^0(t), \mu^0(s))}{|t-s|} &< (j_0 \Delta_{j_0})^{-1} \\ &\leq (j \Delta_j)^{-1} \text{ if } j \geq j_0. \end{aligned}$$

This implies that for any $j \geq j_0$ and $\mu(\cdot) \in D(j_0)$

$$\sup_{t,s \in [0,T], |t-s| \leq \Delta_j} r(\mu^0(t), \mu^0(s)) < \frac{1}{j}.$$

For $j < j_0$ we can show

$$(3.11) \quad \sup_{t,s \in [0,T], |t-s| \leq \Delta_j} r(\mu^0(t), \mu^0(s)) = \max_{|t_k - t_l| \leq \Delta_j} r(\mu^0(t_k), \mu^0(t_l)).$$

where t [resp. s] belongs to some interval $[t_k, t_{k+1}]$ [resp. $[t_l, t_{l+1}]$]. If we can define $\|\mu\| = \sum_{n=0}^{\infty} \frac{1}{2^n} |\mu(n)|$ and $r(\mu, \nu) = \|\mu - \nu\|$, then the argument used in [4].section V also works here. To complete the argument note that for $h \in [t_k \vee (t_l + t - s), t_{k+1} \wedge (t_{l+1} + t - s)]$, $h \rightarrow \mu^0(h) - \mu^0(h + s - t)$ is an affine function, and use the fact that $\mu \rightarrow \|\mu\|$ is a convex function to conclude that $\mu^0(h) - \mu^0(h + s - t)$ achieves its maximum on the boundary of the interval. Thus in order to prove (3.11) it suffices to verify that the $\|\mu^0(t) - \mu^0(t_k)\|$ is less than or equal to the right-hand side of (3.11) when $t \in (t_l, t_{l+1})$ and $|t - t_k| \leq \Delta_j$. Since Δ_j / Δ_{j_0} is an integer we get that $|t_l - t_k| \vee |t_{l+1} - t_k| \leq \Delta_j$. This combined with the convexity of $h \rightarrow \|\mu^0(h) - \mu^0(t_k)\|$ implies the result.

Hence we have

$$(3.12) \quad D(j_0) \cap \{\mu^0(\cdot) \notin K'_a\} \subset \cup_{k,l,j < j_0} \{\mu(\cdot) \in D(j_0); r(\mu(t_k), \mu(t_l)) \geq \frac{1}{j}\}.$$

where the union extends to all pairs (k, l) with $0 \leq k < l \leq m_0 \wedge (\frac{T}{\Delta_j} + \frac{\Delta_j}{\Delta_{j_0}})$ and $j < j_0$. For any fixed k, l, j , Choose $n_j \in \mathbf{N}$ such that $\sum_{n > n_j} \frac{1}{2^n} < \frac{1}{2j}$. Then we have

$$\begin{aligned} & \hat{P}_u^{(N)}\{\mu(\cdot) \in D(j_0); r(\mu(t_k), \mu(t_l)) \geq \frac{1}{j}\} \\ & \leq \sum_{m=0}^{n_j} P_u^{(N)}\left\{\left|\frac{1}{N} \sum_{i=1}^N (f_m(x_i^{(N)}(t_k)) - f_m(x_i^{(N)}(t_l)))\right| \geq 1/4j, \right. \\ & \quad \left. \varepsilon_{X^{(N)}(\cdot)} \in \mathcal{D}_R\right\} \\ & \leq \sum_{m=0}^{n_j} \exp[-N(\beta/4j - 2c(f_m, R)\Delta_j(e^{2\beta} - 1))] \\ & \leq n_j \exp[-N(\frac{\beta}{2j} - 2\hat{c}(j, R)\Delta_j(e^{2\beta} - 1))] \end{aligned}$$

Let $\beta = 8j(a+1)$. We conclude

$$\hat{P}_u^{(N)}\{\mu(\cdot) \in D(j_0); r(\mu(t_k), \mu(t_l)) \geq \frac{1}{j}\} \leq n_j \exp[-N(a+1)]$$

Using the rough upper bound $m_0\Delta_j/\Delta_{j_0}$ on the pairs (k, l) we get

$$(3.13) \quad \hat{P}_u^{(N)}\{D(j_0) \cap \{\mu^0(\cdot) \notin K'_a\}\} \leq m_0 \left(\sum_{j=2}^{j_0-1} n_j \Delta_j / \Delta_{j_0}\right) \exp[-N(a+1)].$$

Combined with (3.6), (3.8) and (3.10) we have

$$\begin{aligned} & \hat{P}_u^{(N)}\{r_{0T}(\mu(\cdot), K_a) \geq \delta\} \\ & \leq \hat{P}_u^{(N)}\{\mu(\cdot) \in \mathcal{D}_R; r_{0T}(\mu(\cdot), K_a) \geq \delta\} + \hat{P}_u^{(N)}\{(\mathcal{D}_R)^c\} \\ & \leq \hat{P}_u^{(N)}\{\mathcal{D}_R \cap D(j_0)^c\} + \hat{P}_u^{(N)}\{(\mathcal{D}_R)^c\} \\ & \quad + \hat{P}_u^{(N)}\{\mu(\cdot) \in D(j_0); r_{0T}(\mu(\cdot), K_a) \geq \delta\} \\ & \leq \frac{T}{\Delta_{j_0}} n_0 e^{[-N(a+1)]} + \exp[-N(a+1)] \\ & \quad + m_0 \left(\sum_{j=2}^{j_0-1} n_j \Delta_j / \Delta_{j_0}\right) \exp[-N(a+1)]. \end{aligned}$$

Since all the coefficients depends only on j_0 , and j_0 is independent of N , there exists an integer N_2 such that $\forall N \geq N_2$ (3.5) is true. $\square$

The proof of theorem 5.3.2. For any $\gamma > 0, \varepsilon > 0$ and $s > 0$, choose $R \geq 1$ and a compact subset K_s of $\mathcal{D}_R$ from lemma 5.3.4, with $a = s$, such that there exists an integer $\hat{N}$, for every $N \geq \hat{N}$ we have

$$\hat{\mathcal{P}}_u^{(N)}\{(\mathcal{D}_R)^c\} \leq \exp[-Ns]$$

$$\Phi_u(s) \subset \mathcal{D}_R.$$

Let $K_s^\varepsilon = K_s \cap \{r_{0T}(\mu(\cdot), \Phi_u(s)) \geq \varepsilon/2\}$. Then K_s^ε is also compact in $\mathcal{D}_R$ in the uniform topology and thus the Skorohod topology.

$\forall \bar{\mu}(\cdot) \in K_s^\varepsilon$, apply lemma 5.3.3 with $I = s$ to show that there exists $N_{\bar{\mu}}$ and $\delta_{\bar{\mu}} < \varepsilon$ such that $\forall N \geq N_{\bar{\mu}}$ we have

$$(3.14) \quad \hat{\mathcal{P}}_u^{(N)}\{r_{0T}(\mu(\cdot), \bar{\mu}(\cdot)) < \delta_{\bar{\mu}}\} \leq \exp[-Ns]$$

Let $\{\{\mu(\cdot) \in \mathcal{D}_R; r_{0T}(\mu(\cdot), \bar{\mu}_i(\cdot)) < \delta_{\bar{\mu}_i}\}; i = 1, 2, \dots, L\}$ be a finite open covering of K_s^ε , $\delta = \min\{\frac{1}{2}\delta_{\bar{\mu}_i}; i = 1, 2, \dots, L\}$. For any $\mu(\cdot) \in \mathcal{D}_R$ satisfying $r_{0T}(\mu(\cdot), K_s) \leq \delta$ and $r_{0T}(\mu(\cdot), \Phi_u(s)) \geq \varepsilon$ there exists $\hat{\mu}(\cdot) \in K_s$ such that $r_{0T}(\mu(\cdot), \hat{\mu}(\cdot)) \leq \delta$. This implies that $r_{0T}(\hat{\mu}(\cdot), \Phi_u(s)) \geq \varepsilon/2$, i.e., $\hat{\mu}(\cdot) \in K_s^\varepsilon$. Hence

$$(3.15) \quad \begin{aligned} & \hat{\mathcal{P}}_u^{(N)}\{\mu(\cdot) \in \mathcal{D}_R; r_{0T}(\mu(\cdot), K_s) \leq \delta, r_{0T}(\mu(\cdot), \Phi_u(s)) \geq \varepsilon\} \\ & \leq \sum_{i=1}^L \hat{\mathcal{P}}_u^{(N)}\{\mu(\cdot) \in \mathcal{D}_R; r_{0T}(\mu(\cdot), \bar{\mu}_i(\cdot)) < \delta_{\bar{\mu}_i}\} \leq L \exp[-Ns]. \end{aligned}$$

Finally we have

$$(3.16) \quad \begin{aligned} & \hat{\mathcal{P}}_u^{(N)}\{r_{0T}(\mu(\cdot), \Phi_u(s)) \geq \varepsilon\} \\ & \leq \hat{\mathcal{P}}_u^{(N)}\{(\mathcal{D}_R)^c\} + \hat{\mathcal{P}}_u^{(N)}\{\mu(\cdot) \in \mathcal{D}_R; r_{0T}(\mu(\cdot), K_s) > \delta\} \\ & \quad + \hat{\mathcal{P}}_u^{(N)}\{\mu(\cdot) \in \mathcal{D}_R; r_{0T}(\mu(\cdot), K_s) \leq \delta, r_{0T}(\mu(\cdot), \Phi_u(s)) \geq \varepsilon\} \\ & \leq (L + 2) \exp[-Ns]. \end{aligned}$$

If we take N large enough we get (3.1). $\square$

Now we are ready to prove the main result of this section.

Theorem 5.3.5 (Upper bound on $\mathcal{D}_\infty$) Under assumptions (1.1)-(1.5) of chapter 4 and (1.3), for any closed subset A of $\mathcal{D}_\infty$ we have

$$(3.17) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \hat{\mathcal{P}}_u^{(N)}(A) \leq - \inf_{\mu \in A} S_u(\mu).$$

Proof: Let A be any closed subset of $\mathcal{D}_\infty$. If $\inf_{\mu \in A} S_u(\mu) = 0$ then the result is trivial. Now we assume that $s = \inf_{\mu \in A} S_u(\mu) > 0$. Choose $R \geq 1$ and $\tilde{N}$ such that for any $N \geq \tilde{N}$ we have

$$(3.18) \quad \hat{\mathcal{P}}_u^{(N)}\{(\mathcal{D}_R)^c\} \leq \exp[-Ns], \quad \Phi_u(s) \subset \mathcal{D}_R.$$

Obviously $\forall s > \gamma > 0$ $A_1 = A \cap \mathcal{D}_R$ and $\Phi_u(s - \gamma)$ are disjoint closed subsets of $\mathcal{D}_R$ closed in both the Skorohod and the uniform topology. Thus there exists a strictly positive constant $\varepsilon > 0$ such that $r_{0T}(A_1, \Phi_u(s - \gamma)) \geq \varepsilon$. This implies

$$(3.19) \quad \hat{\mathcal{P}}_u^{(N)}\{\mathcal{D}_R \cap A\} \leq \hat{\mathcal{P}}_u^{(N)}\{r_{0T}(\mu(\cdot), \Phi_u(s - \gamma)) \geq \varepsilon\}.$$

Since γ is arbitrary, combining (3.18), (3.19) and theorem 5.3.2 we get (3.17). $\square$

Chapter 6

Large Deviation: Upper Bound On Metric Space

6.1 Upper Bound

Let X be any set, d_1, d_2 be two metrics on X . Let $\mathcal{B}_1$ be the Borel σ -algebra of (X, d_1) with respect to which for every $x_0 \in X$ the function $f_2(x) = d_2(x, x_0)$ is measurable.

We assume

- (a). The topology of (X, d_1) is weaker than the topology of (X, d_2) .
- (b). Any bounded closed ball in (X, d_2) is compact in (X, d_1) .

Since $X = \bigcup_{n \geq 1} \{x: d_2(x, x_0) \leq n\}$ for some $x_0 \in X$, assumption (b) implies that (X, d_1) is separable. Let $X_0 \subset X$, $\mathcal{C}$ be a σ -algebra on X_0 containing $\mathcal{B}_1 \cap X_0$ but contained in the Borel σ -algebra of (X_0, d_2) . Let $\mathcal{B} = \{A \subset X; A \cap X_0 \in \mathcal{C}, A \text{ is in the Borel } \sigma\text{-algebra of the space } (X, d_2)\}$. Then $\mathcal{B}$ contains $\mathcal{B}_1$ and is contained in the Borel σ -algebra of space (X, d_2) . Let $\{P_N\}_{N \geq 1}$ be a sequence of probability measures on $(X_0, \mathcal{C})$, Z be the set of all bounded continuous functions on (X, d_2) whose restriction to X_0 are $\mathcal{C}$ -measurable, $(Z^*, \sigma\{Z^*, Z\})$ be the dual space of Z with *weak** topology $\sigma\{Z^*, Z\}$ and $(Z^*, \|\cdot\|^*)$ be the dual space of Z with strong topology $\|\cdot\|^* = \sup\{|\langle \cdot, g \rangle|; g \in Z \text{ and } \|g\|_\infty = \sup_{x \in X} |g(x)| \leq 1\}$. By definition we can see that $Z = C_b(X, d_2, \mathcal{B})$, the set of all $\mathcal{B}$ -measurable, bounded continuous functions on (X, d_2) and for each $N \geq 1$ we can extend P_N to $(X, \mathcal{B})$ by

letting $P_N(X \setminus X_0) = 0$. We will use the same notation for this extension.

If we introduce a σ -algebra $\tilde{\mathcal{B}}$ on Z^* as

$$\tilde{\mathcal{B}} = \{B \subset Z^*; \{x \in X : \delta_x \in B\} \in \mathcal{B}\}$$

then we obviously have

Lemma 6.1.1 *The map Θ defined by*

$$\Theta : (X, \mathcal{B}) \longrightarrow (Z^*, \tilde{\mathcal{B}}), \quad x \mapsto \delta_x \in Z^*.$$

is measurable.

We can check directly that the family $\mathcal{V}$ of open sets

$$V_{z_0^*}(f_1, \dots, f_k; \varepsilon_1, \dots, \varepsilon_k)$$

$$= \{z^* \in Z^*; |\langle z^*, f_i \rangle - \langle z_0^*, f_i \rangle| < \varepsilon_i, \quad i = 1, \dots, k\}, \quad z_0^* \in Z^*, \quad f_i \in Z, \quad \varepsilon_i > 0.$$

is contained in $\tilde{\mathcal{B}}$.

Let $Q_N(\cdot) = P_N \circ \Theta^{-1}(\cdot)$. Then by lemma 6.1.1, Q_N is a sequence of probability measures on $(Z^*, \tilde{\mathcal{B}})$. $\forall f \in Z, z^* \in Z^*$, we define

$$H_N(f) = \frac{1}{N} \log \int_{Z^*} \exp(N \langle z^*, f \rangle) Q_N(dz^*),$$

$$H(f) = \limsup_{N \rightarrow \infty} H_N(f),$$

$$H^*(z^*) = \sup_{f \in Z} \{\langle z^*, f \rangle - H(f)\}.$$

By definition we have $|H(f)| \leq \sup_x |f(x)|, \forall f \in Z$. The next lemma gives the upper bound of $\{Q_N\}_{N=1, \dots}$ on $(Z^*, \tilde{\mathcal{B}})$ which is crucial for the upper bound of $\{P_N\}_{N=1, \dots}$ on (X, d_2) .

Lemma 6.1.2 *For any strong bounded, $\tilde{\mathcal{B}}$ -measurable and weak* closed subset A of Z^* , we have*

$$(1.1) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log Q_N(A) \leq - \inf_{z^* \in A} H^*(z^*).$$

Note: The strong bounded and weak closed sets are compact in the weak* topology.*

Proof: Let A be any strong bounded, $\tilde{\mathcal{B}}$ -measurable and *weak** closed subset of Z^* . For any $z_0^* \in A$ and $\varepsilon > 0$, there exists an $f \in Z$ and a $\tilde{\mathcal{B}}$ -measurable *weak** open neighbourhood $V \subset Z^*$ of z_0^* such that

$$H^*(z_0^*) \leq \langle z_0^*, f \rangle - H(f) + \varepsilon,$$

$$\langle z_0^*, f \rangle - H(f) \leq \inf_{z^* \in A \cap V} \{\langle z^*, f \rangle - H(f)\} + \varepsilon.$$

For this fixed $f \in Z$, $V \subset Z^*$, we define

$$Y_{N,f}(z^*) = \exp[N(\langle z^*, f \rangle - H_N(f))].$$

Obviously $\int_{Z^*} Y_{N,f}(z^*) Q_N(dz^*) = 1$. This implies

$$\begin{aligned} Q_N(A \cap V) &\leq Q_N\{z^*; Y_{N,f}(z^*) \in Y_{N,f}(A \cap V)\} \\ &\leq \frac{\int_{Z^*} Y_{N,f}(z^*) Q_N(dz^*)}{\inf_{z^* \in A \cap V} Y_{N,f}(z^*)} = \frac{1}{\inf_{z^* \in A \cap V} Y_{N,f}(z^*)}. \end{aligned}$$

By definition there exists an N_0 such that

$$\forall N > N_0, H_N(f) \leq H(f) + \varepsilon.$$

Combining all these we obtain

$$\begin{aligned} (1.2) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log Q_N(A \cap V) &\leq - \inf_{z^* \in A \cap V} \{\langle z^*, f \rangle - H_N(f)\} \\ &\leq - \inf_{z^* \in A \cap V} \{\langle z^*, f \rangle - H(f)\} + \varepsilon \\ &\leq -\{\langle z_0^*, f \rangle - H(f)\} + 2\varepsilon \\ &\leq -H^*(z_0^*) + 3\varepsilon \leq - \inf_{z^* \in A \cap V} H^*(z^*) + 3\varepsilon. \end{aligned}$$

Since A is compact in the *weak** topology we can choose finite number of V_i , $i = 1, \dots, r$ such that $A \subset \bigcup_{i=1}^r V_i$ and each V_i satisfies (1.2). Then we have

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log Q_N(A) &\leq \max\{- \inf_{z^* \in A \cap V_i} H^*(z^*); i = 1, \dots, r\} \\ &\leq - \inf_{z^* \in A} H^*(z^*). \end{aligned}$$

□

The next lemma gives a sufficient condition for an element of Z^* to be a probability measure on $(X, \mathcal{B})$.

Lemma 6.1.3 *If $H^*(z^*) < \infty$ and for any decreasing sequence of functions $\{f_n\} \subset Z$ converging to zero pointwise we have $H(f_n) \rightarrow 0$, then z^* is a probability measure on $(X, \mathcal{B})$.*

Proof: Since z^* is a linear functional on Z , by the Daniell extension theorem (cf.[33],p.21) to prove the lemma it suffices to show that the following three conditions are satisfied:

$$(a). \langle z^*, f \rangle \geq 0 \text{ for } f \in Z, f \geq 0;$$

$$(b). \langle z^*, 1 \rangle = 1;$$

$$(c). \langle z^*, f_n \rangle \rightarrow 0 \text{ as } f_n \downarrow 0 \text{ pointwise and } f_n \in Z.$$

Suppose $f \geq 0, \lambda < 0$, we have $H(\lambda f) \leq 0$. This implies

$$\begin{aligned} \sup_{\lambda < 0} \lambda \langle z^*, f \rangle &\leq \sup_{\lambda < 0} \{ \lambda \langle z^*, f \rangle - H(\lambda f) \} \\ &\leq H^*(z^*) < \infty. \end{aligned}$$

This implies condition (a).

Similarly we have

$$\begin{aligned} \sup_{\lambda} \lambda \{ \langle z^*, 1 \rangle - 1 \} &= \sup_{\lambda} \{ \lambda \langle z^*, 1 \rangle - H(\lambda 1) \} \\ &\leq H^*(z^*) < \infty. \end{aligned}$$

which gives (b).

Finally let $\{f_n\} \subset Z$ be a decreasing sequence converging to zero. For any $\lambda > 0$ we have

$$\lambda \langle z^*, f_n \rangle \leq H^*(z^*) + H(\lambda f_n), \text{ this implies}$$

$$\lambda \limsup_{n \rightarrow \infty} \langle z^*, f_n \rangle \leq H^*(z^*).$$

which is just condition (c). $\square$

Now we are ready to establish the upper bound for $\mathcal{B}$ -measurable closed subsets of (X, d_2) .

Theorem 6.1.4 *Let (X, d_1, d_2) satisfy the assumptions (a) and (b). We have*

(i). *Under the assumption that*

$$(1.3) \quad \text{For every sequence } \{f_n\} \subset Z, f_n \downarrow 0, \text{ we have } H(f_n) \rightarrow 0.$$

then for any $\mathcal{B}$ -measurable, closed subset A of (X, d_2) we have

$$(1.4) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_N(A) \leq - \inf_{x \in A} L(x).$$

(ii). *If*

$$(1.5) \quad \text{For every sequence } \{f_n\} \subset C_b(X, d_2, \mathcal{B}_1), f_n \downarrow 0, \text{ we have } H(f_n) \rightarrow 0,$$

then for any $\mathcal{B}$ -measurable, closed subset A of (X, d_2) we have

$$(1.6) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_N(A) \leq - \inf_{x \in A} \tilde{L}(x).$$

where $L(x) = H^*(\Theta(x))$, $\tilde{L}(x) = \sup_{f \in C_b(X, d_2, \mathcal{B}_1)} \{f(x) - H(f)\}$. Obviously $\tilde{L}(x) \leq L(x)$, $\forall x \in X$.

Proof: (i). Let A be any $\mathcal{B}$ -measurable closed subset of (X, d_2) , then $\Theta(A) \subset Z^*$ and is contained in the strong unit ball of Z^* . Let $\overline{\Theta(A)}$ denote the closure of $\Theta(A)$ in Z^* in the *weak** topology. Since the strong unit ball of Z^* is compact in the *weak** sense, $\overline{\Theta(A)}$ is contained in the strong unit ball of Z^* and thus is compact in the *weak** topology. From (1.9) below we can see that $\Theta^{-1}(\overline{\Theta(A)}) = A$. Thus $\overline{\Theta(A)}$ is $\tilde{\mathcal{B}}$ -measurable. If $\inf_{z^* \in \overline{\Theta(A)}} H^*(z^*) < \infty$ then by lemma 6.1.2, lemma 6.1.3 and condition (1.3) we have

$$(1.7) \quad \begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_N(A) &\leq \limsup_{N \rightarrow \infty} \frac{1}{N} \log Q_N(\overline{\Theta(A)}) \\ &\leq - \inf_{z^* \in \overline{\Theta(A)}} H^*(z^*) = - \inf_{z^* \in \overline{\Theta(A)} \cap M_1(X, \mathcal{B})} H^*(z^*). \end{aligned}$$

where $M_1(X, \mathcal{B})$ denotes the set of probability measures on $(X, \mathcal{B})$. If $\inf_{z^* \in \overline{\Theta(A)}} H^*(z^*) = \infty$, then $\inf_{z^* \in \overline{\Theta(A)}} H^*(z^*) \leq \inf_{z^* \in \overline{\Theta(A)} \cap M_1(X, \mathcal{B})} H^*(z^*) = \infty$. Thus in any case (1.7) is true.

For any $z^* \in \overline{\Theta(A)} \cap M_1(X, \mathcal{B})$ we can find a sequence $\{x_n\} \subset A$ such that

$$(1.8) \quad \forall f \in Z, \langle \delta_{x_n}, f \rangle \longrightarrow \langle z^*, f \rangle \text{ as } n \rightarrow \infty.$$

Suppose $\{x_n\}$ does not have any convergent subsequence in (X, d_2) . Then the set $S = \{x_1, x_2, \dots\}$ is closed and so is any subset of S . Let $S_0 = \{x_{n_k}\}_{k=1, \dots}$ be any infinite subset of S . For any $M \geq 1$, let $f_M(x) = d_2(x, S_0) \wedge M$ where $d_2(x, S_0) = \inf\{d_2(x, y); y \in S_0\}$. By definition we have that $\{f_M\}_{M \geq 1} \subset Z$. Thus by (1.8), for any $M \geq 1$, $\langle \delta_{x_{n_k}}, f_M \rangle \rightarrow \langle z^*, f_M \rangle = 0$. Since $d_2(x, S_0) > 0$ for $x \notin S_0$, we get that $z^*(S_0) = 1$. This is a contradiction since z^* is a probability measure on $(X, \mathcal{B})$. Let x, y be any two limit points of S , then for any $f \in Z$, $\langle \delta_x, f \rangle = \langle \delta_y, f \rangle$. Since Z is a separating set of X , $x = y$. This implies that there is an $x_0 \in A$ such that $\delta_{x_0} = z^*$ on $\mathcal{B}$, i.e., $\forall f \in Z, \langle \delta_{x_0}, f \rangle = \langle z^*, f \rangle$. This shows

$$(1.9) \quad \overline{\Theta(A)} \cap M_1(X, \mathcal{B}) = \Theta(A).$$

Combining (1.7) and (1.9) we get (1.4).

The proof for (ii) is almost the same except that Z should be replaced by $C_b(X, d_2, \mathcal{B}_1)$. $\square$

Corollary 6.1.5 *Let $\bar{X}_0$ be the closure of X_0 in the space (X, d_2) . Then we have*

(i). *If $\bar{X}_0 \cap \{x \in X; L(x) < \infty\} \subset X_0$ and (1.3) is true, then for any $\mathcal{C}$ -measurable, closed subset A of (X_0, d_2) we have*

$$(1.10) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_N(A) \leq - \inf_{x \in A} L(x).$$

(ii). *If $\bar{X}_0 \cap \{x \in X; \tilde{L}(x) < \infty\} \subset X_0$ and (1.3) is true, then for any $\mathcal{C}$ -measurable, closed subset A of (X_0, d_2) we have*

$$(1.11) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_N(A) \leq - \inf_{x \in A} \tilde{L}(x).$$

Proof: Noting that the closure $\bar{A}$ in (X, d_2) of any closed subset A of (X_0, d_2) is in $\mathcal{B}$, the corollary follows directly from theorem 6.1.4. $\square$

The following theorem shows that the condition (1.3) is implied by the exponential tightness of $\{P_N\}$ on (X, d_2) .

Theorem 6.1.6 *If the condition*

$$(1.12) \quad \forall a > 0, \exists \text{ compact subset } K_a \subset (X, d_2), \exists \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_N\{x: x \notin K_a\} \leq -a,$$

is satisfied, then $\forall \{f_n\} \subset \mathcal{C}, f_n \downarrow 0$ pointwise, we have $H(f_n) \rightarrow 0$.

Proof: $\forall N \geq 1$ we have

$$\begin{aligned} H_N(f_n) &= \frac{1}{N} \log \int_X \exp[N f_n(x)] P_N(dx) \\ &\leq \frac{1}{N} \log \left\{ \int_{K_a} \exp[N f_n(x)] P_N(dx) + \exp[N \|f_1\|_\infty] P_N(K_a^c) \right\} \\ &\leq \frac{\log 2}{N} + \max \left\{ \frac{1}{N} \log \int_{K_a} \exp[N f_n(x)] P_N(dx), \|f_1\|_\infty - a \right\}. \end{aligned}$$

Since K_a is compact in (X, d_2) , $f_n \rightarrow 0$ uniformly on it. Let $a = \|f_1\|_\infty$. Then $\forall \varepsilon > 0$ there exists $N_0 > 0$ and $n_0 > 0$ such that

$$0 \leq H_N(f_n) \leq \varepsilon, \quad \forall N > N_0, \quad n > n_0.$$

Thus we get $H(f_n) \rightarrow 0$. $\square$

The following theorem gives some connection between condition (1.5) and the exponential decay rate of the complement of balls.

Theorem 6.1.7 *If the condition*

$$(1.13) \quad \forall a > 0, \exists b \geq 0, x_0 \in X, \exists \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_N\{x: d_2(x, x_0) > b\} \leq -a,$$

is satisfied, then $\forall \{f_n\} \subset C_b(X, d_1), f_n \downarrow 0$ pointwise, we have $H(f_n) \rightarrow 0$.

Proof: Noting that condition (1.13) is just the exponential tightness of $\{P_N\}$ on (X, d_1) , the proof follows easily from the proof of theorem 6.1.6. $\square$

Remarks. 1. The exponential tightness is crucial in the passage from a weak to a full large deviation principle. (cf. lemma 2.1.5 of [13]) But it is not a necessary condition in general.

2. If X is a locally compact topological space or a Polish space, then it is necessary in order to get a full large deviation principle (cf. exercises 1.2.19 and 4.1.10 in [12] or theorem (P) of [35]).

3. Our discussion shows that the exponential tightness implies condition (1.3); While condition (1.3) implies condition (1.5). But condition (1.5) may not be implied by (1.13). The problem whether the exponential tightness is equivalent to (1.3) and the exact relation between (1.5) and (1.13) are not clear.

The reason that we do not use the Borel σ -algebra on the space (X, d_2) is that in general the space may be non-separable and thus many interesting probability distributions may not have extensions to the Borel σ -algebra. For more detail discussions, see for instance theorem 5.5 and lecture 23 in [18].

6.2 Application

In this section a nice application of theorem 6.1.4 will be given by using the vector-valued L^p space.

6.2.1 Some Properties of Vector-valued L^p Space

Let $E = \{0, 1, \dots\}$ be equipped with the discrete topology, φ be a strictly positive function on E satisfying: $\lim_{x \rightarrow \infty} \frac{\varphi(x)}{x} = +\infty$ and

$$\mathcal{X} = \{g; g \text{ is a real valued function on } E,$$

$$\|g\|_{\varphi} < \infty \text{ and } \lim_{|x| \rightarrow \infty} \frac{g(x)}{\varphi(x)} \text{ exists}\}$$

where

$$\|g\|_{\varphi} = \sup_{x \in E} \frac{|g(x)|}{\varphi(x)}.$$

2 OF/DE 2

PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

Then $(\mathcal{X}, ||| \cdot |||_{\varphi})$ is a separable normed vector space.

$(\mathcal{X}^*, ||| \cdot |||_{\varphi}^*)$ is the strong dual space of $(\mathcal{X}, ||| \cdot |||_{\varphi})$. Where

$$\forall \nu \in \mathcal{X}^*, ||| \nu |||_{\varphi}^* = \sup \{ |(\nu, g)|; g \in \mathcal{X}, ||| g |||_{\varphi} \leq 1 \}$$

$\sigma\{\mathcal{X}^*, \mathcal{X}\}$ will be used to denote *weak** topology of $\mathcal{X}^*$.

Definition 6.2.1 A map F from $[0, T]$ to $(\mathcal{X}, ||| \cdot |||_{\varphi})$ is called measurable if $\forall \varepsilon > 0$, there exists a compact set $K \subset [0, T]$ such that $\lambda([0, T] \setminus K) < \varepsilon$ and the restriction $F|_K$ of F on K is continuous. Where λ denotes the Lebesgue measure.

Let

$$\begin{aligned} L_{\mathcal{X}}^1 &= L^1([0, T], \mathcal{X}) = \{F : [0, T] \rightarrow \mathcal{X}, F \text{ is measurable and} \\ &|F|_1 = \int_0^T ||| F(t) |||_{\varphi} dt < \infty\}. \end{aligned}$$

and any two elements are considered to be the same if they are equal to each other t -almost surely.

Let

$$C_{\mathcal{X}} = C([0, T], \mathcal{X})$$

$$= \{F : [0, T] \rightarrow \mathcal{X}; F \text{ is continuous}\}.$$

$$S_{\mathcal{X}} = \{F \in L_{\mathcal{X}}^1;$$

there exists finite number of disjoint measurable set $A_i \subset [0, T]$

$$\text{and } g_i \in \mathcal{X} \ni F = \sum_i g_i \chi_{A_i}\}.$$

Note. Elements of $S_{\mathcal{X}}$ are called simple functions.

Let $C_{\mathcal{X}}, S_{\mathcal{X}}$ be endowed with the subspace topology of $(L_{\mathcal{X}}^1, |\cdot|_1)$. Then as in the real-valued case we have

Lemma 6.2.1 $C_{\mathcal{X}}$ and $S_{\mathcal{X}}$ are dense subset of $L_{\mathcal{X}}^1$.

Proof: It is easy to check that any element of L^1_X can be approximated in $\|\cdot\|_1$ norm by bounded element of L^1_X . By the uniform continuity, any element of C_X can be approximated in the same sense by simple functions. Thus to prove the lemma it suffices to show that any bounded element of L^1_X can be approximated by elements of C_X . Let F be any bounded element of L^1_X with bound M . Then by definition $\forall \varepsilon > 0, \exists$ a compact set $K \subset [0, T]$ such that $\lambda([0, T] \setminus K) < \varepsilon/2$ and $F|_K$ is continuous. Let $a = \inf\{t \in K\}, b = \sup\{t \in K\}$. Then $a, b \in K$. Let $K^c = [0, T] \setminus K$. Then there are at most countable number of disjoint open intervals $\{(a_i, b_i)\}_{i=1, \dots}$ such that $K^c = [0, a) \cup (b, T] \cup (\cup_i (a_i, b_i))$ and $a_i, b_i \in K$. Now we introduce an X -valued function $F_\varepsilon(t)$ as

$$F_\varepsilon(t) = \begin{cases} F(a) & \text{for } t \in [0, a) \\ F(b) & \text{for } t \in (b, T] \\ \frac{t-a_i}{b_i-a_i} F(a_i) + \frac{b_i-t}{b_i-a_i} F(b_i) & \text{for } t \in (a_i, b_i) \ i = 1, \dots \\ F(t) & \text{for } t \in K \end{cases}$$

Then $F_\varepsilon \in C_X$ and $\|F - F_\varepsilon\|_1 \leq 2M\varepsilon$ which gives the result. $\square$

Remark. (a). Since X is separable, L^1_X is separable by lemma 6.2.1.

(b). Since L^1_X is a separable metric space, it is second countable. As the subspace of a second countable space, C_X is also second countable and thus separable.

(c). Our definition of measurability is equivalent to the definition in [19] on page 106. But in general if the interval $[0, T]$ is replaced by an arbitrary metric space then our definition is stronger.

Next we define

$$\Lambda_{X^*} = \{\nu(\cdot) : [0, T] \rightarrow X^*, \exists t \mapsto \langle \nu(t), g \rangle$$

$$\text{is measurable for any } g \in X \text{ and } \|\nu(\cdot)\|_\infty < \infty\}.$$

where $\|\nu(\cdot)\|_\infty = \text{ess sup}_{0 \leq t \leq T} \|\nu(t)\|_{X^*}$.

Since any element $F \in C_X$ can be approximated by a sequence of simple functions we can see that for any $\nu(\cdot) \in \Lambda_{X^*}$, the map $t \rightarrow \langle \nu(t), F(t) \rangle$ is measurable.

Here any two elements $\nu(\cdot), \mu(\cdot)$ are also considered to be the same if they are equal t -almost everywhere. The next two theorems will show that the strong dual $(L_X^1)^*$ of L_X^1 can be identified as Λ_{X^*} .

Theorem 6.2.2 *Let*

$$L^1 = \{h : [0, T] \rightarrow R, t \mapsto h(t) \text{ is measurable and}$$

$$\|h\|_1 = \int_0^T |h(t)| dt < \infty\}$$

$$L^\infty = \{h : [0, T] \rightarrow R, t \mapsto h(t) \text{ is measurable and}$$

$$\|h\|_\infty = \sup_{0 \leq t \leq T} |h(t)| < \infty\}.$$

and $U : X \rightarrow L^\infty$ be any linear continuous form. Then there exists an $\nu(\cdot) \in \Lambda_{X^*}$ such that

$$(i). \quad U(g)(t) = \langle \nu(t), g \rangle;$$

$$(ii). \quad \|U\| = \|\nu(\cdot)\|_\infty;$$

$$(iii). \quad H : \Lambda_{X^*} \rightarrow L_b(X, L^\infty),$$

$$\nu(\cdot) \mapsto \langle \nu(\cdot), \cdot \rangle \text{ is an one-to-one isometry.}$$

where $L_b(X, L^\infty)$ is the space of all continuous linear forms from X to L^∞ .

Proof: Since for any $t \in [0, T]$, $g \mapsto U(g)(t)$ is a continuous linear form on X , there exists an $\nu(t) \in X^* \ni U(g)(t) = \langle \nu(t), g \rangle$. Since for any $g \in X$, $U(g) \in L^\infty$, we conclude that $\langle \nu(t), g \rangle$ is measurable. By definition we have

$$\begin{aligned} \operatorname{ess\,sup}_{0 \leq t \leq T} \|\nu(t)\|_\phi^* &= \sup_{0 \leq t \leq T} \sup_{\|g\|_\phi \leq 1} \{|\langle \nu(t), g \rangle|\} \\ &= \sup_{0 \leq t \leq T} \sup_{\|g\|_\phi \leq 1} \{|U(g)(t)|\} = \sup_{\|g\|_\phi \leq 1} \|U(g)\|_\infty \\ &= \|U\| < \infty. \end{aligned}$$

Thus $\nu(\cdot) \in \Lambda_{X^*}$ and (i), (ii) are true. In order to prove (iii) it suffices to show that the map H is one-to-one. Condition (i) implies that H is onto. Let's assume that

$H(\nu(\cdot)) = H(\nu^*(\cdot))$. Then for any $g \in \mathcal{X}$ there exists a zero measure set N_g such that $\langle \nu(t), g \rangle = \langle \nu^*(t), g \rangle$ for $t \in [0, T] \setminus N_g$. Let $\mathcal{X}_0$ be a countable dense subset of $\mathcal{X}$, $N = \bigcup_{g \in \mathcal{X}_0} N_g$. Then N is a zero measure set and $\langle \nu(t), g \rangle = \langle \nu^*(t), g \rangle$ for $g \in \mathcal{X}_0, t \in [0, T] \setminus N$. Thus $\nu(t) = \nu^*(t)$ λ -a.e. By definition we have $\nu(\cdot) = \nu^*(\cdot)$. $\square$

Theorem 6.2.3 *The strong dual of $L^1_{\mathcal{X}}$ can be identified as $\Lambda_{\mathcal{X}^*}$.*

Proof: For any $\nu(\cdot) \in \Lambda_{\mathcal{X}^*}$ and any $F \in S_{\mathcal{X}}$ the function $h(t) = \langle \nu(t), F(t) \rangle \in L^1$.

By lemma 6.2.1 this is true for any $F \in L^1_{\mathcal{X}}$. If we define

$$M_{\nu(\cdot)}(F) = \int_0^T \langle \nu(t), F(t) \rangle dt, \forall F \in L^1_{\mathcal{X}}.$$

Then $M_{\nu(\cdot)}$ is a linear functional on $L^1_{\mathcal{X}}$. By direct calculation we have that

$$\|M_{\nu(\cdot)}\| = \|\nu(\cdot)\|_{\infty} < \infty.$$

Thus $M_{\nu(\cdot)}$ is a continuous linear form on $L^1_{\mathcal{X}}$. On the other hand, let M be any continuous linear form on $L^1_{\mathcal{X}}$. For any $g \in \mathcal{X}, h \in L^1$ let $hg(t) = h(t)g$. Then $hg \in L^1_{\mathcal{X}}$ and $M(hg)$ is a continuous linear form on L^1 for any fixed $g \in \mathcal{X}$. Thus there exists $H(g) \in L^{\infty}$ such that $\langle H(g), h \rangle = M(hg)$. Obviously H is linear. By direct calculation

$$\begin{aligned} \|H\| &= \sup_{\|g\|_{\infty} \leq 1} \|H(g)\|_{\infty} \\ &= \sup_{\|g\|_{\infty} \leq 1} \sup_{\|h\|_1 \leq 1} \{|M(hg)|\} \leq \|M\| < \infty. \end{aligned}$$

This means that H is a continuous linear form from $\mathcal{X}$ to L^{∞} . By theorem 6.2.2 there exists an element $\nu(\cdot)$ of $\Lambda_{\mathcal{X}^*}$ such that $H(g) = \langle \nu(\cdot), g \rangle$ which implies that $M(hg) = \int_0^T h(t) \langle \nu(t), g \rangle dt$. Thus $M = M_{\nu(\cdot)}$ on $S_{\mathcal{X}}$.

By lemma 6.2.1 we actually have $M = M_{\nu(\cdot)}$ on $L^1_{\mathcal{X}}$. By an argument similar to the proof of theorem 6.2.2 we can show that

$$G : \Lambda_{\mathcal{X}^*} \rightarrow L_b(L^1, R); \nu(\cdot) \mapsto M_{\nu(\cdot)}$$

is an one-to-one isometry. Thus the result is proved. $\square$

Since C_X is dense in L_X^1 , the strong dual $(C_X)^*$ of C_X can also be identified with Λ_{X^*} .

Since C_X is separable, we can introduce one more topology on Λ_{X^*} as follow. Let C be a countable dense subset of C_{X^*} . $\forall \mu(\cdot), \nu(\cdot) \in \Lambda_{X^*}$ we define

$$\gamma(\mu(\cdot), \nu(\cdot)) = \sum_{F_n \in C} \frac{1}{2^n} |\langle \mu(\cdot), F_n \rangle - \langle \nu(\cdot), F_n \rangle| \wedge 1.$$

which is a metric on Λ_{X^*} . It is easy to see that the topology of space (Λ_{X^*}, γ) is weaker than the *weak** topology on Λ_{X^*} and (Λ_{X^*}, γ) is separable.

For more details about the vector-valued L^p space, refer to [19] or [20].

6.2.2 Upper Bounds on Space of Measure-valued Cadlag Paths

Let $\mathcal{B}_1$ denote the Borel σ -algebra of (Λ_{X^*}, γ) . Let $M_1(E)$ denote the set of all probability measures on E with the usual weak topology and the Prohorov metric ρ . For $m \geq 1$, we define $M_{\varphi, m}^w(E) = \{u \in M_1(E); \langle u, \varphi \rangle \leq m\}$ endowed with the subspace topology of $M_1(E)$. Define $\mathcal{D}_m = D([0, T], M_{\varphi, m}^w(E))$, $\mathcal{D}_\infty = \bigcup_{m \geq 1} \mathcal{D}_m$. Then we have

Lemma 6.2.4 $\mathcal{D}_\infty \subset \Lambda_{X^*}$.

Proof: To prove the lemma it suffices to verify the measurability of the map $t \mapsto \langle \mu(t), f \rangle$ for any $\mu(\cdot) \in \mathcal{D}$, $f \in X$. Let $f_n(x) = f(x) \wedge n$, then $\langle \mu(t), f_n \rangle$ is right continuous and thus measurable. Since $\langle \mu(t), f \rangle = \lim_{n \rightarrow \infty} \langle \mu(t), f_n \rangle$, the measurability follows. By definition we have $\mathcal{D}_\infty \subset \Lambda_{X^*}$. $\square$

Let $(\mathcal{D}_m, d_{sko})$ denote the space $\mathcal{D}_m$ endowed with the Skorohod topology, d_{sko} the Skorohod metric. $\mathcal{D}_\infty^{ind}$ denotes the space $\mathcal{D}_\infty$ endowed with the “inductive topology” of $(\mathcal{D}_m, d_{sko})$, i.e., V is open in $\mathcal{D}_\infty^{ind}$ if and only if for each $m \geq 1$, $V \cap \mathcal{D}_m$ is open in $\mathcal{D}_m$. Let $\mathcal{D}_\infty^{sko}$ be the space $\mathcal{D}_\infty$ with the Skorohod topology, $\mathcal{C}_{ind}$, $\mathcal{C}_{sko}$ denote the Borel σ -algebra of $\mathcal{D}_\infty^{ind}$ and $\mathcal{D}_\infty^{sko}$ respectively. Let $\mathcal{C}_0 = \mathcal{B}_1 \cap \mathcal{D}_\infty$. Then we have

Lemma 6.2.5 $\mathcal{C}_{ind} = \mathcal{C}_{sko} \supset \mathcal{C}_0$.

Proof: Since the "inductive topology" is stronger than the Skorohod topology we have

$$\mathcal{C}_{ind} \supset \mathcal{C}_{sko}.$$

On the other hand $\forall m \geq 1, (\mathcal{D}_m, d_{sko})$ is a closed subset of $\mathcal{D}_{\infty}^{sko}$. Any closed subset of $\mathcal{D}_m$ is $\mathcal{C}_{sko}$ -measurable and so is the open subset. Let V be an arbitrary open set in $\mathcal{D}_{\infty}^{ind}$. Then $V \cap \mathcal{D}_m \in \mathcal{C}_{sko} \forall m \geq 1$. Finally $V = \bigcup_{m \geq 1} (V \cap \mathcal{D}_m) \in \mathcal{C}_{sko}$. Thus $\mathcal{C}_{ind} = \mathcal{C}_{sko}$.

It's easy to see that $\forall F \in \mathcal{C}, n \geq 1, \langle \mu(\cdot), F \wedge n \rangle = \int_0^T \langle \mu(t), F(t) \wedge n \rangle dt$ is continuous on $\mathcal{D}_{\infty}^{sko}$. Letting n go to ∞ then we get that $\forall F \in \mathcal{C}, \langle \mu(\cdot), F \rangle$ is $\mathcal{C}_{sko}$ -measurable. This implies that γ is $\mathcal{C}_{sko}$ -measurable. Since (Λ_X, γ) is separable, we conclude that $\mathcal{C}_0 \subset \mathcal{C}_{sko}$. Hence the lemma follows. $\square$

Remark. By proposition 7.1 of chapter 3 in [21] we actually have that $\mathcal{C}_{ind} = \mathcal{C}_{sko} = \mathcal{C}_0$. But here lemma 6.2.5 is enough for applications. We will use $\mathcal{C}$ to denote $\mathcal{C}_{sko}$ in the sequel.

Now let $D([0, T], E^{\otimes N})$ be equipped with the Skorohod topology, $\{P_N\}$ be a sequence of probability measures on $D([0, T], E^{\otimes N})$. $\forall N \geq 1$, let $\mathcal{P}_N$ be defined as

$$\mathcal{P}_N = P_N \circ \varepsilon_{x^{(N)}}^{-1}.$$

$$\varepsilon_{x^{(N)}} : (x_i^{(N)})_{1 \leq i \leq N} \in D([0, T], E^{\otimes N}) \longrightarrow (t \rightarrow \frac{1}{N} \sum_{i=1}^N \delta_{x_i^{(N)}(t)}) \in \mathcal{D}_{\infty}.$$

Then $\{\mathcal{P}_N\}_{N \geq 1}$ is a sequence of probability measures on $(\mathcal{D}_{\infty}, \mathcal{C}_{sko})$. The main result of this section is the following:

Theorem 6.2.6 Assume that there exists a function $h(x)$ with $\lim_{x \rightarrow \infty} h(x) = \infty$ and $\beta > 0$ such that

$$(2.1) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log E^{P_N} \left\{ \exp \left[\beta \sum_{i=1}^N \int_0^T h(x_i^{(N)}(t)) dt \right] \right\} < \infty.$$

Then we have that

(i). If (1.3) is satisfied, then for any $\mathcal{C}$ -measurable, closed subset A of $\mathcal{D}_{\infty}^{ind}$ we have

$$(2.2) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_N(A) \leq - \inf_{x \in A} L(x).$$

(ii). If (1.5) is satisfied, then for any $\mathcal{C}$ -measurable, closed subset A of $\mathcal{D}_\infty^{\text{ind}}$ we have

$$(2.3) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_N(A) \leq - \inf_{x \in A} \dot{L}(x).$$

where

$$(2.4) \quad L(\mu) = \sup\{F(\mu) - H(F); F \in \mathcal{F}\}, \mu \in \Lambda_X.$$

$$(2.5) \quad \dot{L}(\mu) = \sup\{F(\mu) - H(F); F \in \tilde{\mathcal{F}}\}, \mu \in \Lambda_X.$$

$$(2.6) \quad H(F) = \limsup_{N \rightarrow \infty} \frac{1}{N} \log E^{P_N} \left\{ \exp \left[NF \left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i^{(N)}(\cdot)} \right) \right] \right\}.$$

where $\mathcal{F}$ and $\tilde{\mathcal{F}}$ are the sets of all bounded, $|\cdot|_\infty$ -continuous functions which are $\mathcal{B}$ and $\mathcal{B}_1$ measurable respectively. $\mathcal{B}$ is defined in the following.

In order to prove this theorem we need a series of lemmas. Let

$$\mathcal{R} = \{A; A \text{ is closed in } (\Lambda_{X^*}, |\cdot|_\infty) \text{ and } A \cap \mathcal{D}_\infty \in \mathcal{C}_{sko}\}$$

We introduce the σ -algebra $\mathcal{B}$ on Λ_{X^*} to be the family of all subset $B \subset \Lambda_{X^*}$ such that B is a Borel measurable subset of space (X, d_2) and $B \cap \mathcal{D}_\infty^{\text{ind}} \in \mathcal{C}$. Then it's easy to see that $\mathcal{B}$ contains $\mathcal{B}_1$ and $\mathcal{R}$. We can extend P_N to $(\Lambda_{X^*}, \mathcal{B})$ by letting $P_N(\Lambda_{X^*} \setminus \mathcal{D}_\infty) = 0$. We use the same symbol for this extension.

Let $\mathcal{X}_w^*$ denote the space of $\mathcal{X}^*$ with the coarsest topology with respect to which $\forall f \in C_b(E), \langle \mu, f \rangle$ is continuous. $\mathcal{D} = \{\mu(\cdot) \in D([0, T], \mathcal{X}_w^*), |\mu(\cdot)|_\infty < \infty\}$. Then $\mathcal{D}$ is a subset of Λ_{X^*} . We have

Lemma 6.2.7 *The closure $\bar{\mathcal{D}}_\infty$ of $\mathcal{D}_\infty$ in the space $(\Lambda_{X^*}, |\cdot|_\infty)$ is a subset of $\mathcal{D}$.*

Proof: Let $\mu(\cdot)$ be an arbitrary element of $\bar{\mathcal{D}}_\infty$. If $\mu(\cdot) \in \mathcal{D}_\infty$ then we are done. Otherwise let $\mu(\cdot) \in \bar{\mathcal{D}}_\infty \setminus \mathcal{D}_\infty$ and $\{\mu_n(\cdot)\}_{n=1,2,\dots} \subset \mathcal{D}_\infty$ be a $|\cdot|_\infty$ -convergent sequence with limit point $\mu(\cdot)$. Then there exists a zero measure set $N_1 \subset [0, T]$ such that $\sup\{||\mu_n(t) - \mu(t)|||_\infty^*; t \notin N_1\} \rightarrow 0$ as $n \rightarrow \infty$. This implies that $\forall f \in C_b(E), \sup\{|\langle \mu_n(t), f \rangle - \langle \mu(t), f \rangle|; t \notin N_1\} \rightarrow 0$ as $n \rightarrow \infty$. Applying the result of

problem 2 of [21] on page 150 we have

$$\begin{aligned} \forall n, m \geq 1, d_{sko}(\mu_n(\cdot), \mu_m(\cdot)) \\ \leq \sup_{t \in [0, T]} \rho(\mu_n(t), \mu_m(t)) \leq (|\mu_n(\cdot) - \mu_m(\cdot)|_\infty)^{1/2}. \end{aligned}$$

which implies that $\{\mu_n(\cdot)\}_{n=1, \dots}$ is Cauchy in the Skorohod space $D([0, T], M_1(E))$. Since $D([0, T], M_1(E))$ is a Polish space one get that there exists an element $\bar{\mu}(\cdot)$ of $D([0, T], M_1(E))$ such that $d_{sko}(\mu_n(\cdot), \bar{\mu}(\cdot)) \rightarrow 0$ as $n \rightarrow \infty$. Thus there again exists a zero measure set $N_2 \subset [0, T]$ such that $\forall t \in N_2, \rho(\mu_n(t), \bar{\mu}(t)) \rightarrow 0$ as $n \rightarrow \infty$. Since $\sup_n \{|\mu_n(\cdot)|_\infty\}$ is finite we have that $\bar{\mu}(\cdot) \in \mathcal{D}_\infty$.

Let $N = N_1 \cup N_2$. Then the above argument implies that $\forall t \in N^c, f \in C_b(E), \langle \mu(t), f \rangle = \langle \bar{\mu}(t), f \rangle$. If we define

$$\tilde{\mu}(t) = \begin{cases} \mu(t) & \text{for } t \in N^c \\ \bar{\mu}(t) & \text{for } t \in N \end{cases}$$

then $\tilde{\mu}(\cdot) \in \mathcal{D}$. Since two elements are considered to be the same in $\Lambda_{\mathcal{X}}$ if they coincide t -almost surely, we conclude the lemma. $\square$

The fact that $\mathcal{P}_N$ is concentrated on $\mathcal{D}_\infty$ implies

Lemma 6.2.8 *For any $\mu(\cdot) \notin \bar{\mathcal{D}}_\infty, L(\mu(\cdot)) = \tilde{L}(\mu(\cdot)) = \infty$.*

Proof: Let $\mu(\cdot) \notin \bar{\mathcal{D}}_\infty$. Since $\bar{\mathcal{D}}_\infty$ is closed in $(\Lambda_{\mathcal{X}}, |\cdot|_\infty)$ by a construction similar to the construction in the proof of Urysohn's lemma we can find a $\mathcal{B}_1$ -measurable, $|\cdot|_\infty$ -continuous bounded function F such that $F(\mu(\cdot)) = 1$ and F is zero on $\bar{\mathcal{D}}_\infty$. Let $F_n = nF$, then $H(F_n) = 0$. Let $n \rightarrow \infty$, we get the result. $\square$

The next two lemmas will show that $\{\mu(\cdot) \in \Lambda_{\mathcal{X}}; \tilde{L}(\mu(\cdot)) < \infty\} \cap \bar{\mathcal{D}}_\infty \subset \mathcal{D}_\infty$ which implies that $\{\mu(\cdot) \in \Lambda_{\mathcal{X}}; L(\mu(\cdot)) < \infty\} \cap \bar{\mathcal{D}}_\infty \subset \mathcal{D}_\infty$ since $\tilde{L} \leq L$. This is crucial in proving theorem 6.2.6.

Applying theorem 1 on page 315 of [27] on E we have

Lemma 6.2.9 *Let u be any element of $\mathcal{X}^*$. If the following three conditions are satisfied:*

$$(a). \forall f \in C_b(E), f \geq 0 \text{ we have } \langle u, f \rangle > 0;$$

$$(b). \langle u, 1 \rangle = 1;$$

$$(c). \lim_{n \rightarrow \infty} \langle u, \chi_{E_n} \rangle = 0;$$

where $E_n = \{n+1, \dots\}$. Then $u \in M_\varphi(E)$.

Proof: Conditions (a) and (c) implies conditions (1) and (3) in theorem 1 on page 315 of [27]. Thus u is a finite measure on E . By (b) we have that u is in fact a probability measure. Since $u \in X^*$ we get that $u \in M_\varphi(E)$. $\square$

Lemma 6.2.10 Assume that the sequence $\{\mathcal{P}_N\}_{N=1, \dots}$ satisfies condition (2.1). Then $\{\mu(\cdot) \in \Lambda_{X^*}; \tilde{L}(\mu(\cdot)) < \infty\} \cap \bar{\mathcal{D}}_\infty \subset \mathcal{D}_\infty$.

Proof: First we prove that for any $\mu(\cdot) \in \bar{\mathcal{D}}_\infty$ with $L(\mu(\cdot)) < \infty$ we have

$$(a)'. \forall f \in C_b(E), f \geq 0, \text{ we have } \int_0^T \langle \mu(t), f \rangle dt \geq 0,$$

$$(b)'. \forall 0 \leq s \leq T, \int_0^s \langle \mu(t), 1 \rangle dt = s.$$

$$(c)'. \lim_{n \rightarrow \infty} \int_0^T \langle \mu(t), \chi_{E_n} \rangle dt = 0, \text{ for } n \rightarrow \infty.$$

For any $k \geq 1, f \in C_b(E)$ let's define

$$F_f^k(\mu(\cdot)) = \langle \mu(\cdot), f \rangle \wedge k = \left(\int_0^T \langle \mu(t), f \rangle dt \right) \wedge k;$$

$$F_s^k(\mu(\cdot)) = \left(\int_0^s \langle \mu(t), 1 \rangle dt \right) \wedge k;$$

$$F_n^k(\mu(\cdot)) = \left(\int_0^T \langle \mu(t), \chi_{E_n} \rangle dt \right) \wedge k.$$

It's easy to see that F_f^k, F_s^k and F_n^k are all bounded $\|\cdot\|_\infty$ -continuous on Λ_{X^*} . Let $G_1(t) = f, G_2(t) = 1\chi_{[0,s]}(t)$ and $G_3(t) = \chi_{E_n}$. Then $G_1(\cdot), G_3(\cdot) \in C_X, G_2(\cdot) \in L_X^1$ and $\sup\{\|\|G_i(t)\|\|_\varphi; t \in [0, T], i = 1, 2, 3\}$ is finite. And also we have

$$F_f^k(\mu(\cdot)) = \langle \mu(\cdot), G_1 \rangle \wedge k;$$

$$F_s^k(\mu(\cdot)) = \langle \mu(\cdot), G_2 \rangle \wedge k;$$

$$F_n^k(\mu(\cdot)) = \langle \mu(\cdot), G_3 \rangle \wedge k.$$

By lemma 6.2.1, we can easily see that G_i can be approximated in $\|\cdot\|_1$ norm by element of C for $i = 1, 2, 3$. Thus F_j^k, F_s^k and F_n^k are elements of $\tilde{\mathcal{F}}$.

$\forall f \geq 0, \gamma < 0$ we have $H(\gamma F_j^k) \leq 0$. Thus

$$\sup_{\gamma < 0, k \geq 1} \gamma F_j^k(\mu(\cdot)) \leq \sup_{\gamma < 0, k \geq 1} \{\gamma F_j^k(\mu(\cdot)) - H(\gamma F_j^k(\cdot))\} \leq \tilde{L}(\mu(\cdot)) < \infty.$$

which implies $F_j^k(\mu(\cdot)) \geq 0$. Since k is arbitrary, (a)' is true.

Since $\forall k > T, H(\gamma F_s^k) = \gamma s$, we have

$$\sup_{\gamma \in \mathbb{R}} \gamma(F_s^k(\mu(\cdot)) - s) \leq \tilde{L}(\mu(\cdot)) < \infty.$$

which combining with the fact that $F_s^k \rightarrow F_s, k \rightarrow \infty$ we simply get (b)'.

By Hölder's inequality, (2.1) implies

$$\lim_{n \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \log E^{P_N} \exp(\lambda \sum_{i=1}^N \int_0^T \chi_{\{x_i^{(N)} \geq n\}} dt) = 0, \\ \forall \lambda \geq 0.$$

which is equivalent to: $\lim_{n \rightarrow \infty} H(\lambda F_n) = 0, \forall \lambda \geq 0$. Where $F_n = \lim_{k \rightarrow \infty} F_n^k$. Since $\forall k \geq 1, F_n^k \leq F_n$ we actually have that for any fixed $k \lim_{n \rightarrow \infty} H(\lambda F_n^k) = 0, \forall \lambda \geq 0$. Thus $\forall \lambda \geq 0$ we have

$$\lambda \limsup_{n \rightarrow \infty} F_n^k(\mu(\cdot)) = \limsup_{n \rightarrow \infty} \{\lambda F_n^k(\mu(\cdot)) - H(\lambda F_n^k)\} \leq \tilde{L}(\mu(\cdot)) < \infty.$$

which implies $\lim_{n \rightarrow \infty} F_n^k(\mu(\cdot)) = 0$. Since

$$|F_n^k(\mu(\cdot))| \leq |\mu|_\infty \|\chi_{E_n}\|_\varphi < \infty.$$

Taking k bigger enough we get (c)'.

By the right continuity of $\mu(\cdot)$ and (a)', (b)' and (c)' we conclude that for every $t \in [0, T], \mu(t)$ satisfies conditions (a), (b) and (c) of lemma 6.2.9. Thus $\mu(\cdot) \in \mathcal{D}_\infty$.
□

Now we are ready to prove theorem 6.2.6.

Proof of theorem 6.2.6: By lemma 6.2.10 and applying corollary 6.1.5 to $X_0 = \mathcal{D}_\infty^{nd}, X = \Lambda_{X^*}, d_1 = \gamma, d_2 = \|\cdot\|_\infty, \mathcal{C}, \mathcal{B}_1$ and $\mathcal{B}$ we get the result. □

Chapter 7

Upper Bound And Exponential Tightness

In chapter 6 we tried to find some technically simple conditions for getting the large deviation upper bound. Usually, from a weak large deviation system to a large deviation system we need the following exponential tightness condition: $\forall a < \infty$, there exists a compact set $K_a \subset X$ such that

$$(0.2) \quad \limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_\varepsilon(K_a^c) \leq -a.$$

But in many cases it is extremely difficult to check this condition. Hence, we are led to a natural question: Under what circumstance is the exponential tightness necessary for getting the large upper bound? In this chapter we will discuss some situation where the exponential tightness is necessary.

Let X be a topological space. In general a lower semicontinuous mapping $I : X \rightarrow [0, \infty]$ is called a *rate function*. A rate function I is called a *good rate function* if the level sets $\{x \in X : I(x) \leq s\}$ are compact for all $s \geq 0$. The following is an important property of good rate function.

Lemma 7.0.1 *Let I be a good rate function.*

(i). *Let $\{F_n\}_{n \geq 1}$ be a sequence of closed subsets of X satisfying $F_n \supset F_{n+1}$, $\forall n \geq 1$. Let $F_0 = \bigcap_{n=1}^{\infty} F_n$. Then*

$$(0.3) \quad \inf_{x \in F_0} I(x) = \lim_{n \rightarrow \infty} \inf_{x \in F_n} I(x).$$

(ii). Suppose X is a metric space with metric d . Then for any set A

$$(0.4) \quad \inf_{x \in A} I(x) = \lim_{n \rightarrow \infty} \inf_{x \in A_n} I(x),$$

where A denotes the closure of A and A_n is defined as

$$A_n = \{x \in X : d(x, A) = \inf_{y \in A} d(x, y) \leq 1/n\}.$$

Proof: (i). Since $F_0 \subset F_n$, $\forall n \geq 1$, it suffices to prove that for all $\varepsilon > 0$

$$\gamma = \lim_{n \rightarrow \infty} \inf_{x \in F_n} I(x) \geq \inf_{x \in F_0} I(x) - \varepsilon.$$

This is obviously when $\gamma = \infty$. When $\gamma < \infty$, fix ε and let $s = \varepsilon + \gamma$. The sets $G_n = F_n \cap \Phi(s)$ are non-empty, compact and decreasing. Hence

$$G_0 = F_0 \cap \Phi(s)$$

is also non-empty and thus the above inequality follows.

(ii). This is a direct result of (i).. $\square$

Note. This lemma and its proof can be found in lemma 4.1.6 of [12].

The following theorem gives the exact relation between the exponential tightness and the large deviation upper bound and was shown in [35].

Theorem 7.0.2 *Let X be a Polish space with metric d . $\{P_n\}_{n \geq 1}$ is a sequence of probability measures on X satisfying the large deviation upper bound with a good rate function I . Then $\{P_n\}_{n \geq 1}$ is exponential tight on X .*

Proof: Let $\{x_i\}$ be any countable dense subset of X . For any $k \geq 1$ let $B(x, 1/k)$ be the ball with center x and radius $1/k$. For any $n \geq 1$, let $F_n = (\cup_{i=1}^n B(x_i, 1/k))^c$. Then $\{F_n\}$ is a decreasing sequence of closed subset of X with intersection $F_0 = \emptyset$. Applying (0.3) and the large deviation upper bound we get that for any $s > 0$ there exists an n_k such that

$$(0.5) \quad P_N(F_{n_k}) \leq \exp[-Nks].$$

Let

$$K_s = \cap_{k=1}^{\infty} H_{n_k}, \quad H_n = \cup_{i=1}^n B(x_i, 1/k).$$

Then the totally bounded subset K_s is precompact because X is Polish. By (0.5) we have

$$(0.6) \quad P_N((K_s)^c) \leq \sum_{k=1}^{\infty} P_N(F_{n_k}) \leq \frac{\exp[-Ns]}{1 - \exp[-Ns]}$$

which implies the exponential tightness of $\{P_N\}$. $\square$

The new result of this chapter is a generalization of theorem 7.0.2 from Polish space to a space which is not Polish.

Let X be a Polish space with metric d . X_n , $n = 1, \dots$, is a increasing sequence of closed subset of X , i.e., $X_n \subset X_{n+1}$. Then each X_n with subspace topology of X is also Polish. Let

$$X_{\infty} = \bigcup_{n=1}^{\infty} X_n$$

equipped with the inductive topology. X_{∞} is not Polish in general, is a closed subset of X . We then have

Theorem 7.0.3 *If $\{P_N\}_{N \geq 1}$ is a sequence of probability measures on space X , satisfying a large deviation upper bound with a good rate function I and for any $s > 0$ there exists an n_0 such that for any $n \geq n_0$*

$$(0.7) \quad \limsup_{N \rightarrow \infty} N^{-1} \log P_N((X_n)^c) \leq -s$$

then $\{P_N\}_{N \geq 1}$ is exponential tight on X_{∞} .

Proof: Let $\tilde{P}_N$ denote the extension of P_N to X and

$$(0.8) \quad \tilde{I}(x) = \begin{cases} I(x) & \text{for } x \in X_{\infty} \\ \infty & \text{otherwise} \end{cases}$$

Then $\{\tilde{P}_N\}$ satisfies the large deviation upper bound with the good rate function I . Thus by theorem 7.0.2 $\{\tilde{P}_N\}$ is exponential tight on X . For any $s > 0$ choose n_0 such that (0.7) holds and a compact subset $\tilde{K}$ of X such that

$$\limsup_{N \rightarrow \infty} N^{-1} \log P_N(\tilde{K}^c) \leq -s$$

Let $K = \tilde{K} \cap X_{n_0}$. Then the theorem follows from the compactness of K in X .

Note. This result can be used to the spaces discussed in chapter 5.

Appendix

Let (X, d) be an arbitrary separable metric space. $T > 0$ be fixed. $M_1(X)$ represents the space of all probability measures on X with the usual weak topology and metric ρ . $D([0, T], X)$ and $D([0, T], M_1(X))$ are the corresponding cadlag spaces with the Skorohod topology. $(M_1(D([0, T], X)), \tau_{weak})$ denotes the space of all probability measures on space $D([0, T], X)$ with weak topology. Define a map π as follow:

$$\pi : M_1(D([0, T], X)) \rightarrow D([0, T], M_1(X)), P \mapsto P \circ \omega^{-1}(\cdot)(t).$$

where $\omega(\cdot) \in D([0, T], X)$.

CLAIM: π is not continuous.

The following example was suggested by D.Dawson.

Choose $a, b \in X$ such that $d(a, b) > 0$. For any $n \geq 0$, consider

$$P_n = \frac{1}{2}\delta_{f_n(\cdot)} + \frac{1}{2}\delta_{g_n(\cdot)}$$

where

$$f_n(t) = \begin{cases} a & \text{for } 0 \leq t < T/2 - (1/n) \\ b & \text{for } T/2 - (1/n) \leq t \leq T \end{cases}$$

$$g_n(t) = \begin{cases} a & \text{for } 0 \leq t < T/2 + (1/n) \\ b & \text{for } T/2 + (1/n) \leq t \leq T \end{cases}$$

Then P_n converges weakly to $P = \delta_{f(\cdot)}$. Where

$$f(t) = \begin{cases} a & \text{for } 0 \leq t < T/2 \\ b & \text{for } T/2 \leq t \leq T \end{cases}$$

This follows since $f_n \rightarrow f$ and $g_n \rightarrow f$ in $D([0, T], X)$.

Let $t = T/2$, $t_n = T/2 + (1/2n)$, then by direct calculation we have

$$\pi(P_n)(t_n) = \frac{1}{2}\delta_a + \frac{1}{2}\delta_b, \quad \pi(P)(t) = \delta_b$$

and $\pi(P)(t-) = \delta_a$. Thus

$$\rho(\pi(P_n)(t_n), \pi(P)(t)) \wedge \rho(\pi(P_n)(t_n), \pi(P)(t-))$$

does not converge to zero. By proposition 6.5 in chapter 3 of [21] we conclude that $\pi(P_n)$ does not converge to $\pi(P)$.

Bibliography

- [1] Baldi, P.(1988). Large deviation and stochastic homogenization. *Ann. Mat. Pura. Appl.*, **151**:161-177.
- [2] Billingsley, P.(1968). *Convergence of Probability Measures*. John Wiley and Sons, New York.
- [3] Chen, M.F.(1986). *Jump Processes and Particle System*. Beijing Normal University Publishing House, Beijing. (in Chinese).
- [4] Comets, F.(1987). Nucleation for a long range magnetic model. *Ann.Inst.Henri Poincaré*, **23**:137-178.
- [5] Dawson, D.A. and Gärtner, J.(1986). Large deviations and tunnelling for particle systems with mean-field interaction. *C.R. Math. Rep. Acad. Sci. Canada*, **8**:387-392.
- [6] Dawson, D.A. and Gärtner, J.(1987). Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, **20**:247-308.
- [7] Dawson, D.A. and Gärtner, J.(1988). Long-time behavior of interacting diffusions. In: Norris, J. R. (ed.) *Stochastic calculus in application*. Proc. Cambridge Symp. 1987, pp. 29-54. Longman.
- [8] Dawson, D.A. and Gärtner, J.(1989). Large deviations, free energy functional and quasi-potential for a mean field model of interacting diffusions. *Mem. Amer. Math. Soc*, **78**, No.398.

- [9] Dawson, D.A. and Zheng, X.(1991) . Law of large numbers and a central limit theorem for unbounded jump mean-field models. *Advances in Applied Mathematics*, **12**:293-326.
- [10] Dawson, D.A.(1983). Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *J. Statist. Phys.*, **31**:29-85.
- [11] Day, M.V.(1990). Large deviations results for the exit problem with characteristic boundaries. *J. Math Anal. and appl.*, **147**:134-153.
- [12] Dembo, A. and Zeitouni, O.(1993). *Large deviations and Applications*. Jones and Bartlett Publishers, Boston.
- [13] Deuschel, J-D. and Stroock, D.W.(1989). *Large deviations*. Pure and Applied Mathematics, **137**. Academic Press, Boston.
- [14] Donsker, M.D. and Varadhan, S.R.S.(1975) . Asymptotic evaluation of certain Markov process expectations for large time, I. *Comm.Pure Appl.Math.*, **28**:1-47.
- [15] Donsker, M.D. and Varadhan, S.R.S.(1975) . Asymptotic evaluation of certain Markov process expectations for large time, II. *Comm.Pure Appl.Math.*, **28**:279-301.
- [16] Donsker, M.D. and Varadhan, S.R.S.(1976) . Asymptotic evaluation of certain Markov process expectations for large time, III. *Comm.Pure Appl.Math.*, **29**:389-461.
- [17] Donsker, M.D. and Varadhan, S.R.S.(1983) . Asymptotic evaluation of certain Markov process expectations for large time, IV. *Comm.Pure Appl.Math.*, **36**:183-212.
- [18] Dudley, R.M.(1976). *Probabilities and Metrics: Convergence of laws on metric spaces, with a view to statistical testing*. Aarhus: Aarhus University. (Mathematics Institute Lecture Note Series 45.).
- [19] Dunford, N. and Schwartz, J.T.(1967). *Linear Operators, Vol.1*. Interscience, New York.

- [20] Edwards, R.(1965) . *Functional Analysis*. Holt-Rienhart-Winston.
- [21] Ethier, S.N. and Kurtz, T.G.(1985) . *Markov Processes: Characterization and Convergence*. Wiley.
- [22] Feng, S. and Zheng, X.(1992). Solutions of a class of nonlinear master equations. *Stochastic Processes and Their Applications*, **43**: 65-84.
- [23] Feng, S.(1992). Large deviations for Markov processes with mean field interaction and unbounded jumps. *C.R. Math. Rep. Acad. Sci. Canada*, XIV, No.1:37-42.
- [24] Freidlin, M.I. and Wentzell, A.D.(1984). *Random Perturbations of Dynamical Systems*. Springer-Verlag, New York.
- [25] Funaki, T.(1984). A certain class of diffusion processes associated with nonlinear parabolic equations. *Z. Wahrsch. Verw. Gebiete*, **67**:331-348.
- [26] Gärtner, J.(1988). On the McKean-Vlasov limit for interacting diffusions. *Math. Nachr.*, **137**:197-248.
- [27] Gihman, I.I. and Skorohod, A.V.(1974). *The Theory of Stochastic Processes, Vol.1*. Springer-Verlag, Berlin. (English translation).
- [28] Holley, R. and Stroock, D.W.(1976). A martingale approach to infinite system of interacting particles. *Ann. Probab.*, **4**:195-228.
- [29] Ikeda, N. and Watanabe, S.(1981). *Stochastic Differential Equations and Diffusion Processes*. Amsterdam.
- [30] Lanford, O.E.(1973). Entropy and equilibrium states in classical statistical mechanics. in A. Lenard, editor, *Statistical Mechanics and Mathematical Problems*.
- [31] Léonard, C.(1990). Large deviations in the dual of a normed vector space. *Prépublication de l'Université d'Orsay*.
- [32] Léonard, C.(1990). On large deviations for particle systems associated with spatially homogeneous Boltzmann type equations. preprint.

- [33] Meyer, P.A.(1966). *Probability and Potential*. Blaisdell Publishing Company, Waltham.
- [34] Nicolis,G. and Prigogine,I.(1977). *Self-organization in Non-equilibrium Systems*. Wiley-Interscience,New York.
- [35] Pukhalskii, A.A.(1991). On functional principle of large deviations. in V. Sazonov and T. Shervashidze, editors, *New trends in probability and statistics*. VSP Moks'las, Moskva.
- [36] Reed, M.and Simon, B.(1972). *Methods of Modern Mathematical Physics, Vol.1*. Academic Press, New York.
- [37] Schlögl, F.(1972). Chemical reaction models for non-equilibrium phase transitions. *Z.Physik*, **253**:147-161.
- [38] Shiga, T. and Tanaka, H.(1985). Central limit theorem for a system of Markovian particles with mean-field interactions. *Zeit.fur.Wahr.*, **69**:439-459.
- [39] Sowers, R.(1992). Large deviations for a reaction-diffusion equation with non-Gaussian perturbations. *Ann. Probab.*, **20**:504-537.
- [40] Stroock, D.W. and Varadhan, S.R.S.(1979). *Multidimensional Diffusion Processes*. Springer-Verlag, Berlin.
- [41] Stroock, D.W.(1984). *An Introduction to the Theory of Large Deviations*. Springer-Verlag, Berlin.
- [42] Sugiura, M.(1990). Large deviations for Markov processes of jump type with mean-field interactions. preprint.
- [43] Varadhan, S.R.S.(1984). *Large deviations and applications*. Society for Industrial and Applied Mathematics, Philadelphia.
- [44] Zheng, J. and Zheng, X.(1987). A martingale approach to q-process. *Chinese Kezue Tongbao*.

END

1 3-1 2-9 3

FIN