SOME ASYMPTOTIC RESULTS FOR WEAKLY INTERACTING PARTICLE SYSTEMS

Ruoyu Wu

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

Chapel Hill 2016

Approved by:
Amarjit Budhiraja
Shankar Bhamidi
Edward Carlstein
Vladas Pipiras
Kai Zhang
ABSTRACT

RUOYU WU: Some Asymptotic Results for Weakly Interacting Particle Systems
(Under the direction of Amarjit Budhiraja)

Weakly interacting particle systems have been widely used as models in many areas, including, but not limited to, communication systems, mathematical finance, chemical and biological systems, and social sciences. In this dissertation, we establish law of large numbers (LLN), central limit theorems (CLT), large deviation principles (LDP) and moderate deviation principles (MDP) for several types of such systems.

The dissertation consists of two parts. In the first part, we are mainly concerned with LLN and CLT. We begin by studying weakly interacting multi-type particle systems that arise in neurosciences to model a network of interacting spiking neurons. We prove a CLT showing the centered and suitably normalized empirical measures converge in distribution to a Gaussian random field. This result can in particular be applied to single-type systems to characterize the joint asymptotic behavior of large disjoint subpopulations. We then establish a CLT for a multi-type model where each particle is affected by a common source of noise. Here the limit is not Gaussian but rather described through a suitable Gaussian mixture. Next, we consider weakly interacting particle systems in a setting where not every pair of particles interacts, but rather particle interactions are governed by Erdős–Rényi random graphs and an interaction between a pair of particles occurs only when there is a corresponding edge in the graph. Edges can form and break down independently as time evolves. We prove a LLN and CLT under conditions on the edge probabilities.

The second part of this dissertation concerns MDP and LDP for certain weakly interacting particle systems. We study interacting systems of both diffusions and of pure jump Markov processes with a countable state space. We are interested in estimating probabilities of moderate deviations of empirical measure processes, from the LLN limit. For both systems a MDP is established which is formulated in terms of a LDP with an appropriate speed function, for suitably centered and normalized empirical measure processes. Finally, we study particle approximations for certain
nonlinear heat equations using a system of Brownian motions with killing. A LLN and LDP for
sub-probability measure valued processes given as the empirical measure of the alive Brownian
particles are proved. We also give, as a byproduct, a convenient variational representation for ex-
pectations of nonnegative functionals of Brownian motions along with an i.i.d. sequence of random
variables.
ACKNOWLEDGEMENTS

I would like to express my gratitude to all who helped me in completing this dissertation and along my PhD. journey.

First, I extend my deepest gratitude to my advisor, Dr. Amarjit Budhiraja, for his supervision, encouragement and patience. I would like to thank him for numerous pleasant and fruitful discussions, rigorous training, insightful guidance and helpful career suggestions during my PhD. study. It was Dr. Budhiraja who raised me up and gave me the confidence to walk the path of probability. I am grateful to him for his great patience in checking over all the technical details, grammar, and flow of ideas in countless revisions of each work. His strong memory and sharp insight are always key whenever I get stuck on research projects. Finally, I would like to thank him for his financial support, without which I would not be able to fully concentrate on my study and research.

Also, I would like to acknowledge my committee members. I am especially grateful to Dr. Shankar Bhamidi, who provided many valuable comments and lots of guidance, and encouraged me with cool stories and meaningful words. I would like to thank Dr. Kai Zhang for his helpful suggestions and feedback as well. I am indebted to Dr. Vladas Pipiras, who taught me the fundamentals and gave wonderful jokes both inside and outside the classroom. I am also grateful to Dr. Edward Carlstein for suggestions and guidance on my research and teaching.

I would like to express my gratitude to all professors who have taught me, for their discussions and lectures on various topics in probability and statistics, which helped build up my foundation and broaden my interest. My sincere thanks go to Dr. Ross Leadbetter for his wisdom, rigorous training and great sense of humor; to Dr. Andrew Nobel for his enlightening courses and discussions; to Dr. Mark Williams for his lucid explanations of core mathematical courses; and to Dr. Dayue Chen at Peking University and Dr. Haipeng Shen for their countless suggestions and career advice.

I am also thankful to staff in the department for their help and support during my graduate study, especially Ms. Alison Kieber, for her generous help and sunny smile.
I am very grateful to all my friends whose support and care helped me overcome setbacks, keep up a good mood, and stay focused on my graduate study. In particular, my gratitude especially goes to Dr. Xuan Wang, for numerous exciting discussions, extremely strong support and tons of selfless help. I would like to thank Dr. Jiang Chen, Dr. Xuan Wang, Dr. Haojin Zhai, Mr. Jingxiang Chen and Ms. Shuang Lu for wonderful memories at board-game nights. I am also thankful to my officemates, Dr. Wen Jenny Shi, Mr. Hyowon An, Ms. Liuqing Yang and Ms. Yang Yu, for delicious snacks, happy times and meaningful conversations. I want to thank Mr. Jimmy Jin and Mr. Eric Friedlander for great times together and patience with all my confusions about English. My gratitude also goes to Ms. Qunqun Yu, Dr. Dehan Kong, Mr. Gan Liu and Mr. Xi Chen for strong help and advice on my graduate life.

Most importantly, none of this would have been possible without the love and support from my family. I would like to especially express my heart-felt gratitude to my wife Ye Zhao, for her understanding, encouragement, love, care, support and thousands of delicious meals. Lastly I owe my parents a great deal, who have been a constant source of love, care, support and strength all these years.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>PRELIMINARIES</td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>Symmetric statistics and multiple Wiener integrals</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Large deviation and moderate deviation principles</td>
<td>11</td>
</tr>
<tr>
<td>2.3</td>
<td>A variational representation for functionals of Brownian motions</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>WEAKLY INTERACTING MULTI-TYPE PARTICLE SYSTEMS</td>
<td>14</td>
</tr>
<tr>
<td>3.1</td>
<td>Model</td>
<td>16</td>
</tr>
<tr>
<td>3.2</td>
<td>Fluctuations for multi-type particle system</td>
<td>18</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Central limit theorem</td>
<td>18</td>
</tr>
<tr>
<td>3.2.2</td>
<td>An application to single-type particle system</td>
<td>20</td>
</tr>
<tr>
<td>3.2.3</td>
<td>A limit theorem for statistics of multi-type populations</td>
<td>22</td>
</tr>
<tr>
<td>3.3</td>
<td>Proof of Lemma 3.3</td>
<td>24</td>
</tr>
<tr>
<td>3.4</td>
<td>Proof of Theorem 3.4</td>
<td>25</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Asymptotics of statistics of multi-type populations of independent particles</td>
<td>26</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Asymptotics of $J^N$</td>
<td>32</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Completing the proof of Theorem 3.4</td>
<td>36</td>
</tr>
<tr>
<td>3.5</td>
<td>Proof of Theorem 3.5</td>
<td>37</td>
</tr>
<tr>
<td>4</td>
<td>WEAKLY INTERACTING MULTI-TYPE PARTICLE SYSTEMS WITH A COMMON FACTOR</td>
<td>38</td>
</tr>
<tr>
<td>4.1</td>
<td>Model</td>
<td>39</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Well-posedness</td>
<td>40</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Central limit theorem</td>
<td>41</td>
</tr>
</tbody>
</table>
6.2.2 Tightness of $\tilde{Y}^m$ ......................................................... 106
6.2.3 Laplace upper bound ....................................................... 108
6.2.4 Laplace lower bound ....................................................... 112
6.2.5 $I$ is a rate function ....................................................... 116

7 MODERATE DEVIATION PRINCIPLES FOR WEAKLY INTERACTING
PURE JUMP MARKOV PROCESSES ........................................... 119
7.1 Model and main results ...................................................... 120
7.1.1 Model and law of large numbers ...................................... 120
7.1.2 Moderate deviation principle .......................................... 123
7.2 Proofs for the pure jump case ............................................. 125
7.2.1 Proof of Theorem 7.4 ..................................................... 125
7.2.2 Proof of Theorem 7.5 ..................................................... 137
7.3 Proofs of Theorem 7.2 and Remark 7.2(iii) ............................ 139

8 LARGE DEVIATIONS FOR PARTICLE APPROXIMATION OF A NONLIN-
EAR HEAT EQUATION BY BROWNIAN MOTIONS WITH KILLING ............... 143
8.1 Introduction ................................................................. 143
8.2 Model and notation ......................................................... 144
8.3 Large deviation principle .................................................. 145
8.3.1 Canonical space and processes ...................................... 145
8.3.2 Rate function and statement of the LDP .......................... 145
8.4 A general variational representation formula .......................... 148
8.5 Variational representation and tightness properties ................. 159
8.6 Laplace upper bound ....................................................... 169
8.7 Laplace lower bound ....................................................... 170
8.8 $I$ is a rate function ....................................................... 172
8.9 Sketch of proof of Theorem 8.1 ......................................... 176

REFERENCES ................................................................. 177
LIST OF ABBREVIATIONS AND SYMBOLS

BM Brownian motion
CLT Central limit theorem
MDP Moderate deviation principle
MWI Multiple Wiener integral
LDP Large deviation principle
LLN Law of large numbers
ODE Ordinary differential equation
PDE Partial differential equation
POC Propagation of chaos
PRM Poisson random measure
RCLL Right continuous functions (stochastic processes) with left limits
$\mathbb{N}$ Set of natural numbers
$\mathbb{N}_0$ Set of non-negative integers
$\mathbb{R}$ Set of real numbers
$\mathbb{R}_+$ Set of non-negative real numbers
$\mathbb{R}^d$ Set of $d$-dimensional real vectors
$\mathbb{Z}$ Set of integers
$\mathcal{B}(\mathcal{E})$ Borel $\sigma$-field on a topological space $\mathcal{E}$
$\mathcal{C}_b(\mathcal{S})$ The collection of bounded continuous functions from space $\mathcal{S}$ to $\mathbb{R}$
$\mathcal{M}_b(\mathcal{S})$ The collection of bounded $\mathcal{B}(\mathcal{S})/\mathcal{B}(\mathbb{R})$-measurable maps
$\mathcal{P}(\mathcal{S})$ Collection of all probability measures on space $\mathcal{S}$
$\mathcal{M}_{FC}(\mathcal{S})$ Collection of all measures which are finite on compacts on space $\mathcal{S}$
$\mathcal{C}([0, T] : \mathcal{S})$ The space of continuous functions from $[0, T]$ to $\mathcal{S}$
$\mathcal{D}([0, T] : \mathcal{S})$ The space of right continuous functions with left limits from $[0, T]$ to $\mathcal{S}$
$\|f\|_\infty$ The supremum norm of $f : \mathcal{S} \to \mathbb{R}^d$ (see Page 8)
$\|f\|_{BL}$ The bounded Lipschitz norm of $f : \mathcal{S} \to \mathbb{R}^d$ (see Page 8)
$\mathcal{L}(X)$ Probability law of a random variable $X$
$X_n \Rightarrow X$ Weak convergence of random variables with values in some metric space
$\mu_n \to \mu$  Weak convergence of probability measures $\mu_n$ to $\mu$

$L^2(S, \nu, H)$  The space of measurable functions $f: S \to H$ such that $\int_S \|f(x)\|_H^2 \nu(dx) < \infty$, where $\| \cdot \|_H$ is the norm on $H$, for a measure $\nu$ on $S$ and a Hilbert space $H$

$x_t$ and $x(t)$  The evaluation of $x$ at $t \in [0, T]$ for a function $x: [0, T] \to S$, and similar convention will be used for stochastic processes

$l_2$  Space of square-summable sequences
CHAPTER 1
INTRODUCTION

This dissertation concerns certain asymptotic results, including law of large numbers (LLN), central limit theorems (CLT), large deviation principles (LDP) and moderate deviation principles (MDP), for stochastic processes that describe the evolution of a collection of weakly interacting particles. The study of such systems dates back to the classical works of Boltzmann, Vlasov, McKean and others, and has a long history (see [67, 46] and references therein). Over the past few decades, such models have arisen in many different areas, including communication systems (e.g. loss network models [33, 1, 69], random medium access protocols [37, 5]), mathematical finance (e.g. mean field games [50, 40, 38, 19], default clustering in large portfolios [58, 59, 34]), chemical and biological systems (e.g. biological aggregation, chemotactic response dynamics [64, 31, 61, 16]), social sciences (e.g. opinion dynamics models [21, 35]). A typical such system is given as a Markov process $Z^N = (Z^1_N, \ldots, Z^N_N)$ with values in $\mathbb{S}^N$, where $\mathbb{S}$ is some Polish space, $N$ is the number of particles in the system and $Z^i_N$ represents the state process for the $i$-th particle in the system. The evolution equation for $Z^i_N$ depends on not only the state of $Z^i_t$, at any given time instant, but also on the empirical measure of all particles at that moment, namely $\frac{1}{N} \sum_{i=1}^N \delta_{Z^i_t}$, where $\delta_x$ is the Dirac probability measure at point $x$. As the number of particles grows, the contribution of any given particle to the empirical measure becomes small (of order $\frac{1}{N}$), and in that sense the interaction between any two particles is weak.

To motivate the questions of interest, we will begin with a basic example of weakly interacting diffusions of the following form (cf. [53], see also [67]), and review some well known asymptotic results. Suppose $b: \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is a bounded Lipschitz function. Let $\{W^i\}_{i \in \mathbb{N}}$ be i.i.d. $d$-dimensional standard $\{\mathcal{F}_t\}$-Brownian motions (BM) given on some filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\})$. Then fixing a time horizon $[0, T]$, $W^i$ can be regarded as a random variable with values in $C_d = C([0, T] : \mathbb{R}^d)$, the space of continuous functions from $[0, T]$ to $\mathbb{R}^d$, equipped with the uniform topology. Let $\{Z^i_0\}_{i \in \mathbb{N}}$ be i.i.d. $d$-dimensional $\mathcal{F}_0$-measurable random variables with
common distribution $\mu_0$. For $N \in \mathbb{N}$, let $\{Z_i^{i,N}\}_{i=1}^N$ be given as the unique solution of the following system of stochastic differential equations (SDE):

$$Z_i^{i,N}_t = Z_0^i + \frac{1}{N} \sum_{j=1}^N \int_0^t b(Z_i^{i,N}_s, Z_j^{j,N}_s) \, ds + W^i_t, \quad i = 1, \ldots, N, \quad t \in [0, T]. \quad (1.1)$$

It can be shown (cf. [67], Section I.1) that as $N \to \infty$, each $Z_i^{i,N}$ converges in distribution to $X_i$ given as the unique solution to the following SDE:

$$X_i^t = Z_0^i + \int_0^t \left( \int_{\mathbb{R}^d} b(X_i^s, y) \mu_s(dy) \right) ds + W^i_t, \quad (1.2)$$

where $\mu_t = \mathcal{L}(X_i^t)$ for $t \in [0, T]$. Since $\{Z_0^i\}_{i \in \mathbb{N}}$ are i.i.d., so are $\{X_i\}_{i \in \mathbb{N}}$. Since the evolution of the process $X_i$ depends on its own probability law $\mu = \mathcal{L}(X_i)$, the equations describing the evolution of the probability distributions (i.e. the Kolmogorov Forward equations and the Fokker–Planck equations for the probability densities) are nonlinear and for this reason the process is referred to as a nonlinear diffusion (or more generally a nonlinear Markov process).

The following moment estimate holds (see e.g. [67], Theorem 1.4).

**Theorem 1.1.** For any $i \in \mathbb{N}$ and $T \geq 0$,

$$\sup_{N \geq 1} \sqrt{N} \mathbb{E} \left( \sup_{t \leq T} \|Z_i^{i,N}_t - X_i^t\| \right) < \infty,$$

where $\| \cdot \|$ is the Euclidean norm on $\mathbb{R}^d$.

It is easy to deduce from Theorem 1.1 that for any $T \geq 0$ and $k \geq 1$, $(Z_1^{1,N}, \ldots, Z_k^{k,N})$ converges to $(X^1, \ldots, X^k)$ in probability as $N \to \infty$, in $C_T$. Since the latter vector is i.i.d., the result says that the (asymptotic) independence of the law of any $k$-particles at time zero, propagates to future time instants, and for this reason this result is referred to as the *Propagation of Chaos* (POC). We will see in Proposition 1.3 that such a property is equivalent to a certain law of large numbers result. We begin with some notation. For a metric space $\mathcal{E}$, let $C_b(\mathcal{E})$ denote the space of real continuous and bounded functions on $\mathcal{E}$. For a measure $\nu$ on $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$ and a $\nu$-integrable function $f : \mathcal{E} \to \mathbb{R}$, we denote $\int f \, d\nu$ as $\langle f, \nu \rangle$. Denote by $\mathcal{P}(\mathcal{E})$ the space of probability measures on $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$, equipped with the usual topology of weak convergence.
**Definition 1.2.** Let $\mathcal{E}$ be a separable metric space and $u_N$ be a sequence of exchangeable probability measures on $\mathcal{E}^N$. We say that $u_N$ is $u$-chaotic, where $u$ is a probability measure on $\mathcal{E}$, if for all $\phi_1, \ldots, \phi_k \in \mathcal{C}_b(\mathcal{E})$ and $k \geq 1$,

$$
\lim_{N \to \infty} \langle \phi_1 \otimes \cdots \otimes \phi_k \otimes 1 \otimes \cdots \otimes 1, u_N \rangle = \prod_{i=1}^{k} \langle \phi_i, u \rangle.
$$

Note that since $\{Z_i^0\}$ are i.i.d. and the same function $b$ appears in the evolution equation of each $Z_i^N$, the probability law $\nu^N$ of $(Z_1^N, \ldots, Z_k^N)$ is exchangeable and thus according to the above definition and Theorem 1.1, $\nu^N$ is $\mu$-chaotic, where $\mu = \mathcal{L}(X^i)$. The next proposition says that the property of POC is equivalent to a certain LLN result.

**Proposition 1.3.** Let $u_N$ be a sequence of exchangeable probability measures on $\mathcal{E}^N$. Denote by $Y_i : \mathcal{E}^N \to \mathcal{E}$ the canonical maps $Y_i(y) = y_i$ for $y = (y_1, \ldots, y_N) \in \mathcal{E}^N$. Then $u_N$ is $u$-chaotic for some $u \in \mathcal{P}(\mathcal{E})$ if and only if $\theta_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{Y_i}$ converges in probability, in $\mathcal{P}(\mathcal{E})$, to $u$.

Proposition 1.3 in particular says that $\mu^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{Z_i^N}$ converges in probability to $\mu$, in $\mathcal{P}(\mathcal{C}_d)$. It is then natural to study the fluctuations of $\mu^N$ from $\mu$. Although $\mu^N$ is not the empirical measure of i.i.d. random variables, the fact that the interaction is weak, which is also reflected in the asymptotic independence result noted below Theorem 1.1, suggests that $\sqrt{N} \langle \phi, \mu^N - \mu \rangle$ should satisfy a CLT for suitable $\phi : \mathcal{C}_d \to \mathbb{R}$. Indeed, the classical works of Sznitman [67] and Shiga–Tanaka [62] show that such a result holds, and the limiting variance suitably accounts for the dependence between the particles. We now briefly describe this result. We start with some notation.

Fix $T > 0$ and let $\Omega_d = \mathcal{C}_d \times \mathcal{C}_d$. For $\omega = (\omega_1, \omega_2) \in \Omega_d$, define canonical processes $(W_s, X_s)$ as $(W_s(\omega), X_s(\omega)) = (\omega_1, \omega_2)$. Let $\nu = \mathcal{L}(W^i, X^i) \in \mathcal{P}(\Omega_d)$, where $X^i$ and $W^i$ are as in (1.2). Note that $\nu(2) = \mu$, where $\nu(2)$ denotes the marginal distribution on the second coordinate. With $\mu_t = \mathcal{L}(X_t^i)$, define $b_t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ by

$$
b_t(x,y) = b(x,y) - \int_{\mathbb{R}^d} b(x,z) \mu_t(dz), \quad x,y \in \mathbb{R}^d.
$$
In addition, define function $h: \Omega_d \times \Omega_d \to \mathbb{R} \ (\nu \ \text{a.s.})$ as 

$$h(\omega, \omega') = \int_0^T b_t(X_{s,t}(\omega), X_{s,t}(\omega')) \cdot dW_{s,t}(\omega), \ (\omega, \omega') \in \Omega_d \times \Omega_d.$$ 

Consider the Hilbert space $L^2(\Omega_d, \nu)$, on which the integral operator $A$ is defined by

$$Af(\omega) = \int_{\Omega_d} h(\omega', \omega) f(\omega') \nu(d\omega'), \ f \in L^2(\Omega_d, \nu), \ \omega \in \Omega_d.$$ 

Denote by $I$ the identity operator on $L^2(\Omega_d, \nu)$. For $\phi \in L^2(C_d, \mu)$, let

$$\xi_N^N(\phi) = \frac{1}{\sqrt{N}} \sum_{i=1}^N (\phi(Z_{i,N}) - \langle \phi, \mu \rangle)$$

and $\phi \equiv \phi(X_\ast) - \langle \phi, \mu \rangle \in L^2(\Omega_d, \nu)$. Then the following theorem holds (cf. [66, 62]).

**Theorem 1.4.** The operator $I - A$ is invertible. The collection $\{\xi_N^N(\phi) : \phi \in L^2(C_d, \mu)\}$ converges as $N \to \infty$ to a mean zero Gaussian field $\{\xi(\phi) : \phi \in L^2(C_d, \mu)\}$ in the sense of convergence of finite dimensional distributions, where for $\phi, \psi \in L^2(C_d, \mu)$,

$$E[\xi(\phi)\xi(\psi)] = \langle (I - A)^{-1} \phi, (I - A)^{-1} \psi \rangle_{L^2(\Omega_d, \nu)}.$$ 

**Remark 1.1.** (i) Note that when $b \equiv 0$, $\{Z_{i,N}\}_{i=1}^N$ are i.i.d. and $A = 0$. In this case the above theorem reduces to the classical CLT for i.i.d. random variables.

(ii) Theorem 1.4 in particular says that the random variable $\frac{1}{\sqrt{N}} \sum_{i=1}^N \phi(Z_{i,N})$ converges in probability to $\langle \phi, \mu \rangle$ for all $\phi \in L^2(C_d, \mu)$. Note that the propagation of chaos result below Theorem 1.1 only implied this convergence in probability for $\phi \in C_b(C_d)$.

**New Central Limit Results.** In this dissertation we will study several variants of the above central limit type result for different models in Chapters 3, 4 and 5. In Chapter 3, CLT for weakly interacting multi-type particle systems is established. Next, in Chapter 4, a model with a common source of noise and more general coefficients is studied. Chapter 5 proves POC and CLT in a setting where not every pair of particles interacts, but rather particle interactions are governed by Erdős–Rényi random graphs and an interaction between a pair of particles occurs only when there is a corresponding edge in the graph.
In addition to LLN and CLT results we will study the tail behavior of \( \mu^N \) through certain large and moderate deviation principles. In order to motivate the problem of interest we begin with a simple example. Let \( \{Y_i\}_{i \in \mathbb{N}} \) be a sequence of i.i.d. real-valued mean zero random variables with common probability law \( \rho \in \mathcal{P}(\mathbb{R}) \). A LDP for \( \bar{Y}_N = \frac{1}{N} \sum_{i=1}^{N} Y_i \) under suitable integrability conditions will formally say that for \( c > 0 \),

\[
P(|\bar{Y}_N| \geq c) \approx \exp \left\{ -N \inf_{|y| \geq c} I(y) \right\}, \tag{1.3}
\]

where for \( y \in \mathbb{R}, I(y) = \sup_{\alpha \in \mathbb{R}} \{\langle \alpha, y \rangle - \log \int_{\mathbb{R}} \exp(\langle \alpha, z \rangle \rho(dz)) \} \) is the rate function. Such a result gives probabilities of deviations of order 1 from the LLN. Roughly speaking, MDP gives estimates on deviations of orders that lie between 1 and \( \frac{1}{\sqrt{N}} \). More precisely, let \( \{a(N)\} \) be a positive sequence such that \( a(N) \to 0 \) and \( a(N) \sqrt{N} \to \infty \) as \( N \to \infty \) (e.g. \( a(N) = N^{-\delta} \) for some \( \delta \in (0, \frac{1}{2}) \)). Then a MDP for \( \bar{Y}_N \) says that

\[
P \left( a(N) \sqrt{N} |\bar{Y}_N| \geq c \right) \approx \exp \left\{ -\frac{1}{[a(N)]^2} \inf_{|y| \geq c} I^0(y) \right\}, \tag{1.4}
\]

where \( I^0(y) = y^2 / 2\sigma^2 \) and \( \sigma^2 \equiv \text{Var}(Y_i) \). Thus the MDP gives estimates on probabilities of deviations from the LLN limit of order \( \frac{1}{a(N)\sqrt{N}} \), which is lower order than 1 and higher than \( \frac{1}{\sqrt{N}} \). Note that the rate function for a MDP which is a quadratic form has a much simpler expression than that for the LDP. Since \( a(N) \) can converge to 0 as slowly as desired, a moderate deviation principle bridges the gap between a central limit approximation and a large deviations approximation.

**New Large and Moderate Deviation Principles.** In Chapters 6, 7 and 8 we will present some large and moderate deviation results for empirical measures associated with several types of weakly interacting particle systems. In Chapter 6, a MDP for weakly interacting diffusions is established. Next, in Chapter 7, a model for systems with pure jumps and countable state space is studied. Finally, in Chapter 8, we study a LDP for a weakly interacting particle system associated with a nonlinear heat equation.

**Organization and Overview of the Chapters.** This dissertation is organized as follows. We begin with some background material in Chapter 2. Chapters 3 – 8 will present our main asymptotic results for several different types of weakly interacting particle systems.
In Chapter 3 we introduce a weakly interacting multi-type particle system. The system consists of $K$ different kinds of particles such that each type interacts with the particles of other types through a drift coefficient that depends only on the particle types. Particles of each type have i.i.d. initial distributions although these distributions may differ between different types. The system models a network of $N$ interacting spiking neurons and has been previously studied by Baladron et al. [2] where the authors study the law of large numbers behavior and prove a propagation of chaos result. In our work we study fluctuations about the law of large numbers limit and establish a central limit theorem. The classical central limit theorem for single-type system relies crucially on the fact that statistical distribution of state process of $N$ particles is exchangeable. However in the multi-type setting considered in our work, this exchangeability property breaks down. There is no obvious way to view the collection as a $Kd$-dimensional exchangeable vector process, and hence new ideas are needed.

In Chapter 4 we consider a variant of the model from Chapter 3, where neurons of each type are influenced by a common noise source (common factor). The coefficients in both particle and common factor dynamics depend on not only the state of process itself, but also empirical measures of particles belonging to each type. Due to the presence of the common factor, which does not get averaged out, the limit of empirical measures will in general be a random measure, and the centering in the fluctuation theorem will typically be random as well. Fluctuations are characterized, instead of by a Gaussian limit, through a mixture of Gaussian distributions. The results of Chapters 3 and 4 are taken from the paper [18].

In Chapter 5, we consider another variant of the model in (1.1) in a setting where the interaction graph is not complete but rather given in terms of an Erdős–Rényi random graph on $N$ vertices. More precisely, we are given at time 0 an Erdős–Rényi random graph with parameter $p_N(0)$ on $N$ vertices, namely any pair of vertices have an edge with probability $p_N(0)$ independently for different pairs of vertices. As time evolves, edges form and break down independently, so that at time $t$ the random graph is still of Erdős–Rényi type but with parameter $p_N(t)$. An $O(\frac{1}{N})$ interaction occurs between a pair of particles if and only if there is an edge between them at that time instant, otherwise the particles do not interact. It will be shown in Chapter 5 that under conditions on $p_N(\cdot)$, a propagation of chaos result and a central limit theorem hold. A paper [3] based on this work is near completion.
In Chapters 6 and 7 we are interested in estimating probabilities of moderate deviations of empirical measure processes, from the mean field limit, for weakly interacting Markov processes. Two families of models are considered and for each one a MDP is established. The model in Chapter 6 corresponds to a system of interacting diffusions whereas Chapter 7 describes a collection of pure jump Markov processes with a countable state space. Although not treated here, one can use similar techniques to develop moderate deviation results for settings that have both Brownian and Poisson type noises. For both cases the MDP is formulated in terms of a LDP with an appropriate speed function, for suitably centered and normalized empirical measure processes. For the first family of models the LDP is established in the path space of an appropriate Schwartz distribution space whereas for the second family the LDP is proved in the space of $l_2$ (the Hilbert space of square summable sequences)-valued paths. The results of these two Chapters are taken from the paper [17].

Finally, in Chapter 8 we study the particle approximation for a nonlinear heat equation using a system of Brownian motions with killing. The system that we consider is described by a collection of i.i.d. Brownian particles, and each particle is killed independently at a rate determined by the empirical sub-probability measure of the states of the alive particles. A law of large numbers result and large deviation principle for such sub-probability measure processes are established. As a byproduct, we also give a convenient variational representation for expectations of positive functionals of Brownian motions along with an i.i.d. sequence of random variables. A paper [14] based on this work is near completion.

**Notation.** The following notation will be used. For a Polish space $(\mathcal{S}, d(\cdot, \cdot))$, denote the corresponding Borel σ-field by $\mathcal{B}(\mathcal{S})$. For a signed measure $\mu$ on $\mathcal{S}$ and $\mu$-integrable function $f : \mathcal{S} \to \mathbb{R}$, we will use $\langle \mu, f \rangle$ and $\langle f, \mu \rangle$ interchangeably for $\int f \, d\mu$. Let $\mathcal{P}(\mathcal{S})$ be the space of probability measures on $\mathcal{S}$, equipped with the topology of weak convergence. A convenient metric for this topology is the bounded-Lipschitz metric $d_{BL}$, which metrizes $\mathcal{P}(\mathcal{S})$ as a Polish space, defined as

$$d_{BL}(\nu_1, \nu_2) \doteq \sup_{\|f\|_{BL} \leq 1} |\langle \nu_1 - \nu_2, f \rangle|, \quad \nu_1, \nu_2 \in \mathcal{P}(\mathcal{S}),$$
where \( \| \cdot \|_{BL} \) is the bounded Lipschitz norm, i.e. for \( f : S \to \mathbb{R} \),

\[
\| f \|_{BL} = \max \{ \| f \|_\infty, \| f \|_L \}, \quad \| f \|_\infty = \sup_{x \in S} |f(x)|, \quad \| f \|_L = \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}.
\]

Denote by \( M_b(S) \) (resp. \( C_b(S) \)) the space of real bounded \( \mathcal{B}(S)/\mathcal{B}(\mathbb{R}) \)-measurable functions (resp. real bounded and continuous functions). Let \( C^k_b(\mathbb{R}^d) \) be the space of functions on \( \mathbb{R}^d \), which have continuous and bounded derivatives up to the \( k \)-th order. For a measure \( \nu \) on \( S \) and a Hilbert space \( H \), let \( L^2(S, \nu, H) \) denote the space of measurable functions \( f \) from \( S \) to \( H \) such that

\[
\int_S \| f(x) \|^2_H \nu(dx) < \infty,
\]

where \( \| \cdot \|_H \) is the norm on \( H \). When \( H = \mathbb{R} \) and \( S \) is clear from the context we write \( L^2(S, \nu) \) or simply \( L^2(\nu) \).

Fix \( T < \infty \). All stochastic processes will be considered over the time horizon \([0, T]\). We will use the notations \( \{ X_t \} \) and \( \{ X(t) \} \) interchangeably for stochastic processes. For a Polish space \( S \), denote by \( C([0, T] : S) \) (resp. \( D([0, T] : S) \)) the space of continuous functions (resp. right continuous functions with left limits) from \([0, T]\) to \( S \), endowed with the uniform (resp. Skorokhod) topology. For a map \( f : [0, T] \to S \), where \( S \) is a normed space with norm \( \| \cdot \| \), let \( \| f \|_{*, t} = \sup_{0 \leq s \leq t} \| f(s) \| \) for \( t \in [0, T] \).

Probability law of an \( S \)-valued random variable \( X \) will be denoted as \( \mathcal{L}(X) \). We will use \( \mathbb{E}^P \) and \( \mathbb{E}_P \) interchangeably for expected value under some probability law \( P \). We say a collection \( \{ X_n \} \) of \( S \)-valued random variables is tight if \( \{ \mathcal{L}(X_n) \} \) are tight in \( \mathcal{P}(S) \). Convergence of a sequence \( \{ X_n \} \) of \( S \)-valued random variables in distribution to \( X \) will be written as \( X_n \Rightarrow X \).

We will usually denote by \( \kappa, \kappa_1, \kappa_2, \ldots \), the constants that appear in various estimates within a proof. The value of these constants may change from one proof to another. Cardinality of a finite set \( A \) will be denoted as \( |A| \). For \( x, y \in \mathbb{R}^d \), \( x \cdot y = \sum_{i=1}^d x_i y_i \).
CHAPTER 2
PRELIMINARIES

2.1 Symmetric statistics and multiple Wiener integrals

In this section we give some background on symmetric statistics, multiple Winer integrals, and related results, which will be used in Chapters 3, 4 and 5.

Let $\mathbb{S}$ be a Polish space and let $\{X_n\}_{n \in \mathbb{N}}$ be a sequence of i.i.d. $\mathbb{S}$-valued random variables having common probability law $\nu$. For $k \in \mathbb{N}$, let $L^2(\nu^{\otimes k})$ be the space of all real valued square integrable functions on $(\mathbb{S}^k, \mathcal{B}(\mathbb{S})^{\otimes k}, \nu^{\otimes k})$. Denote by $L^2_c(\nu^{\otimes k})$ the subspace of centered functions, namely $\phi \in L^2(\nu^{\otimes k})$ such that for all $1 \leq j \leq k$,

$$\int_{\mathbb{S}} \phi(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_k) \nu(dx) = 0, \ \nu^{\otimes k-1} \text{ a.e. } (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k).$$

Denote by $L^2_{\text{sym}}(\nu^{\otimes k})$ the subspace of symmetric functions, namely $\phi \in L^2(\nu^{\otimes k})$ such that for every permutation $\pi$ on $\{1, \ldots, k\}$,

$$\phi(x_1, \ldots, x_k) = \phi(x_{\pi(1)}, \ldots, x_{\pi(k)}), \ \nu^{\otimes k} \text{ a.e. } (x_1, \ldots, x_k).$$

Also, denote by $L^2_{c,\text{sym}}(\nu^{\otimes k})$ the subspace of centered symmetric functions in $L^2(\nu^{\otimes k})$, namely $L^2_{c,\text{sym}}(\nu^{\otimes k}) \doteq L^2_c(\nu^{\otimes k}) \cap L^2_{\text{sym}}(\nu^{\otimes k})$. Given $\phi_k \in L^2_{c,\text{sym}}(\nu^{\otimes k})$ define the symmetric statistic $\sigma^n_k(\phi_k)$ as

$$\sigma^n_k(\phi_k) \doteq \begin{cases} 
\sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \phi_k(X_{i_1}, \ldots, X_{i_k}) & \text{for } n \geq k, \\
0 & \text{for } n < k.
\end{cases}$$

In order to describe the asymptotic distributions of such statistics consider a Gaussian field $\{I_1(h) : h \in L^2(\nu)\}$ such that

$$\mathbf{E}(I_1(h)) = 0, \ \mathbf{E}(I_1(h)I_1(g)) = \langle h, g \rangle_{L^2(\nu)}, \ h, g \in L^2(\nu).$$
For $h \in L^2(\nu)$, define $\phi^h_k \in L^2_{sym}(\nu \otimes^k)$ as $\phi^h_k(x_1, \ldots, x_k) \doteq h(x_1) \cdots h(x_k)$ and set $\phi^h_0 \doteq 1$.

The multiple Wiener integral (MWI) of $\phi^h_k$, denoted as $I_k(\phi^h_k)$, is defined through the following formula. For $k \geq 1$,

$$I_k(\phi^h_k) \doteq \left\lfloor \frac{k}{2} \right\rfloor \sum_{j=0}^{\left\lfloor \frac{k}{2} \right\rfloor} (-1)^j C_{k,j} \|h\|^2_{L^2(\nu)} (I_1(h))^{k-2j},$$

$$C_{k,j} \doteq \frac{k!}{(k-2j)!2^j j!}, \quad j = 0, \ldots, \left\lfloor \frac{k}{2} \right\rfloor. \quad (2.1)$$

The following representation gives an equivalent way to characterize the MWI of $\phi^h_k$:

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} I_k(\phi^h_k) = \exp \left( t I_1(h) - \frac{t^2}{2} \|h\|^2_{L^2(\nu)} \right), \quad t \in \mathbb{R},$$

where we set $I_0(\phi^h_0) \doteq 1$. We extend the definition of $I_k$ to the linear span of $\{\phi^h_k : h \in L^2(\nu)\}$ by linearity. It can be checked that for all $f$ in this linear span,

$$\mathbb{E}(I_k(f)^2) = k! \|f\|^2_{L^2(\nu \otimes^k)}. \quad (2.3)$$

Using this identity and standard denseness arguments, the definition of $I_k(f)$ can be extended to all $f \in L^2_{sym}(\nu \otimes^k)$ and the identity (2.3) holds for all $f \in L^2_{sym}(\nu \otimes^k)$. The following theorem describes asymptotic distributions of symmetric statistics.

**Theorem 2.1** (Dynkin–Mandelbaum [29]). Let $\{\phi_k\}_{k=1}^{\infty}$ be such that, for each $k \geq 1$, $\phi_k \in L^2_{c,sym}(\nu \otimes^k)$. Then as $n \to \infty$,

$$\left( \frac{1}{n^2} \sigma^2_k(\phi_k) \right)_{k \geq 1} \Rightarrow \left( \frac{1}{k!} I_k(\phi_k) \right)_{k \geq 1}$$

as a sequence of $\mathbb{R}^\infty$-valued random variables.

We will use the following lemma taken from Shiga–Tanaka [62] when proving central limit theorems in Chapters 3, 4 and 5. Let $a(\cdot, \cdot) \in L^2(\nu \otimes \nu)$ and denote by $A$ the integral operator on $L^2(\nu)$ associated with $a$: For $x \in \mathbb{S}$ and $\phi \in L^2(\nu)$

$$A\phi(x) = \int_{\mathbb{S}} a(x, y) \phi(y) \nu(dy).$$
Then $A$ is a Hilbert-Schmidt operator. Also, $AA^*$, and for $n \geq 2$, $A^n$, are trace class operators. Furthermore, the following lemma holds.

**Lemma 2.2** (Shiga–Tanaka [62]). Suppose that $\text{Trace}(A^n) = 0$ for all $n \geq 2$. Then $E[e^{\frac{1}{2}I_2(f)}] = e^{\frac{1}{4}\text{Trace}(AA^*)}$, where $f(x,y) = a(x,y) + a(y,x) - \int_S a(x,z)a(y,z)\nu(dz)$ for $x,y \in S$. Moreover, $I - A$ is invertible and for any $\phi \in L^2(\nu)$,

$$E\left[\exp \left(\frac{1}{2}I_2(f)\right)\right] = \exp \left[-\frac{1}{2} \left(\|(I - A)^{-1}\phi\|_{L^2(\nu)}^2 - \text{Trace}(AA^*)\right)\right],$$

where $I$ is the identity operator on $L^2(\nu)$.

### 2.2 Large deviation and moderate deviation principles

Definition and results of this section are used in Chapters 6, 7 and 8.

Let $\{X_n\}_{n \in \mathbb{N}} \equiv \{X_n\}$ be a sequence of $S$-valued random variables defined on some probability space $(\Omega, \mathcal{F}, P)$, where $S$ is a Polish space. The theory of large deviations concerns events $A \in \mathcal{B}(S)$, for which $P(X_n \in A) \to 0$ exponentially fast as $n \to \infty$. The exponential decay rate is usually given through a rate function $I: S \to [0, \infty]$.

**Definition 2.3** (Rate function). A function $I: S \to [0, \infty]$ is called a rate function on $S$ if for each $M < \infty$, the level set $\{x \in S : I(x) \leq M\}$ is a compact subset of $S$.

**Definition 2.4** (Large deviation principle). Let $I$ be a rate function on $S$. The sequence $\{X_n\}$ is said to satisfy the large deviation principle on $S$, as $n \to \infty$, with rate function $I$ and speed function $\alpha(n)$ if the following two conditions hold.

(a) Large deviation upper bound. For each closed set $F \subset S$,

$$\limsup_{n \to \infty} \alpha(n) \log P(X_n \in F) \leq - \inf_{x \in F} I(x).$$

(b) Large deviation lower bound. For each open set $G \subset S$,

$$\liminf_{n \to \infty} \alpha(n) \log P(X_n \in G) \geq - \inf_{x \in G} I(x).$$
Formally, the above definition says that if \( \{X_n\} \) satisfies the large deviation principle with rate function \( I \) and speed function \( \alpha(n) \), then
\[
P(X_n \in A) \approx \exp \left\{ -\frac{1}{\alpha(n)} \inf_{x \in A} I(x) \right\}, \quad A \subset \mathbb{S}.
\]

For example, when \( X_n = \bar{Y}_n \), where \( \bar{Y}_n \) is as introduced above (1.3), Cramer’s theorem gives a LDP for \( X_n \) with \( \alpha(n) = \frac{1}{n} \). However, there are natural problem settings where one is interested in a different speed function. For instance, in many problems one is interested in the asymptotic decay of probabilities of deviations that are of smaller order than those given by a LDP. Such an asymptotic result is given in terms of a moderate deviation principle and it is convenient to formulate this result as a LDP with a different speed function. See for example (1.4), where a MDP for \( \bar{Y}_n \) is expressed as a LDP for \( a(n)\sqrt{n}\bar{Y}_n \) with speed function \([a(n)]^2\). In Chapters 6 and 7 of this work we will present some moderate deviation principles for weakly interacting Markov processes. Finally in Chapter 8, we will study a large deviation principle for a class of some weakly interacting systems. These systems arise as particle approximations for certain nonlinear heat equations.

It is well known that if a sequence of random variables satisfies the large deviation principle with some rate function, then the rate function is unique (cf. [27], Theorem 1.3.1). In many problems one is interested in exponential decays of positive functionals which are more general than indicator functions of closed or open sets. Such asymptotic results are usually formulated in terms of a Laplace principle.

**Definition 2.5 (Laplace principle).** Let \( I \) be a rate function on \( \mathbb{S} \). The sequence \( \{X_n\} \) is said to satisfy the Laplace principle upper bound (respectively lower bound) on \( \mathbb{S} \), as \( n \to \infty \), with rate function \( I \) and speed function \( \alpha(n) \) if for all \( h \in \mathcal{C}_b(\mathbb{S}) \),
\[
\limsup_{n \to \infty} \alpha(n) \log \mathbb{E} \left\{ \exp \left[ -\frac{1}{\alpha(n)} h(X_n) \right] \right\} \leq -\inf_{x \in \mathbb{S}} \{h(x) + I(x)\},
\]
and, respectively,
\[
\liminf_{n \to \infty} \alpha(n) \log \mathbb{E} \left\{ \exp \left[ -\frac{1}{\alpha(n)} h(X_n) \right] \right\} \geq -\inf_{x \in \mathbb{S}} \{h(x) + I(x)\}.
\]
Laplace principle is said to hold for \( \{X_n\} \) with rate function \( I \) and speed function \( \alpha(n) \) if both the Laplace upper and lower bounds are satisfied for all \( h \in \mathcal{C}_b(\mathbb{S}) \).

One of the main results in the theory of large deviations is the following equivalence between the Laplace principle and the large deviation principle. For a proof we refer to Section 1.2 of [27].

**Theorem 2.6.** The family \( \{X_n\} \) satisfies the Laplace principle upper (respectively lower) bound with speed function \( \alpha(n) \) and rate function \( I \) on \( \mathbb{S} \) if and only if \( \{X_n\} \) satisfies the large deviation upper (respectively lower) bound for all closed sets (respectively open sets) with speed function \( \alpha(n) \) and rate function \( I \).

Theorem 2.6 will be key in the proofs of the LDP and MDP results presented in Chapters 6, 7 and 8.

### 2.3 A variational representation for functionals of Brownian motions

In this section we present a variational representation for functionals of Brownian motions, which will be used to prove the Laplace principle in Chapter 6. Such a representation was obtained in [8]. For our purpose it is convenient to use a form of the representation given in [10] that allows for an arbitrary filtration. Let \((\Omega, \mathcal{F}, P)\) be a probability space with an increasing family of right continuous \( P \)-complete \( \sigma \)-fields \( \{\mathcal{F}_t\} \). Let for \( n \in \mathbb{N} \), \( \{B(t) \doteq (B_1(t), \ldots, B_n(t)), 0 \leq t \leq T\} \) be an \( n \)-dimensional standard \( \mathcal{F}_t \)-Brownian motions on this probability space. Let \( \mathcal{A} \) be the collection of \( \mathbb{R}^n \)-valued \( \mathcal{F}_t \)-progressively measurable processes \( \{u(t) \doteq (u_1(t), \ldots, u_n(t)), 0 \leq t \leq T\} \) such that \( P \left( \int_0^T \|u(s)\|^2 \, ds < \infty \right) = 1 \).

The following variational representation follows from [8, 10].

**Theorem 2.7.** Let \( f \in M_b(\mathcal{C}([0,T]:\mathbb{R}^n)) \). Then

\[
- \log E \exp\{-f(B)\} = \inf_{u \in \mathcal{A}} E \left( \frac{1}{2} \int_0^T \|u(s)\|^2 \, ds + f \left( B + \int_0^T u(s) \, ds \right) \right).
\]

We will generalize the above result in Chapter 8 to allow for functionals that depend on Brownian motions and a collection of i.i.d. random variables. Using this new representation we will then prove a Laplace principle and hence the corresponding LDP for certain empirical measures valued processes.
CHAPTER 3
WEAKLY INTERACTING MULTI-TYPE PARTICLE SYSTEMS

In this chapter we are interested in the following weakly interacting multi-type particle systems. For \( N \geq 1 \), let \( Z_{1,N}, \ldots, Z_{N,N} \) be \( \mathbb{R}^d \)-valued stochastic processes, representing trajectories of \( N \) particles, each of which belongs to one of \( K \) types (populations) with the membership map denoted by \( p: \{1, \ldots, N\} \to \{1, \ldots, K\} \equiv K \), namely \( i \)-th particle is type \( \alpha \) if \( p(i) = \alpha \). The dynamics is given in terms of a collection of stochastic differential equations (SDE) driven by mutually independent Brownian motions (BM) with each particle’s initial condition governed independently by a probability law that depends only on its type. The \( N \) stochastic processes interact with each other through the coefficients of the SDE which, for the \( i \)-th process, with \( p(i) = \alpha \), depend on not only the \( i \)-th state process and the \( \alpha \)-th type, but also the empirical measures \( \mu_{i,N}^{\gamma,N} = \frac{1}{N_\gamma} \sum_{j : p(j) = \gamma} \delta_{Z^j,N} \) for all \( \gamma \in K \). Here \( N_\gamma \) is the total number of particles that belong to the \( \gamma \)-th type.

Classical works for weakly interacting particle systems that study the law of large number (LLN) results and central limit theorems (CLT) include McKean [52, 53], Braun and Hepp [9], Dawson [24], Tanaka [68], Oelschläger [56], Sznitman [66, 67], Graham and Méliard [36], Shiga and Tanaka [62], Méliard [54]. All the above papers consider exchangeable populations, i.e. \( K = 1 \). The goal of this chapter is to study fluctuations for multi-type particle systems. Since these systems are not exchangeable (there is also no natural way to regard the system as a \( K \)-vector of \( d \)-dimensional exchangeable particles), classical techniques for proving CLT, developed in the above papers [66, 62, 54], are not directly applicable.

Systems with multi-type particles show up in many different areas, including statistical mechanics [20], social sciences [22], neurosciences [2], etc. These also arise naturally, even if the original model is single-type, when one is interested in the joint asymptotic behavior of disjoint subpopulations whose sizes grow to infinity as \( N \to \infty \). In particular, to study the global behavior of large-scale neural assemblies with special structures that characterize specific brain functions,
Baladron et al. [2] considers interacting diffusions of the form studied in this chapter and establishes a LLN result and a propagation of chaos (POC) property. Our results in particular will provide asymptotic results on fluctuations from the LLN behavior obtained in [2].

Our main objective is to establish a suitable CLT where the summands are quite general functionals of the trajectories of the particles with suitable integrability properties. Specifically, letting $N_\alpha$ denote the set of indices $i$ such that $p(i) = \alpha$ and

$$
\xi_\alpha^N(\phi) = \frac{1}{\sqrt{N_\alpha}} \sum_{i \in N_\alpha} \phi(Z_{i,N}^),
$$

for functions $\phi$ on the path space of the particles, we will establish (see Theorem 3.4) the weak convergence of the family $\{\xi_\alpha^N(\phi), \phi \in A_\alpha, \alpha \in K\}$, in the sense of finite dimensional distributions, to a mean zero Gaussian field $\{\xi_\alpha(\phi), \phi \in A_\alpha, \alpha \in K\}$. Here $A_\alpha$ is a family of functions on the path space that are suitably centered and have appropriate integrability properties (see Section 3.2.1 for definitions).

CLT established in this chapter also leads to new fluctuation results for the classical single-type setting. Suppose one is interested in the joint asymptotic distribution of $(\xi_1^N(\phi_1), \xi_2^N(\phi_2))$, where

$$
\xi_1^N(\phi_1) = \frac{1}{\sqrt{\lambda N}} \sum_{i=1}^{\lfloor \lambda N \rfloor} \phi_1(Z_{i,N}^), \quad \xi_2^N(\phi_2) = \frac{1}{\sqrt{N - \lfloor \lambda N \rfloor}} \sum_{i=\lfloor \lambda N \rfloor + 1}^{N} \phi_2(Z_{i,N}^) \quad \text{(3.1)}
$$

and $\lambda \in (0,1)$. Existing results on central limit theorems for $K = 1$ (eg. [66, 62, 54]) do not give information on the joint limiting behavior of the above random variables. Indeed, a naive guess that the propagation of chaos property should imply the asymptotic independence of $\xi_1^N(\phi_1)$ and $\xi_2^N(\phi_2)$ is in general false. In Section 3.2.2 we will illustrate through a simple example how one can characterize the joint asymptotic distribution of the above pair.

We are also interested in asymptotic behavior of path functionals of particles of multiple type. Specifically, we will study in Section 3.2.3 the limiting distribution of multi-type statistics of the form

$$
\xi^K(\phi) = \frac{1}{\sqrt{N_1 \cdots N_K}} \sum_{i_1 \in N_1} \cdots \sum_{i_K \in N_K} \phi(Z_{i_1,N}^, \ldots, Z_{i_K,N}^) \quad \text{(3.2)}
$$
where \( \phi \) is a suitably centered function on the path space of \( K \times d \) dimensional stochastic processes with appropriate integrability. In the classical case (cf. [29]) where the particles have independent and identical dynamics, limit distributions of analogous statistics (with \( N_1 = \cdots = N_K \)) are given through certain multiple Wiener integrals. In the setting considered here \( \{Z^{i,N}\} \) are neither independent nor identical and we need to suitably extend the classical result for U-statistics to a multi-type setting and apply techniques as in the proof of Theorem 3.4 to establish weak convergence of (3.2) and characterize the limit distributions.

This chapter is organized as follows. In Section 3.1 we begin by introducing our model of multi-type weakly interacting diffusions. A basic condition (Condition 3.1) is stated, under which both SDE for the pre-limit \( N \)-particle system and for the corresponding limiting nonlinear diffusion process have unique solutions, and a law of large numbers and a propagation of chaos property holds. These results are taken from [2]. We consider fluctuations in Section 3.2 and present a CLT (Theorem 3.4) in Section 3.2.1. As noted previously, this CLT gives new asymptotic results for a single type population as well. This point is illustrated through an example in Section 3.2.2. We also give a limit theorem (Theorem 3.5) for multi-type statistics of the form as in (3.2) in Section 3.2.3. Finally proofs are provided in Sections 3.3, 3.4 and 3.5.

3.1 Model

Consider an infinite collection of particles where each particle belongs to one of \( K \) different populations. Letting \( K = \{1, \ldots, K\} \), define a function \( p: \mathbb{N} \rightarrow K \) by \( p(i) = \alpha \) if the \( i \)-th particle belongs to \( \alpha \)-th population. For \( N \in \mathbb{N} \), let \( N = \{1, \ldots, N\} \). For \( \alpha \in K \), let \( N_\alpha = \{i \in N : p(i) = \alpha\} \) and denote by \( N_\alpha \) the number of particles belonging to the \( \alpha \)-th population, namely \( N_\alpha = |N\alpha| \). Assume that \( N_\alpha/N \rightarrow \lambda_\alpha \in (0,1) \) as \( N \rightarrow \infty \). Let \( \mathbb{N}_\alpha = \{i \in \mathbb{N} : p(i) = \alpha\} \).

For fixed \( N \geq 1 \), consider the following system of equations for the \( \mathbb{R}^d \)-valued continuous stochastic processes \( Z^{i,N} \), \( i \in N \), given on a filtered probability space \((\Omega, \mathcal{F}, P, \{\mathcal{F}_t\})\) that satisfies the usual conditions (namely \( \mathcal{F} \) and \( \mathcal{F}_t \) are \( P \)-complete and the filtration is right continuous). For \( i \in N_\alpha, \alpha \in K \),

\[
Z_t^{i,N} = Z_0^{i,N} + \int_0^t f_\alpha(s, Z_s^{i,N}) \, ds + \int_0^t \sum_{\gamma=1}^K \langle b_{\alpha\gamma}(Z_s^{i,N}, \cdot), \mu_s^{\gamma,N} \rangle \, ds + W_t^i, \tag{3.3}
\]
where \( f_\alpha : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \) and \( b_{\alpha \gamma} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \) are suitable functions, and \( \mu_{s}^{\gamma,N} = \frac{1}{N_{\gamma}} \sum_{j \in N_{\gamma}} \delta_{Z_{i,N}}^{i} \). Here \( \{W^i : i \in N\} \) are mutually independent \( d \)-dimensional \( \{F_t\}\)-Brownian motions. We assume that \( \{Z_{i,N}^{i} : i \in N\} \) are \( F_0 \)-measurable and mutually independent with \( \mathcal{L}(Z_{i,N}^{i}) = \mu_{0}^{\alpha} \) for \( i \in N_{\alpha}, \alpha \in K \).

Conditions on the various coefficients will be introduced shortly. Along with the \( N \)-particle equation (3.3) we will also consider a related infinite system of equations for \( \mathbb{R}^d \)-valued continuous McKean-Vlasov processes \( X_i, i \in N \), given (without loss of generality) on \( (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}) \). For \( \alpha \in K \) and \( i \in N_{\alpha} \),

\[
X_i^t = X_0^i + \int_0^t f_\alpha(s, X_i^s) \, ds + \int_0^t \sum_{\gamma=1}^{K} \langle b_{\alpha \gamma}(X_i^s, \cdot), \mu_{s}^{\gamma} \rangle \, ds + W_i^t, \quad (3.4)
\]

where \( \mu_{s}^{\alpha} \equiv \mathcal{L}(X_i^s) \). We assume that \( \{X_0^i : i \in N\} \) are \( F_0 \)-measurable and mutually independent, with \( \mathcal{L}(X_0^i) = \mu_{0}^{\alpha} \) for \( \alpha \in K \) and \( i \in N_{\alpha} \).

The existence and uniqueness of pathwise solutions of (3.3) and (3.4) can be shown under following conditions on the coefficients (cf. [66, 67], see also [2]).

**Condition 3.1.** (a) For all \( \alpha \in K \), the functions \( f_\alpha \) are locally Lipschitz in the second variable, uniformly in \( t \in [0, T] \): For every \( r > 0 \) there exists \( L_r \in (0, \infty) \) such that for all \( x, y \in \{z \in \mathbb{R}^d : \|z\| \leq r\} \), and \( t \in [0, T] \),

\[
\|f_\alpha(t, x) - f_\alpha(t, y)\| \leq L_r \|x - y\|.
\]

(b) There exists \( L \in (0, \infty) \) such that for all \( \alpha \in K \), \( t \in [0, T] \) and \( x \in \mathbb{R}^d \),

\[
x \cdot f_\alpha(t, x) \leq L(1 + \|x\|^2).
\]

(c) For all \( \alpha, \gamma \in K \), \( b_{\alpha \gamma} \) are bounded Lipschitz functions: There exists \( L \in (0, \infty) \) such that \( \|b_{\alpha \gamma}\|_{BL} \leq L \).

The paper [2] also proves the following POC property: For any \( n \)-tuple \( (i_1^N, \ldots, i_n^N) \in N^n \) with \( i_j^N \neq i_k^N \) whenever \( j \neq k \), and \( p(i_j^N) = \alpha_j, j = 1, \ldots, n \),

\[
\mathcal{L}(\{Z_1^{i_1^N}, \ldots, Z_n^{i_n^N}\}) \to \mu_{\alpha_1} \otimes \cdots \otimes \mu_{\alpha_n} \quad \text{as} \quad N \to \infty, \quad (3.5)
\]
in \( \mathcal{P}(\mathcal{C}_d^\alpha) \), where for \( \alpha \in K \) and \( i \in \mathbb{N}_\alpha \), \( \mu^\alpha \equiv L(X^i) \in \mathcal{P}(\mathcal{C}_d) \), \( X^i \) is as in (3.4), and \( \mathcal{C}_d \equiv \mathbb{C}([0,T]: \mathbb{R}^d) \). Using the above result and a straightforward argument similar to [67] one can show the following law of large numbers. Proof is omitted.

**Theorem 3.2.** Suppose Condition 3.1 holds. For all \( f \in \mathcal{C}_b(\mathcal{C}_d^K) \), as \( N \to \infty \),

\[
\frac{1}{N_1 \cdots N_K} \sum_i f(Z^{i_1:N}, \ldots, Z^{i_K:N}) \Rightarrow \langle f, \mu^1 \otimes \cdots \otimes \mu^K \rangle, \tag{3.6}
\]

where the summation is taken over all \( K \)-tuples \( i = (i_1, \ldots, i_K) \in \mathbb{N}_1 \times \cdots \times \mathbb{N}_K \).

Note that Theorem 3.2 implies in particular that for all \( \phi \in \mathcal{C}_b(\mathcal{C}_d) \) and \( \alpha \in K \), as \( N \to \infty \),

\[
\frac{1}{N_\alpha} \sum_{i \in \mathbb{N}_\alpha} \phi(Z^{i:N}) \Rightarrow \langle \phi, \mu^\alpha \rangle. \tag{3.7}
\]

In this chapter we are concerned with the fluctuations of expressions as in the LHS of (3.6) and (3.7) about their LLN limits given by the RHS of (3.6) and (3.7), respectively. Limit theorems that characterize these fluctuations are given in Theorems 3.4 and 3.5.

### 3.2 Fluctuations for multi-type particle system

Throughout this section Condition 3.1 will be assumed and will not be noted explicitly in the statement of results.

#### 3.2.1 Central limit theorem

We first introduce the following canonical spaces and stochastic processes. Let \( \Omega_d \equiv \mathcal{C}_d \times \mathcal{C}_d \). For \( \alpha \in K \), denote by \( \nu_\alpha \in \mathcal{P}(\Omega_d) \) the law of \((W^i, X^i)\) where \( i \in \mathbb{N}_\alpha \) and \( X^i \) is given by (3.4). Let \( \hat{\nu} \equiv \nu_1 \otimes \cdots \otimes \nu_K \). Define for \( N \in \mathbb{N} \) the probability measure \( \mathbb{P}^N \) on \( \Omega_d^N \) as

\[
\mathbb{P}^N \equiv L\left( (W^1, X^1), (W^2, X^2), \ldots, (W^N, X^N) \right) = \nu_{p(1)} \otimes \nu_{p(2)} \otimes \cdots \otimes \nu_{p(N)},
\]

For \( \bar{\omega} = (\omega_1, \omega_2, \ldots, \omega_N) \in \Omega_d^N \), let \( V^i(\bar{\omega}) \equiv \omega_i, i \in \mathbb{N} \). Abusing notation, write

\[
V^i = (W^i, X^i), \quad i \in \mathbb{N}. \tag{3.8}
\]
Also define the canonical processes $V_\ast \doteq (W_\ast, X_\ast)$ on $\Omega_d$ as

$$V_\ast(\omega) = (W_\ast(\omega), X_\ast(\omega)) \equiv (\omega_1, \omega_2), \quad \omega = (\omega_1, \omega_2) \in \Omega_d.$$  

We will need the following functions for stating our first main theorem. Define for $\alpha, \gamma \in K$ and $t \in [0, T]$, the function $b_{\alpha \gamma, t}$ from $\mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^d$ as

$$b_{\alpha \gamma, t}(x, y) \doteq b_{\alpha \gamma}(x, y) - \langle b_{\alpha \gamma}(x, \cdot), \mu_t^\gamma \rangle, \quad (x, y) \in \mathbb{R}^d \times \mathbb{R}^d.$$  

Define for $\alpha, \gamma \in K$, the function $h_{\alpha \gamma}$ from $\Omega_d \times \Omega_d$ to $\mathbb{R}$ ($\nu_\alpha \otimes \nu_\gamma$ a.s.) as

$$h_{\alpha \gamma}(\omega, \omega') \doteq \sqrt{\frac{\lambda_\alpha}{\lambda_\gamma}} \int_0^T b_{\alpha \gamma, t}(X_{\ast, t}(\omega), X_{\ast, t}(\omega')) \cdot dW_{\ast, t}(\omega), \quad (\omega, \omega') \in \Omega_d \times \Omega_d.$$  

We now define lifted functions $\hat{h}_{\alpha \gamma}, \hat{h}$ from $\Omega_d^K \times \Omega_d^K$ to $\mathbb{R}$ ($\nu_\alpha \otimes \nu_\gamma$ a.s.) for $\alpha, \gamma \in K$ as follows: For $\omega = (\omega_1, \ldots, \omega_K)$ and $\omega' = (\omega'_1, \ldots, \omega'_K)$ in $\Omega_d^K$, let

$$\hat{h}_{\alpha \gamma}(\omega, \omega') \doteq h_{\alpha \gamma}(\omega_\alpha, \omega'_\gamma), \quad \hat{h}(\omega, \omega') \doteq \sum_{\alpha=1}^K \sum_{\gamma=1}^K \hat{h}_{\alpha \gamma}(\omega, \omega').$$  

Now consider the Hilbert space $L^2(\Omega_d^K, \nu)$. For $\alpha, \gamma \in K$, define integral operators $A_{\alpha \gamma}$ and $A$ on $L^2(\Omega_d^K, \nu)$ as follows: For $f \in L^2(\Omega_d^K, \nu)$ and $\omega \in \Omega_d^K$,

$$A_{\alpha \gamma} f(\omega) \doteq \int_{\Omega_d^K} \hat{h}_{\alpha \gamma}(\omega', \omega) f(\omega') \nu(\omega') d\omega', \quad Af(\omega) \doteq \int_{\Omega_d^K} \hat{h}(\omega', \omega) f(\omega') \nu(\omega') d\omega'.$$  

It’s clear from (3.12) and (3.13) that $A = \sum_{\alpha=1}^K \sum_{\gamma=1}^K A_{\alpha \gamma}$. Denote by $I$ the identity operator on $L^2(\Omega_d^K, \nu)$. The following lemma will be proved in Section 3.3.

**Lemma 3.3.** (a) Trace$(AA^*) = \sum_{\alpha, \gamma=1}^K \frac{\lambda_\alpha}{\lambda_\gamma} \int_0^T \int_{\Omega_d \times \Omega_d} ||b_{\alpha \gamma, t}(X_{\ast, t}(\omega), X_{\ast, t}(\omega'))||^2 \nu_\alpha(\omega) \nu_\gamma(\omega') dt.$

(b) Trace$(A^n) = 0$ for all $n \geq 2$. (c) $I - A$ is invertible.

We can now present the main result of this chapter. For $\alpha \in K$, let $L^2_c(\Omega_d, \nu_\alpha)$ be the collection of $\phi \in L^2(\Omega_d, \nu_\alpha)$ such that $\int_{\Omega_d} \phi(\omega) \nu_\alpha(\omega) d\omega = 0$. Denote by $A_\alpha$ the collection of all measurable maps $\phi: \mathcal{C}_d \to \mathbb{R}$ such that $\phi(\omega) \in L^2_c(\Omega_d, \nu_\alpha)$. For $\phi \in A_\alpha$, let $\xi^N_{\alpha}(\phi) \doteq \frac{1}{\sqrt{N_\alpha}} \sum_{i \in N_\alpha} \phi(Z_i^N)$. 

19
Given $\alpha \in K$ and $\phi \in L^2(\Omega_\alpha, \nu_\alpha)$, define lifted function $\hat{\phi}_\alpha \in L^2(\Omega^K_\alpha, \hat{\nu})$ as follows:

$$\hat{\phi}_\alpha(\omega) \doteq \phi(\omega_\alpha), \quad \omega = (\omega_1, \ldots, \omega_K) \in \Omega^K_\alpha.$$  \hspace{1cm} (3.14)

**Theorem 3.4.** \{$\xi^N_\alpha(\phi) : \phi \in A_\alpha, \alpha \in K$\} converges as $N \to \infty$ to a mean zero Gaussian field \{$\xi(\phi) : \phi \in A_\alpha, \alpha \in K$\} in the sense of convergence of finite dimensional distributions, where for $\alpha, \gamma \in K$ and $\phi \in A_\alpha, \psi \in A_\gamma$,

$$E[\xi(\phi)\xi(\psi)] = \langle (I - A)^{-1} \hat{\phi}_\alpha, (I - A)^{-1} \hat{\psi}_\gamma \rangle_{L^2(\Omega^K_\alpha, \hat{\nu})},$$

with $\phi = \phi(X_\gamma)$, $\hat{\phi}_\alpha$ defined as in (3.14), and $\psi, \hat{\psi}_\gamma$ given similarly.

Proof of the theorem is given in Section 3.4.

### 3.2.2 An application to single-type particle system

Consider the special case of (3.3) where $K = 1$. Here, Theorem 3.4 can be used to describe joint asymptotic distributions of suitably scaled sums formed from disjoint sub-populations of the particle system. As an illustration, consider the single-type system given through the following collection of equations:

$$Z^{i,N}_t = Z^{i,N}_0 + \int_0^t f(s, Z^{i,N}_s) \, ds + \frac{1}{N} \int_0^t \sum_{j=1}^N b(Z^{i,N}_s, Z^{j,N}_s) \, ds + W^i_t, \quad i \in N. \hspace{1cm} (3.15)$$

Suppose that we are interested in the joint asymptotic distribution of $(\xi^N_1(\phi_1), \xi^N_2(\phi_2))$ for suitable path functionals $\phi_1$ and $\phi_2$, where for $\lambda \in (0, 1)$, $\xi^N_1(\phi_1), \xi^N_2(\phi_2)$ are as in (3.1). The propagation of chaos property in (3.5) says that any finite collection of summands on the right sides of the above display are asymptotically mutually independent. However, in general, this does not guarantee the asymptotic independence of $\xi^N_1(\phi_1)$ and $\xi^N_2(\phi_2)$. In fact, one can use Theorem 3.4 with $K = 2$ to show that $(\xi^N_1(\phi_1), \xi^N_2(\phi_2))$ converges in distribution to a bivariate Gaussian random variable and provide expressions for the asymptotic covariance matrix. We illustrate this below through a toy example.

**Example 3.1.** Suppose $Z^{i,N}_0 \equiv 0$, $f \equiv 0$, $T = 1$, $d = 1$ and $b(x, y) = \beta(y)$ for some bounded Lipschitz function $\beta$. Further suppose that $\beta$ is an odd function, namely $\beta(y) = -\beta(-y)$ for all
\( y \in \mathbb{R} \). Due to the special form of \( b \), one can explicitly characterize the measure \( \nu_i, i = 1, 2 \). Indeed, noting that for a one dimensional BM, \( \{W_t : t \in [0,1]\} \)

\[
W_t = \lambda \int_0^t \langle \beta, \mu_s \rangle \, ds + (1 - \lambda) \int_0^t \langle \beta, \mu_s \rangle \, ds + W_t, \quad \mu_s \doteq \mathcal{L}(W_s),
\]

we see that \( \nu_i = \mathcal{L}(W,W) \) for \( i = 1, 2 \). Consider for \( \omega \in \mathcal{C}_1 \),

\[
\phi_i(\omega) = \kappa_i \left( \omega_1 - \int_0^1 \beta(\omega_t) \, dt \right), \quad \kappa_i \in \mathbb{R}, \quad i = 1, 2.
\]

Note that \( \phi_i \in \mathcal{A}_i \) since \( \beta \) is odd, \( i = 1, 2 \). For this example one can explicitly describe the asymptotic distribution of \( (\xi_1^N(\phi_1), \xi_2^N(\phi_2)) \) by regarding (3.15) as a 2-type population with \( N_1 = \lceil N\lambda \rceil \) and \( N_2 = N - \lfloor N\lambda \rfloor \). Following (3.12) and (3.14), define for \( \omega = (\omega_1, \omega_2) \in \Omega_\beta^2 \),

\[
\hat{\phi}_i(\omega) = \kappa_i \left( X_{s,1}(\omega_1) - \int_0^1 \beta(X_{s,t}(\omega_t)) \, dt \right), \quad i = 1, 2,
\]

and for \( \omega' = (\omega'_1, \omega'_2) \in \Omega_\beta^2 \),

\[
\hat{h}(\omega', \omega) = \int_0^1 \left( \lambda \beta(X_{s,t}(\omega_1)) + \sqrt{\lambda(1-\lambda)} \beta(X_{s,t}(\omega_2)) \right) dW_{s,t}(\omega'_1) \\
+ \int_0^1 \left( \sqrt{\lambda(1-\lambda)} \beta(X_{s,t}(\omega_1)) + (1-\lambda) \beta(X_{s,t}(\omega_2)) \right) dW_{s,t}(\omega'_2).
\]

The operator \( A \) is then defined by (3.13). The special form of \( \phi_i \) allows us to determine \( (I - A)^{-1} \hat{\phi}_i \), \( i = 1, 2 \). Indeed, for \( \omega = (\omega_1, \omega_2) \in \Omega_\beta^2 \), let

\[
\psi_1(\omega) = \kappa_1 \left[ X_{s,1}(\omega_1) + \int_0^1 \left( \sqrt{\lambda(1-\lambda)} \beta(X_{s,t}(\omega_2)) - (1-\lambda) \beta(X_{s,t}(\omega_1)) \right) dt \right],
\]

\[
\psi_2(\omega) = \kappa_2 \left[ X_{s,1}(\omega_2) + \int_0^1 \left( \sqrt{\lambda(1-\lambda)} \beta(X_{s,t}(\omega_1)) - \lambda \beta(X_{s,t}(\omega_2)) \right) dt \right].
\]

Then

\[
A\psi_1(\omega) = \int_{\Omega_\beta^2} \hat{h}(\omega', \omega) \psi_1(\omega') \, d\nu'(\omega') \\
= \kappa_1 \int_0^1 \left( \lambda \beta(X_{s,t}(\omega_1)) + \sqrt{\lambda(1-\lambda)} \beta(X_{s,t}(\omega_2)) \right) dt
\]
and similarly
\[ A\psi_2(\omega) = \kappa_2 \int_0^1 \left( \sqrt{\lambda(1-\lambda)} \beta(X_{s,t}(\omega_1)) + (1-\lambda)\beta(X_{s,t}(\omega_2)) \right) dt. \]

This shows that \( \psi_i = (I - A)^{-1} \hat{\varphi}_i \) for \( i = 1, 2 \). From Theorem 3.4 we then have that 
\( (\xi^N_1(\phi_1), \xi^N_2(\phi_2)) \) converges in distribution to a bivariate Gaussian random variable \((X, Y)\) with mean zero and covariance matrix \((\sigma_{ij})_{i,j=1}^2\), where
\[
\begin{align*}
\sigma_{11} &\equiv \kappa_1^2 E \left[ \left( W_1 - (1-\lambda) \int_0^1 \beta(W_t) dt \right)^2 + \lambda(1-\lambda) \left( \int_0^1 \beta(W_t) dt \right)^2 \right], \\
\sigma_{22} &\equiv \kappa_2^2 E \left[ \left( W_1 - \lambda \int_0^1 \beta(W_t) dt \right)^2 + (1-\lambda) \left( \int_0^1 \beta(W_t) dt \right)^2 \right], \\
\sigma_{12} &\equiv \sqrt{\lambda(1-\lambda)} \kappa_1 \kappa_2 E \left[ (2W_1 - \int_0^1 \beta(W_t) dt) \int_0^1 \beta(W_t) dt \right],
\end{align*}
\]

and \( W \) is a one dimensional BM.

### 3.2.3 A limit theorem for statistics of multi-type populations

In this section we provide a CLT for statistics of the form given on the left side of (3.6).

Denote by \( L^2_{\nu}(\Omega^K_d, \hat{\nu}) \) the subspace of \( \phi \in L^2(\Omega^K_d, \hat{\nu}) \) such that for each \( \alpha \in K \),
\[
\int_{\Omega_d} \phi(\omega_1, \ldots, \omega_K) \nu_\alpha(d\omega_\alpha) = 0,
\]
for \( \nu_1 \otimes \cdots \otimes \nu_{\alpha-1} \otimes \nu_{\alpha+1} \otimes \cdots \otimes \nu_K \) a.e. \((\omega_1, \ldots, \omega_{\alpha-1}, \omega_{\alpha+1}, \ldots, \omega_K)\). Denote by \( A^K \) the collection of measurable maps \( \phi: \mathcal{C}^K_d \to \mathbb{R} \) such that for \( \omega = (\omega_1, \ldots, \omega_K) \in \mathcal{C}^K_d \), \( \omega \mapsto \phi(\omega) \equiv \phi(X_1(\omega_1), \ldots, X_K(\omega_K)) \in L^2(\Omega^K_d, \hat{\nu}) \). For \( \phi \in A^K \) and with \( \{Z^{i,N}\} \) as in (3.3), let \( \xi^N(\phi) \) be as in (3.2). We will like to study the asymptotic behavior of \( \xi^N(\phi) \) as \( N \to \infty \).

Let \( \{I_k(h), h \in L^2_{\text{sym}}(\mathbb{S}^k, \nu^{\otimes k})\}_{k \geq 1} \) be the MWI defined as in Section 2.1 with \( \mathbb{S} = \Omega^K_d \) and \( \nu = \hat{\nu} \), on some probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\nu})\). Recall \( \hat{h} \) defined in (3.12). Define \( f \in L^2(\Omega^K_d \times \Omega^K_d, \hat{\nu} \otimes \hat{\nu}) \) as follows: For \((\omega, \omega') \in \Omega^K_d \times \Omega^K_d \),
\[
f(\omega, \omega') = \hat{h}(\omega, \omega') + \hat{h}(\omega', \omega) - \int_{\Omega^K_d} \hat{h}(\omega'', \omega) \hat{h}(\omega'', \omega') \hat{\nu}(d\omega''). \quad (3.16)
\]
Define a random variable \( J \) on \((\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})\) as

\[
J = \frac{1}{2} [I_2(f) - \text{Trace}(AA^*)],
\]  

(3.17)

where \( A \) is as introduced in (3.13). For \( \phi \in L^2_c(\Omega^K_d, \nu) \), define lifted symmetric function \( \hat{\phi}^\text{sym} \in L^2((\Omega^K_d)^K, \nu^\otimes K) \) as follows: For \( \omega^j = (\omega_1^j, \ldots, \omega_K^j) \in \Omega^K_d, j \in K \),

\[
\hat{\phi}^\text{sym}(\omega^1, \omega^2, \ldots, \omega^K) \triangleq \frac{1}{K!} \sum_\pi \phi(\omega_{\pi(1)}^{(1)}, \omega_{\pi(2)}^{(2)}, \ldots, \omega_{\pi(K)}^{(K)}),
\]  

(3.18)

where the summation is taken over all permutations \( \pi \) on \( K \). The following result characterizes the asymptotic distribution of \( \xi^N(\phi) \).

**Theorem 3.5.** We have \( E_{\tilde{\mathbb{P}}} \exp(J) = 1 \). Denote by \( \tilde{\mathbb{Q}} \) the probability measure on \((\bar{\Omega}, \bar{\mathcal{F}})\) such that \( d\tilde{\mathbb{Q}} = \exp(J) \cdot d\tilde{\mathbb{P}} \). Then for \( \phi \in \mathcal{A}^K \), \( L(\xi^N(\phi)) \to \tilde{\mathbb{Q}} \circ (I_K(\hat{\phi}^\text{sym}))^{-1} \) as \( N \to \infty \), where \( \phi = \phi(X_s(\cdot), \ldots, X_s(\cdot)) \) and \( \hat{\phi}^\text{sym} \) is defined as in (3.18).

Proof of the theorem is given in Section 3.5.

**Remark 3.2.** We believe that in Theorems 3.4 and 3.5 the assumption that diffusion coefficients in (3.3) are identity matrices can be relaxed as follows: Replace \( W^i_t \) in (3.3) by \( \int_0^t \Sigma_{\alpha}(s, Z^i_{s,N}) \, d\tilde{W}^i_s + \int_0^t \sigma_{\alpha}(s, Z^i_{s,N}) \, d\bar{W}^i_s \), where \( \Sigma_{\alpha} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times d} \) and \( \sigma_{\alpha} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times r} \) are bounded Lipschitz (in the second variable, uniformly in the first variable) functions such that \( \Sigma_{\alpha} \) is invertible and \( \Sigma_{\alpha}^{-1} \) is bounded. Here \( \{\tilde{W}^i, i \in \mathbf{N}\} \) and \( \{\bar{W}^i, i \in \mathbf{N}\} \) are mutually independent \( d \) and \( r \) dimensional BM respectively defined on \((\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})\). An approach for the proof of the central limit theorem under this weaker condition will be to take the change of measure introduced in Section 3.4 to be of a slightly different form. Roughly speaking, the change of measure will be such that under the new measure \( \{\tilde{W}^i, i \in \mathbf{N}\} \) are still independent Brownian motions and suitable absolutely continuous translations of \( \{\tilde{W}^i, i \in \mathbf{N}\} \) are independent Brownian motions that are independent of \( \{\bar{W}^i, i \in \mathbf{N}\} \). Specifically in defining the change of measure the function \( b_{\alpha \gamma,t} \) appearing in (3.23) and (3.24) (see (3.10)) is replaced by

\[
b_{\alpha \gamma,t}(x, y) \triangleq \Sigma_{\alpha}^{-1}(t, x) \left( b_{\alpha \gamma}(x, y) - \langle b_{\alpha \gamma}(x, \cdot), \mu^\gamma_t \rangle \right), \quad (x, y) \in \mathbb{R}^d \times \mathbb{R}^d, \quad \alpha, \gamma \in K,
\]
and \( \{W^i, i \in \mathbb{N}\} \) is replaced by \( \{\hat{W}^i, i \in \mathbb{N}\} \).

### 3.3 Proof of Lemma 3.3

Recall the definition of \( b_{\alpha \gamma, t}, \hat{b}_{\alpha \gamma} \) and \( \hat{h} \) in (3.10) and (3.12). Let \( \Omega_d, X, \{\nu_{\alpha}\}_{\alpha \in \mathbb{K}}, \hat{\nu} \) be as in Section 3.2.1. Define for \( \alpha, \beta, \gamma \in \mathbb{K} \), the functions \( b_{\alpha \beta \gamma, t} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \) and \( l_{\alpha, \beta \gamma} \) from \( \Omega \times \Omega \) to \( \mathbb{R} (\nu_{\beta} \otimes \nu_{\gamma} \text{ a.s.}) \) as follows:

\[
b_{\alpha \beta \gamma, t}(x, y) = \int_{\mathbb{R}^d} b_{\alpha \beta, t}(z, x) \cdot b_{\alpha \gamma, t}(z, y) \mu_t(dz), \quad (x, y) \in \mathbb{R}^d \times \mathbb{R}^d, \tag{3.19}
\]

\[
l_{\alpha, \beta \gamma} = \frac{\lambda_{\alpha}}{\sqrt{\lambda_{\beta} \lambda_{\gamma}}} \int_0^T b_{\alpha \beta \gamma, t}(X_{\ast, t}^\beta(\omega), X_{\ast, t}^\beta(\omega')) dt, \quad (\omega, \omega') \in \Omega \times \Omega. \tag{3.20}
\]

Also define lifted function \( \hat{l}_{\alpha, \beta \gamma} \) from \( \Omega^K \times \Omega^K \) to \( \mathbb{R} (\hat{\nu} \otimes \hat{\nu} \text{ a.s.}) \) as

\[
\hat{l}_{\alpha, \beta \gamma}(\omega, \omega') = l_{\alpha, \beta \gamma}(\omega', \omega), \quad \omega = (\omega_1, \ldots, \omega_K) \in \Omega_d^K, \omega' = (\omega'_1, \ldots, \omega'_K) \in \Omega_d^K. \tag{3.21}
\]

It is easily seen that for \( \alpha, \alpha', \beta, \gamma \in \mathbb{K} \),

\[
\int_{\Omega_d^K} \hat{h}_{\alpha \beta}(\omega'', \omega) \hat{h}_{\alpha' \gamma}(\omega'', \omega') \hat{\nu}(d\omega'') = 1_{\{\alpha = \alpha'\}} \hat{l}_{\alpha, \beta \gamma}(\omega, \omega'). \tag{3.22}
\]

Recall the definition of the integral operator \( A_{\alpha \gamma} \) for \( \alpha, \gamma \in \mathbb{K} \) in (3.13), and note that its adjoint \( A^*_{\alpha \gamma} : L^2(\Omega_d^K, \hat{\nu}) \to L^2(\Omega_d^K, \hat{\nu}) \) is given as follows: For \( f \in L^2(\Omega_d^K, \hat{\nu}) \) and \( \omega \in \Omega_d^K \),

\[
A^*_{\alpha \gamma} f(\omega) = \int_{\Omega_d^K} \hat{h}_{\alpha \gamma}(\omega, \omega') f(\omega') \hat{\nu}(d\omega').
\]

Then for \( \alpha, \alpha', \beta, \gamma \in \mathbb{K} \), the operator \( A_{\alpha \beta} A^*_{\alpha' \gamma} : L^2(\Omega_d^K, \hat{\nu}) \to L^2(\Omega_d^K, \hat{\nu}) \) is given as follows: For \( f \in L^2(\Omega_d^K, \hat{\nu}) \) and \( \omega \in \Omega_d^K \),

\[
A_{\alpha \beta} A^*_{\alpha' \gamma} f(\omega) = \int_{\Omega_d^K} \left( \int_{\Omega_d^K} \hat{h}_{\alpha \beta}(\omega', \omega) \hat{h}_{\alpha' \gamma}(\omega', \omega'') \hat{\nu}(d\omega') \right) f(\omega'') \hat{\nu}(d\omega'').
\]

\[
= 1_{\{\alpha = \alpha'\}} \int_{\Omega_d^K} \hat{l}_{\alpha, \beta \gamma}(\omega, \omega'') f(\omega'') \hat{\nu}(d\omega''),
\]

24
where the last equality follows from (3.22). In particular, $A_{\alpha\beta}A_{\alpha\gamma}^*$ = 0 if $\alpha \neq \alpha'$. Moreover, for $\alpha, \gamma, \gamma' \in K$,

$$\text{Trace}(A_{\alpha\gamma}^* A_{\alpha\gamma'}) = \int_{\Omega^K \times \Omega^K_d} \hat{h}_{\alpha\gamma'}(\omega, \omega')\hat{h}_{\alpha\gamma}(\omega, \omega') \nu(d\omega) \nu(d\omega')$$

$$= \frac{\lambda_\alpha}{\sqrt{\lambda_\gamma \lambda_\gamma'}} \int_{\Omega^K \times \Omega^K_d} \int_0^T b_{\alpha\gamma',t}(X_{s,t}(\omega), X_{s,t}(\omega')) \cdot b_{\alpha\gamma,t}(X_{s,t}(\omega), X_{s,t}(\omega')) dt \nu(d\omega) \nu(d\omega')$$

$$= 1_{\{\gamma' = \gamma\}} \frac{\lambda_\alpha}{\lambda_\gamma} \int_0^T \int_{\Omega_d \times \Omega_d} \|b_{\alpha\gamma,t}(X_{s,t}(\omega), X_{s,t}(\omega'))\|^2 \nu_\alpha(d\omega) \nu_\gamma(d\omega') dt.$$

In particular, $\text{Trace}(A_{\alpha\gamma}^* A_{\alpha\gamma'}) = 0$ if $\gamma \neq \gamma'$. Hence we have

$$\text{Trace}(AA^*) = \text{Trace}\left(\sum_{\alpha=1}^K \sum_{\gamma=1}^K A_{\alpha\gamma}\left(\sum_{\gamma'=1}^K A_{\alpha\gamma'}^*\right)\right) = \sum_{\alpha=1}^K \sum_{\gamma=1}^K \text{Trace}(A_{\alpha\gamma}^* A_{\alpha\gamma})$$

$$= \sum_{\alpha,\gamma=1}^K \frac{\lambda_\alpha}{\lambda_\gamma} \int_0^T \int_{\Omega_d \times \Omega_d} \|b_{\alpha\gamma,t}(X_{s,t}(\omega), X_{s,t}(\omega'))\|^2 \nu_\alpha(d\omega) \nu_\gamma(d\omega') dt,$$

which proves part (a) of Lemma 3.3. Noting that

$$\text{Trace}(A^n) = \int_{\Omega^n_d \times \Omega^n_d \times \cdots \times \Omega^n_d} \hat{h}(\omega_1, \omega_2) \hat{h}(\omega_2, \omega_3) \cdots \hat{h}(\omega_n, \omega_1) \nu(d\omega_1) \nu(d\omega_2) \cdots \nu(d\omega_n),$$

part (b) follows from the proof of Lemma 2.7 of [62]. Part (c) is now immediate from Lemma 1.3 of [62] (cf. Lemma 2.2).

### 3.4 Proof of Theorem 3.4

For $N \in \mathbb{N}$, recall the canonical processes $X^i, W^i, V^i$ and probability measure $\mathbb{P}^N$ introduced in Section 3.2. For $t \in [0, T]$, define

$$J^N(t) = J^{N,1}(t) - \frac{1}{2} J^{N,2}(t),$$

where

$$J^{N,1}(t) = \sum_{\alpha=1}^K \sum_{i \in \mathcal{N}_\alpha} \int_0^t \sum_{\gamma=1}^K \sum_{j \in \mathcal{N}_\gamma} \frac{1}{N} b_{\alpha\gamma,s}(X^i_s, X^j_s) \cdot dW^i_s$$

(3.23)

$$J^{N,2}(t) = \sum_{\alpha=1}^K \sum_{i \in \mathcal{N}_\alpha} \int_0^t \sum_{\gamma=1}^K \sum_{\gamma'=1}^K \frac{1}{N} b_{\alpha\gamma,s}(X^i_s, X^j_s) \cdot dW^i_s$$

(3.24)
and
\[
J^{N,2}(t) \doteq \sum_{\alpha=1}^{K} \sum_{i \in \mathcal{N}_\alpha} \int_0^t \left\| \sum_{\gamma=1}^{K} \frac{1}{N_\gamma} \sum_{j \in \mathcal{N}_\gamma} b_{\alpha\gamma,s}(X^i_s, X^j_s) \right\|^2 ds.
\] (3.24)

Let \( \mathcal{F}^N_t \doteq \sigma\{V^i(s), 0 \leq s \leq t, i \in \mathbf{N}\} \). Note that \( \{\exp(J^N(t))\} \) is an \( \{\mathcal{F}^N_t\}\) martingale under \( \mathbb{P}^N \). Define a new probability measure \( Q^N \) on \( \Omega^N_d \) by
\[
dQ^N_d P^N = \exp(J^N(T)).
\]

By Girsanov’s theorem, \((X^1, \ldots, X^N)\) has the same probability distribution under \( Q^N \) as \((Z^1, \ldots, Z^N, \ldots)\) under \( P \). For \( \phi \in \mathcal{A}_\alpha, \alpha \in K \), let
\[
\tilde{\xi}^N_\alpha(\phi) \doteq \frac{1}{\sqrt{N_\alpha}} \sum_{i \in \mathcal{N}_\alpha} \phi(X^i).
\] (3.25)

Thus in order to prove the theorem it suffices to show that for any \( \phi^{(\alpha)} \in \mathcal{A}_\alpha, \alpha \in K \),
\[
\lim_{N \to \infty} E_{Q^N} \exp\left( i \sum_{\alpha \in K} \tilde{\xi}^N_\alpha(\phi^{(\alpha)}) \right) = \exp\left( -\frac{1}{2} \left\| (I - A)^{-1} \sum_{\alpha \in K} \tilde{\phi}^{(\alpha)}_\alpha \right\|_{L^2(\Omega^K_d, \hat{\nu})}^2 \right),
\]
where as in Section 3.2.1, \( \phi^{(\alpha)} \doteq \phi^{(\alpha)}(X^\ast) \), and \( \tilde{\phi}^{(\alpha)}_\alpha \) is defined by (3.14), replacing \( \phi \) with \( \phi^{(\alpha)} \).

This is equivalent to showing
\[
\lim_{N \to \infty} E_{P^N} \exp\left( i \sum_{\alpha \in K} \tilde{\xi}^N_\alpha(\phi^{(\alpha)}) + J^{N,1}(T) - \frac{1}{2} J^{N,2}(T) \right)
\]
\[
= \exp\left( -\frac{1}{2} \left\| (I - A)^{-1} \sum_{\alpha \in K} \tilde{\phi}^{(\alpha)}_\alpha \right\|_{L^2(\Omega^K_d, \hat{\nu})}^2 \right).
\] (3.26)

In order to prove (3.26), we will need to study the asymptotics of \( J^{N,1} \) and \( J^{N,2} \) as \( N \to \infty \). The proof of (3.26) is completed in Section 3.4.3. We begin by studying a generalization of Theorem 2.1 to the case of \( K \) populations.

### 3.4.1 Asymptotics of statistics of multi-type populations of independent particles

Throughout this subsection, let \( \{H^i \doteq (H^i_1, H^i_2, \ldots, H^i_K)\}_{i \geq 1} \) be a sequence of i.i.d. \( \Omega^K_d \)-valued random variables with law \( \hat{\nu} = \nu_1 \otimes \cdots \otimes \nu_K \). We introduce the following notion of lifted functions and lifted symmetric functions. Given \( \alpha, \gamma \in K \) and \( \psi_{\alpha\gamma} \in L^2(\Omega_d \times \Omega_d, \nu_\alpha \otimes \nu_\gamma) \), define \( \tilde{\psi}_{\alpha\gamma} \) and
\( \hat{\psi}_{\alpha\gamma} \) in \( L^2(\Omega_{\upsilon_d}^K \times \Omega_{\upsilon_d}^K, \upsilon \otimes \upsilon) \) as follows: For \( \omega = (\omega_1, \ldots, \omega_K) \) and \( \omega' = (\omega'_1, \ldots, \omega'_K) \) in \( \Omega_{\upsilon_d}^K \),

\[
\hat{\psi}_{\alpha\gamma}(\omega, \omega') = \psi_{\alpha\gamma}(\omega_\alpha, \omega'_\gamma), \quad \hat{\psi}_{\alpha\gamma}^\text{sym}(\omega, \omega') = \frac{1}{2}(\hat{\psi}_{\alpha\gamma}(\omega, \omega') + \hat{\psi}_{\alpha\gamma}(\omega', \omega)).
\] (3.27)

Recall \( \hat{\phi}^\text{sym} \in L^2((\Omega_{\upsilon_d}^K)^K, \upsilon^{\otimes K}) \) defined in (3.18) for \( \phi \in L^2(\Omega_{\upsilon_d}^K, \upsilon) \).

Recall the definition of \( L_c^2(\Omega_d, \nu_\alpha) \) and \( L_c^2(\Omega_{\upsilon_d}^K, \upsilon) \) in Section 3.2.1 and 3.2.3, respectively. Also, for \( \alpha, \gamma \in K \), denote by \( L_c^2(\Omega_d \times \Omega_d, \nu_\alpha \otimes \nu_\gamma) \) the subspace of \( \psi \in L^2(\Omega_d \times \Omega_d, \nu_\alpha \otimes \nu_\gamma) \) such that

\[
\int_{\Omega_d} \psi(\omega', \omega) \nu_\alpha(d\omega') = 0, \nu_\gamma \text{ a.e. } \omega \text{ and } \int_{\Omega_d} \psi(\omega, \omega') \nu_\gamma(d\omega') = 0, \nu_\alpha \text{ a.e. } \omega.
\]

The following lemma gives asymptotic distribution of certain statistics of the \( K \)-type population introduced above.

**Lemma 3.6.** Let \( \{\phi_\alpha, \psi_{\alpha\gamma}, \phi: \alpha, \gamma \in K\} \) be a collection of functions such that, \( \phi_\alpha \in L_c^2(\Omega_d, \nu_\alpha) \) for each \( \alpha \in K \), \( \psi_{\alpha\gamma} \in L_c^2(\Omega_d \times \Omega_d, \nu_\alpha \otimes \nu_\gamma) \) for each pair of \( \alpha, \gamma \in K \), and \( \phi \in L_c^2(\Omega_{\upsilon_d}^K, \upsilon) \). For each \( N \in \mathbb{N} \), let

\[
\zeta^N \doteq (\zeta_1^N, \zeta_2^N, \zeta_3^N, \zeta_4^N), \quad \eta \doteq (\eta_1, \eta_2, \eta_3, \eta_4),
\]

where

\[
\zeta_1^N \doteq \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N \psi_{\alpha\alpha}(H_{\alpha i}^i) \right)_{\alpha=1}^K, \quad \zeta_2^N \doteq \left( \frac{1}{N} \sum_{1 \leq i \neq j \leq N} \psi_{\alpha\alpha}(H_{\alpha i}^i, H_{\alpha j}^j) \right)_{\alpha=1}^K,
\]

\[
\zeta_3^N \doteq \left( \frac{1}{N} \sum_{i,j=1}^N \psi_{\alpha\gamma}(H_{\alpha i}^i, H_{\gamma j}^j) \right)_{1 \leq \alpha \neq \gamma \leq K}, \quad \zeta_4^N \doteq \left( \frac{1}{NK/2} \sum_{i_1, \ldots, i_K=1}^N \phi(H_{i_1}^1, \ldots, H_{i_K}^K) \right),
\]

and

\[
\eta_1 \doteq (I_1(\hat{\phi}_\alpha))_{\alpha=1}^K, \quad \eta_2 \doteq (I_2(\hat{\psi}_{\alpha\alpha}^\text{sym}))_{\alpha=1}^K, \quad \eta_3 \doteq (I_2(\hat{\psi}_{\alpha\gamma}^\text{sym}))_{1 \leq \alpha \neq \gamma \leq K}, \quad \eta_4 \doteq I_K(\hat{\phi}^\text{sym}).
\]

Here \( \{I_k\}_{k \geq 1} \) are as defined in Section 3.2.3. Then \( \zeta^N \Rightarrow \eta \) as a sequence of \( \mathbb{R}^q \) valued random variables, as \( N \to \infty \), where \( q \doteq K + K + K(K - 1) + 1 \).
Proof. Note that for $\alpha \in K$, we have

$$
\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \phi_{\alpha}(H_{\alpha}^{i}) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \hat{\phi}_{\alpha}(H^{i}),
$$

$$
\frac{1}{N} \sum_{1 \leq i \neq j \leq N} \psi_{\alpha\gamma}(H_{\alpha}^{i}, H_{\gamma}^{j}) = \frac{2}{N} \sum_{1 \leq i < j \leq N} \hat{\psi}_{\alpha\gamma}(H^{i}, H^{j}).
$$

Also, for $\alpha, \gamma \in K$ such that $\alpha \neq \gamma$, we have

$$
\frac{1}{N} \sum_{i,j=1}^{N} \psi_{\alpha\gamma}(H_{\alpha}^{i}, H_{\gamma}^{j}) = \frac{2}{N} \sum_{1 \leq i < j \leq N} \hat{\psi}_{\alpha\gamma}(H^{i}, H^{j}) + \frac{1}{N} \sum_{i=1}^{N} \psi_{\alpha\gamma}(H_{\alpha}^{i}, H_{\gamma}^{i}).
$$

Let $S = \{(i_{1}, \ldots, i_{K}) \in N^{K} : i_{1}, \ldots, i_{K} \text{ are distinct}\}$. Then we have

$$
\frac{1}{N^{K/2}} \sum_{i_{1}, \ldots, i_{K}=1}^{N} \phi(H_{1}^{i_{1}}, \ldots, H_{K}^{i_{K}})
$$

$$
= \frac{1}{N^{K/2}} \sum_{(i_{1}, \ldots, i_{K}) \in S} \phi(H_{1}^{i_{1}}, \ldots, H_{K}^{i_{K}}) + \frac{1}{N^{K/2}} \sum_{(i_{1}, \ldots, i_{K}) \notin S} \phi(H_{1}^{i_{1}}, \ldots, H_{K}^{i_{K}})
$$

$$
= \frac{K!}{N^{K/2}} \sum_{1 \leq i_{1} < \cdots < i_{K} \leq N} \hat{\phi}_{\text{sym}}(H_{1}^{i_{1}}, \ldots, H_{K}^{i_{K}}) + \frac{1}{N^{K/2}} \sum_{(i_{1}, \ldots, i_{K}) \notin S} \phi(H_{1}^{i_{1}}, \ldots, H_{K}^{i_{K}}).
$$

Combining above results gives us

$$
\zeta^{N} = \tilde{\zeta}^{N} + R_{N,1} = (\zeta_{1}^{N}, \zeta_{2}^{N}, \zeta_{3}^{N}, \zeta_{4}^{N}) + (0_{K \times 1}, 0_{K \times 1}, R_{3}^{N,1}, R_{4}^{N,1}),
$$

(3.28)

where

$$
\tilde{\zeta}_{1}^{N} = \left( \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \hat{\phi}_{\alpha}(H^{i}) \right)_{\alpha=1}^{K}, \quad \tilde{\zeta}_{2}^{N} = \left( \frac{2}{N} \sum_{1 \leq i < j \leq N} \hat{\psi}_{\alpha\gamma}(H^{i}, H^{j}) \right)_{\alpha=1}^{K},
$$

$$
\tilde{\zeta}_{3}^{N} = \left( \frac{2}{N} \sum_{1 \leq i < j \leq N} \hat{\psi}_{\alpha\gamma}(H^{i}, H^{j}) \right)_{1 \leq \alpha \neq \gamma \leq K}, \quad \tilde{\zeta}_{4}^{N} = \frac{K!}{N^{K/2}} \sum_{1 \leq i_{1} < \cdots < i_{K} \leq N} \hat{\phi}_{\text{sym}}(H_{1}^{i_{1}}, \ldots, H_{K}^{i_{K}}),
$$

and

$$
R_{3}^{N,1} = \left( \frac{1}{N} \sum_{i=1}^{N} \psi_{\alpha\gamma}(H_{\alpha}^{i}, H_{\gamma}^{i}) \right)_{1 \leq \alpha \neq \gamma \leq K}, \quad R_{4}^{N,1} = \frac{1}{N^{K/2}} \sum_{(i_{1}, \ldots, i_{K}) \notin S} \phi(H_{1}^{i_{1}}, \ldots, H_{K}^{i_{K}}).
$$
From Theorem 2.1 it follows that as $N \to \infty$,

$$\tilde{\zeta}^N \Rightarrow \eta.$$  \hspace{1cm} (3.29)

By law of large numbers, we have $\mathcal{R}^{N,1}_3 \to 0_{K(K-1)\times 1}$ in probability as $N \to \infty$. Note that as $N \to \infty$,

$$\mathbb{E}(\mathcal{R}^{N,1}_3)^2 = \frac{1}{N^K} \sum_{(i_1,\ldots,i_K) \notin \mathcal{S}} \mathbb{E}[\phi(H_{i_1}^{i_K})]^2 \to 0,$$

since $|N^K \setminus \mathcal{S}| = N^K - N(N-1) \cdots (N-K+1) = o(N^K)$. Hence

$$\mathcal{R}^{N,1}_3 \to 0_{q \times 1}$$ \hspace{1cm} (3.30)

in probability as $N \to \infty$. The desired result follows by combining (3.28) – (3.30).

For studying the asymptotic behavior of $J^{N,1}$ and $J^{N,2}$, we will need an extension of Lemma 3.6 to a setting where the numbers of particles of different types may differ. Before discussing this extension, we present an elementary lemma (cf. [4], Theorem 3.2) and the proof is omitted.

**Lemma 3.7.** For $m,n \in \mathbb{N}$, let $\xi_{mn}, \xi_n$ be $\mathbb{R}^d$-valued random variables defined on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$, and $\eta_m, \eta$ be $\mathbb{R}^d$-valued random variables defined on some other probability space $(\Omega'_0, \mathcal{F}'_0, \mathbb{P}'_0)$. If for each $m \geq 1$, $\xi_{mn} \Rightarrow \eta_m$ as $n \to \infty$ and the following condition holds:

$$\lim_{m \to \infty} \sup_{n \geq 1} \mathbb{E}_{\mathbb{P}_0}(\|\xi_{mn} - \xi_n\| \wedge 1) = 0, \quad \lim_{m \to \infty} \mathbb{E}_{\mathbb{P}'_0}(\|\eta_m - \eta\| \wedge 1) = 0,$$

then $\xi_n \Rightarrow \eta$ as $n \to \infty$.

We now study the following extension of Lemma 3.6.

**Lemma 3.8.** For $\alpha \in K$, let $N_\alpha$ be a function of $N$ such that $\lim_{N \to \infty} N_\alpha = \infty$. Let $\{\phi_\alpha, \psi_{\alpha \gamma}, \phi, \alpha, \gamma \in K\}$, $\eta$ and $q$ be as in Lemma 3.6. For each $N \in \mathbb{N}$, let

$$\xi^N = (\xi_1^N, \xi_2^N, \xi_3^N, \xi_4^N),$$
where

\[ \xi_1^N = \left( \frac{1}{\sqrt{N_\alpha}} \sum_{i=1}^{N_\alpha} \phi_\alpha(H^1_\alpha) \right)_{\alpha=1}^K, \]

\[ \xi_2^N = \left( \frac{1}{N_\alpha} \sum_{1 \leq i \neq j \leq N_\alpha} \psi_{\alpha\alpha}(H^i_\alpha, H^j_\alpha) \right)_{\alpha=1}^K, \]

\[ \xi_3^N = \left( \frac{1}{\sqrt{N_\alpha N_\gamma}} \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_\gamma} \psi_{\alpha\gamma}(H^i_\alpha, H^j_\gamma) \right)_{1 \leq \alpha \neq \gamma \leq K}, \]

\[ \xi_4^N = \frac{1}{\sqrt{N_1 \cdots N_K}} \sum_{i_1=1}^{N_1} \cdots \sum_{i_K=1}^{N_K} \phi(H^{i_1}_1, \ldots, H^{i_K}_K). \]

Then \( \xi^N \to \eta \) as a sequence of \( \mathbb{R}^d \) valued random variables, as \( N \to \infty \).

**Proof.** For each \( \beta = 1, \ldots, K \), let \( \{1_{\Omega_\beta}, e^1_{\beta}, e^2_{\beta}, \ldots \} \) be a complete orthonormal system (CONS) in \( L^2(\Omega_\beta, \nu_\beta) \). Note that \( \langle \phi_\alpha, 1_{\Omega_\beta} \rangle_{L^2(\Omega_\beta, \nu_\alpha)} = 0 \) for all \( \alpha \in K \), and analogous orthogonality properties with the function \( 1_{\Omega_\beta} \) hold for \( \psi_{\alpha\gamma} \) and \( \phi \), for all \( \alpha, \gamma \in K \). (For example, \( \langle \psi_{\alpha\gamma}, 1_{\Omega_\beta} \otimes e^j_\gamma \rangle_{L^2(\Omega_\beta \times \Omega_\gamma, \nu_\alpha \otimes \nu_\gamma)} = 0 \) for all \( \alpha, \gamma \in K \) and \( j \geq 1 \)). Because of this, \( \psi^M_{\alpha\gamma} \to \psi_{\alpha\gamma} \) in \( L^2(\Omega_\beta \times \Omega_\gamma, \nu_\alpha \otimes \nu_\gamma) \) and \( \phi^M \to \phi \) in \( L^2(\Omega^K_d, \hat{\nu}) \), as \( M \to \infty \), where

\[ \psi^M_{\alpha\gamma}(\omega, \omega') = \sum_{m_\alpha=1}^{M} \cdots \sum_{m_\gamma=1}^{M} c^m_{\alpha\gamma} e^m_\alpha(\omega) e^m_\gamma(\omega'), \quad (\omega, \omega') \in \Omega_\beta \times \Omega_\gamma, \quad \alpha, \gamma \in K, \]

\[ \phi^M(\omega_1, \ldots, \omega_K) = \sum_{m_1=1}^{M} \cdots \sum_{m_K=1}^{M} e^{m_1 \cdots m_K}(\omega_1, \omega_1) \cdots e^{m_K}(\omega_K), \quad (\omega_1, \ldots, \omega_K) \in \Omega^K_d, \]

and \( c^m_{\alpha\gamma} = \langle \psi_{\alpha\gamma}, e^m_\alpha \otimes e^m_\gamma \rangle_{L^2(\Omega_\beta \times \Omega_\gamma, \nu_\alpha \otimes \nu_\gamma)} ) \). It follows that as \( M \to \infty \), \( \hat{\psi}^M_{\alpha\gamma} \to \hat{\psi}_{\alpha\gamma}^{sym} \) in \( L^2(\Omega^K_d, \hat{\nu} \otimes \hat{\nu}) \) and \( \hat{\phi}^M \to \hat{\phi} \) in \( L^2((\Omega^K_d)^K, \hat{\nu} \otimes \hat{\nu}) \). For \( M, N \in \mathbb{N} \), let

\[ \xi^{MN} = (\xi_1^N, \xi_2^N, \xi_3^{MN}, \xi_4^{MN}), \quad \eta^M = (\eta_1, \eta_2, \eta_3^M, \eta_4^M), \]

where \( \xi_3^{MN}, \xi_4^{MN} \) are defined as \( \xi_3^N, \xi_4^N \) but with \( \psi^M_{\alpha\gamma} \) and \( \phi^M \) instead of \( \psi_{\alpha\gamma} \) and \( \phi \), and

\[ \eta_3^M = (I_2(\hat{\psi}^{M,sym}_{\alpha\gamma}))_{1 \leq \alpha \neq \gamma \leq K}, \quad \eta_4^M = I_K(\hat{\phi}^{M,sym}). \]
From Lemma 3.7, in order to finish the proof, it suffices to check the following three properties hold:

\[
\lim_{M \to \infty} \sup_{N \geq 1} E\|\xi^{MN} - \xi^N\|^2 = 0, \quad \lim_{M \to \infty} E\|\eta^M - \eta\|^2 = 0, \quad \xi^{MN} \Rightarrow \eta^M \text{ as } N \to \infty. \tag{3.31}
\]

Note that the first two coordinates of \(\xi^{MN}\) and \(\eta^M\) are the same as those of \(\xi^N\) and \(\eta\), respectively. So for the first two statements in (3.31), we only need to check for the third and last coordinates. Consider the first statement.

\[
\lim_{M \to \infty} \sup_{N \geq 1} E\left(\phi^M(H_1^i, \ldots, H_K^i) - \phi(H_1^i, \ldots, H_K^i)\right)^2 = 0.
\]

Similarly, for each pair of \(\alpha, \gamma \in K\) such that \(\alpha \neq \gamma\),

\[
\lim_{M \to \infty} \sup_{N \geq 1} E\left(\frac{1}{\sqrt{N_\alpha N_\gamma}} \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_\gamma} \psi_{\alpha\gamma}^M(H^{i_1}, H^{i_2}) - \frac{1}{\sqrt{N_\alpha N_\gamma}} \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_\gamma} \psi_{\alpha\gamma}^M(H^{i_1}, H^{i_2})\right)^2 = 0.
\]

This proves the first statement in (3.31). Recalling the isometry property of MWI (2.3), we have for each pair of \(\alpha, \gamma \in K\) such that \(\alpha \neq \gamma\),

\[
\lim_{M \to \infty} E\left(I_2(\hat{\psi}_{\alpha\gamma}^{M,sym}) - I_2(\hat{\psi}_{\alpha\gamma}^{sym})\right)^2 = \lim_{M \to \infty} 2 \left\|\hat{\psi}_{\alpha\gamma}^{M,sym} - \hat{\psi}_{\alpha\gamma}^{sym}\right\|^2_{L^2(\Omega^K, \hat{\nu} \otimes \hat{\nu})} = 0.
\]

Thus the second statement in (3.31) holds with \(\eta^M\) and \(\eta\) replaced by \(\eta_3^M\) and \(\eta_3\), respectively. Similarly, for \(\eta_4^{M}\), we have

\[
\lim_{M \to \infty} E(\eta_4^M - \eta_4)^2 = \lim_{M \to \infty} K! \left\|\hat{\phi}_{\alpha\gamma}^{M,sym} - \hat{\phi}_{\alpha\gamma}^{sym}\right\|^2_{L^2(\Omega^K \times \Omega^K, \hat{\nu} \otimes \hat{\nu})} = 0.
\]

Combining the above observations, we have the second statement in (3.31).
Finally we check the third statement in (3.31). Let \( \{ \tilde{I}_i^n(h), h \in L^2(\nu_\alpha) \}_{\alpha=1}^K \) be \( K \) mutually independent Gaussian fields such that

\[
\mathbf{E}(\tilde{I}_i^n(h)) = 0, \quad \mathbf{E}(\tilde{I}_i^n(h)\tilde{I}_i^n(g)) = \langle h, g \rangle_{L^2(\nu_\alpha)} \text{ for all } h, g \in L^2(\nu_\alpha), \alpha \in K.
\]

Let \( \tilde{\eta}^M = (\tilde{\eta}_1, \tilde{\eta}_2, \tilde{\eta}_3^M, \tilde{\eta}_4^M) \), where

\[
\tilde{\eta}_1 \doteq (\tilde{I}_1^n(\hat{\phi}_\alpha))_{\alpha=1}^K, \quad \tilde{\eta}_2 \doteq (\tilde{I}_2^n(\hat{\psi}_{\alpha\gamma}))_{\alpha=1}^K,
\]

\[
\tilde{\eta}_3^M \doteq \left( \sum_{m_1=1}^M \cdots \sum_{m_K=1}^M c_{m_1}^{m_K} \tilde{I}_1^n(e_{m_1}) \cdots \tilde{I}_K^n(e_{m_K}) \right)_{1 \leq \alpha \neq \gamma \leq K},
\]

\[
\tilde{\eta}_4^M \doteq \sum_{m_1=1}^M \cdots \sum_{m_K=1}^M c_{m_1}^{m_K} \tilde{I}_1^n(e_{m_1}) \cdots \tilde{I}_K^n(e_{m_K}),
\]

and \( \tilde{I}_2^n \) is as defined below (2.2) by replacing \( I_1 \) there with \( \tilde{I}_1^n \). Theorem 2.1 and mutual independence of \( \{ H_i^n: i = 1, \ldots, N_\alpha, \alpha = 1, \ldots, K \} \) imply that for each \( M \geq 1 \), as \( N \to \infty \), \( \xi^{MN} \Rightarrow \tilde{\eta}^M \).

In order to verify the third condition in (3.31), it now suffices to show \( \tilde{\eta}^M \) and \( \eta^M \) have the same probability distribution. However, this follows easily by considering the asymptotic behavior of

\[
\zeta^{MN} \doteq (\zeta_1^N, \zeta_2^N, \zeta_3^MN, \zeta_4^{MN}),
\]

where \( \zeta_1^N \) and \( \zeta_2^N \) are as in Lemma 3.6, and \( \zeta_3^{MN} \) [resp. \( \zeta_4^{MN} \)] are also as in Lemma 3.6 but with \( \psi_{\alpha\gamma} \) [resp. \( \phi \)] replaced with \( \hat{\psi}_{\alpha\gamma}^M \) [resp. \( \hat{\phi}_M \)]. Once more by Theorem 2.1 and mutual independence of \( \{ H_i^n: i = 1, \ldots, N_\alpha, \alpha = 1, \ldots, K \} \), we get as \( N \to \infty \), \( \zeta^{MN} \Rightarrow \tilde{\eta}^M \). On the other hand, Lemma 3.6 implies that as \( N \to \infty \), \( \zeta^{MN} \Rightarrow \eta^M \). Combining these observations, we see that \( \tilde{\eta}^M \) and \( \eta^M \) have the same probability distribution, which finishes the proof.

### 3.4.2 Asymptotics of \( J^N \)

Recall the definition of \( J^{N,1}, V^i, h_{\alpha\gamma}, \hat{h}_{\alpha\gamma} \) and \( \hat{h} \) in (3.23), (3.8), (3.11) and (3.12) respectively. All convergence statements in this section are under \( \mathbb{P}^N \). It follows by law of large numbers that for \( \alpha \in K \), as \( N \to \infty \),

\[
\frac{1}{N_\alpha} \sum_{i \in N_\alpha} h_{\alpha\alpha}(V^i, V^i) \Rightarrow \int_{\Omega_d} h_{\alpha\alpha}(\omega, \omega) \nu_\alpha(d\omega) = 0.
\]
By Lemma 3.8 and the above result, we have the following convergence as $N \to \infty$:

$$
J^{N,1}(T) = \sum_{\alpha=1}^{K} \sum_{i=1}^{N_\alpha} \int_{0}^{T} \sum_{\gamma=1}^{K} \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} b_{x^{i}_{\gamma},t}(X^{i}_{t}, X^{j}_{t}) \cdot dW^{i}_{t}
$$

$$
= \sum_{\alpha=1}^{K} \sum_{\gamma=1}^{K} \sqrt{\lambda_{\alpha}N_{\alpha}} \frac{1}{\sqrt{\lambda_{\alpha}N_{\gamma}}} \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_{\gamma}} \hat{h}_{x^{i}_{\gamma}}(V^{i}, V^{j}) \Rightarrow I_{2}(\hat{h}^{sym}),
$$

where $\hat{h}^{sym}$ is defined as in (3.27) with $\hat{\psi}$ replaced by $\hat{h}$.

Recall the definition of $J^{N,2}$ in (3.24). We split $J^{N,2}$ as follows:

$$
J^{N,2}(T) = \sum_{\alpha=1}^{K} \sum_{i=1}^{N_\alpha} \sum_{\beta=1}^{K} \sum_{j=1}^{N_\beta} \sum_{\gamma=1}^{K} \sum_{k=1}^{N_{\gamma}} \frac{1}{N_{\beta}N_{\gamma}} \int_{0}^{T} b_{x^{i}_{\beta},t}(X^{i}_{t}, X^{j}_{t}) \cdot b_{x^{\gamma}_{\gamma},t}(X^{i}_{t}, X^{k}_{t}) dt
$$

$$
= \sum_{n=1}^{5} \sum_{\alpha=1}^{K} \sum_{i=1}^{N_\alpha} \sum_{\beta=1}^{K} \sum_{j=1}^{N_\beta} \sum_{\gamma=1}^{K} \sum_{k=1}^{N_{\gamma}} \frac{1}{N_{\beta}N_{\gamma}} \int_{0}^{T} b_{x^{i}_{\beta},t}(X^{i}_{t}, X^{j}_{t}) \cdot b_{x^{\gamma}_{\gamma},t}(X^{i}_{t}, X^{k}_{t}) dt \Rightarrow \sum_{n=1}^{5} \mathcal{T}^{N}_{n},
$$

where $\mathcal{S}_{1}$, $\mathcal{S}_{2}$, $\mathcal{S}_{3}$, $\mathcal{S}_{4}$ and $\mathcal{S}_{5}$ are collections of $(\alpha, \beta, \gamma) \in \mathbf{K}^{3}$ and $(i, j, k) \in \mathbf{N}_{\alpha} \times \mathbf{N}_{\beta} \times \mathbf{N}_{\gamma}$ such that \( \{i = j = k\}, \{i = j \neq k\}, \{i = k \neq j\}, \{j = k \neq i\} \) and \( \{i, j, k \text{ distinct}\} \), respectively. For the term $\mathcal{T}^{N}_{1}$, from the boundedness of $b_{x^{i}_{\beta},t}$ it follows that as $N \to \infty$,

$$
\mathcal{T}^{N}_{1} = \sum_{\alpha=1}^{K} \sum_{i=1}^{N_\alpha} \sum_{\beta=1}^{K} \sum_{j=1}^{N_\beta} \frac{1}{N_{\beta}N_{\gamma}} \int_{0}^{T} \|b_{x^{i}_{\beta},t}(X^{i}_{t}, X^{j}_{t})\|^{2} dt \Rightarrow 0.
$$

For $\mathcal{T}^{N}_{2}$, let $\mathcal{S}_{x^{i}_{\beta},t} = \{(i, k) \in \mathbf{N}_{\alpha} \times \mathbf{N}_{\gamma} : i \neq k\}$ for $\alpha, \gamma \in \mathbf{K}$. Given a sequence of random variables $Y_{N}$ on $\Omega^{N}_{d}$, $N \geq 1$, we say $Y_{N}$ converges to 0 in $L^{2}(\Omega^{N}_{d}, \mathbb{P}^{N})$ if $E_{\mathbb{P}_{N}}Y^{2}_{N} \to 0$ as $N \to \infty$. Then

$$
\mathcal{T}^{N}_{2} = \sum_{\alpha=1}^{K} \sum_{\gamma=1}^{K} \frac{1}{N_{\alpha}N_{\gamma}} \int_{0}^{T} b_{x^{i}_{\beta},t}(X^{i}_{t}, X^{j}_{t}) \cdot b_{x^{i}_{\gamma},t}(X^{i}_{t}, X^{k}_{t}) dt \Rightarrow \mathcal{T}^{N}_{2,1} + \mathcal{T}^{N}_{2,2},
$$

where

$$
\mathcal{T}^{N}_{2,1} = \sum_{\alpha=1}^{K} \sum_{\gamma=1}^{K} \frac{1}{N_{\alpha}N_{\gamma}} \sum_{(i, k) \in \mathcal{S}_{x^{i}_{\beta},t}} \int_{0}^{T} \left( b_{x^{i}_{\beta},t}(X^{i}_{t}, X^{j}_{t}) \cdot b_{x^{i}_{\gamma},t}(X^{i}_{t}, X^{k}_{t}) 
\right.
\left. - \int_{\Omega^{d}} b_{x^{i}_{\beta},t}(X^{i}_{t}(\omega), X^{j}_{t}(\omega)) \cdot b_{x^{i}_{\gamma},t}(X^{i}_{t}(\omega), X^{k}_{t}) \nu_{\alpha}(d\omega) \right) dt
$$

$$
\mathcal{T}^{N}_{2,2} = \sum_{\alpha=1}^{K} \sum_{\gamma=1}^{K} \frac{1}{N_{\alpha}N_{\gamma}} \sum_{(i, k) \in \mathcal{S}_{x^{i}_{\beta},t}} \int_{0}^{T} \left( b_{x^{i}_{\beta},t}(X^{i}_{t}, X^{j}_{t}) \cdot b_{x^{i}_{\gamma},t}(X^{i}_{t}, X^{k}_{t}) 
\right.
\left. - \int_{\Omega^{d}} b_{x^{i}_{\beta},t}(X^{i}_{t}(\omega), X^{j}_{t}(\omega)) \cdot b_{x^{i}_{\gamma},t}(X^{i}_{t}(\omega), X^{k}_{t}) \nu_{\alpha}(d\omega) \right) dt
$$

$$
\mathcal{T}^{N}_{2,1} + \mathcal{T}^{N}_{2,2} \Rightarrow 0 \text{ as } N \to \infty.
$$
converges to 0 in $L^2(\Omega^N_d, \mathbb{P}^N)$ as $N \to \infty$, and

$$T_{2,2}^N \doteq \sum_{\alpha, \gamma=1}^K \frac{1}{N_{\alpha} N_{\gamma}} \sum_{(i, k) \in S_{\alpha, \gamma}} \int_0^T \int_{\Omega_d} b_{\alpha, \gamma, t}(X_{s, t}(\omega), X_{s, t}(\omega)) \cdot b_{\alpha, \gamma, t}(X_{s, t}(\omega), X_{t}^k) \nu_{\alpha}(d \omega) \, dt \Rightarrow 0$$

by the law of large numbers. Hence $T_2^N \Rightarrow 0$ as $N \to \infty$. Similarly $T_3^N \Rightarrow 0$ as $N \to \infty$. Consider now $T_4^N$.

$$T_4^N = \sum_{\alpha, \gamma=1}^K \frac{1}{N_{\alpha}^2} \sum_{(i, k) \in S_{\alpha, \gamma}} \int_0^T \sum_k \|b_{\alpha, \gamma, t}(X_{i}^k, X_{t}^k)\|^2 \, dt \doteq T_{4,1}^N + T_{4,2}^N,$$

where

$$T_{4,1}^N = \sum_{\alpha, \gamma=1}^K \frac{1}{N_{\alpha}^2} \sum_{(i, k) \in S_{\alpha, \gamma}} \int_0^T \left( \|b_{\alpha, \gamma, t}(X_{i}^k, X_{t}^k)\|^2 - \int_{\Omega_d} \|b_{\alpha, \gamma, t}(X_{s, t}(\omega), X_{t}^k)\|^2 \nu_{\alpha}(d \omega) \right) \, dt$$

converges to 0 in $L^2(\Omega^N_d, \mathbb{P}^N)$ as $N \to \infty$, and

$$T_{4,2}^N = \sum_{\alpha, \gamma=1}^K \frac{1}{N_{\alpha}^2} \sum_{(i, k) \in S_{\alpha, \gamma}} \int_0^T \left( \int_{\Omega_d} \|b_{\alpha, \gamma, t}(X_{s, t}(\omega), X_{t}^k)\|^2 \nu_{\alpha}(d \omega) \right) \, dt$$

$$+ \int_{\Omega_d} \|b_{\alpha, \gamma, t}(X_{s, t}(\omega), X_{t}^k)\|^2 \nu_{\alpha}(d \omega) \, dt$$

$$\Rightarrow \sum_{\alpha, \gamma=1}^K \frac{1}{N_{\alpha}^2} \sum_{(i, k) \in S_{\alpha, \gamma}} \int_0^T \|b_{\alpha, \gamma, t}(X_{s, t}(\omega), X_{t}^k)\|^2 \nu_{\alpha}(d \omega) \, dt = \text{Trace}(AA^*)$$

as $N \to \infty$, by the law of large numbers and Lemma 3.3. So $T_4^N \Rightarrow \text{Trace}(AA^*)$ as $N \to \infty$. Finally consider $T_5^N$. Let $S_{\alpha, \gamma} = \{(i, j, k) \in N_{\alpha} \times N_{\beta} \times N_{\gamma} : i, j, k \text{ distinct}\}$ for $\alpha, \beta, \gamma \in K$. Recalling the definition of $b_{\alpha, \beta, \gamma, t}$ in (3.19), we have

$$T_5^N = \sum_{\alpha, \beta, \gamma=1}^K \frac{1}{N_{\beta} N_{\gamma}} \sum_{(i, j, k) \in S_{\alpha, \beta, \gamma}} \int_0^T b_{\alpha, \gamma, t}(X_{i}^j, X_{i}^k) \cdot b_{\alpha, \gamma, t}(X_{i}^j, X_{i}^k) \, dt \doteq \sum_{\alpha, \beta, \gamma=1}^K (T_{\alpha, \beta, \gamma, 1}^N + T_{\alpha, \beta, \gamma, 2}^N),$$

where

$$T_{\alpha, \beta, \gamma, 1}^N \doteq \frac{1}{N_{\beta} N_{\gamma}} \sum_{(i, j, k) \in S_{\alpha, \beta, \gamma}} \int_0^T \left( b_{\alpha, \beta, t}(X_{i}^i, X_{i}^j) \cdot b_{\alpha, \gamma, t}(X_{i}^i, X_{i}^k) - b_{\alpha, \beta, t}(X_{i}^j, X_{i}^k) \right) \, dt$$

34
converges to 0 in $L^2(\Omega^N, \mathbb{P}^N)$ as $N \to \infty$, and

$$T_{\alpha\beta\gamma,2}^N = \frac{1}{N_\beta N_\gamma} \sum_{(i,j,k) \in S_{\alpha\beta\gamma}} \int_0^T b_{\alpha\beta\gamma,t}(X^i_t, X^k_t) \, dt.$$  

Recall the definition of $l_{\alpha\beta\gamma}$ and $\hat{l}_{\alpha\beta\gamma}$ in (3.20) and (3.21) respectively. By Lemma 3.8, we have

$$\lim_{N \to \infty} \sum_{\alpha\beta\gamma=1}^K T_{\alpha\beta\gamma,2}^N = \lim_{N \to \infty} \sum_{\alpha\beta\gamma=1}^K \frac{N_\alpha}{N_\beta N_\gamma} \sum_{(j,k) \in S_{\beta\gamma}} \int_0^T b_{\beta\gamma,t}(X^j_t, X^k_t) \, dt$$

$$= \sum_{\alpha\beta\gamma=1}^K I_2(\hat{l}_{\alpha\beta\gamma}^\text{sym}),$$

where $\hat{l}_{\alpha\beta\gamma}^\text{sym}$ is defined as in (3.27) with $\hat{\varphi}$ replaced by $\hat{l}_{\alpha\beta\gamma}$, and the convergence is in distribution.

So $T_{\alpha\beta\gamma,2}^N \Rightarrow \sum_{\alpha\beta\gamma=1}^K I_2(\hat{l}_{\alpha\beta\gamma}^\text{sym})$ as $N \to \infty$. Define $\hat{l}$ from $\Omega^K_d \times \Omega^K_d$ to $\mathbb{R}$ ($\hat{\nu} \otimes \hat{\nu}$ a.s.) as

$$\hat{l}(\omega, \omega') = \sum_{\alpha=1}^K \sum_{\beta=1}^K \sum_{\gamma=1}^K \hat{l}_{\alpha\beta\gamma}(\omega, \omega'), \quad (\omega, \omega') \in \Omega^K_d \times \Omega^K_d.$$

Noting that for all $(\omega, \omega') \in \Omega^K_d \times \Omega^K_d$, $\hat{l}_{\alpha\beta\gamma}(\omega', \omega) = \hat{l}_{\alpha\beta\gamma}(\omega, \omega')$, we have $\sum_{\alpha\beta\gamma=1}^K \hat{l}_{\alpha\beta\gamma}^\text{sym} = \hat{l}$.

Combining the above observations we have as $N \to \infty$,

$$J^{N,2}(T) \Rightarrow \text{Trace}(AA^*) + \sum_{\alpha\beta\gamma=1}^K I_2(\hat{l}_{\alpha\beta\gamma}^\text{sym}) = \text{Trace}(AA^*) + I_2(\hat{l}),$$

In fact by Lemma 3.8 we have as $N \to \infty$, under $\mathbb{P}^N$,

$$(J^{N,1}(T), J^{N,2}(T)) \Rightarrow (I_2(\hat{l}_{\text{sym}}^\text{sym}), \text{Trace}(AA^*) + I_2(\hat{l})).$$  \hfill (3.32)

Recall the function $f \in L^2(\Omega^K_d \times \Omega^K_d, \hat{\nu} \otimes \hat{\nu})$ defined in (3.16). It follows from (3.22) that

$$\hat{l}(\omega, \omega') = \int_{\Omega^K_d} \hat{h}(\omega'', \omega) \hat{h}(\omega'', \omega') \hat{\nu}(d\omega''),$$

35
which implies \( f = 2\hat{h}^{sym} - \hat{l} \). From (3.32) we get as \( N \to \infty \),

\[
J^N(T) \Rightarrow I_2(\hat{h}^{sym}) - \frac{1}{2}(\text{Trace}(AA^*) + I_2(\hat{l})) = \frac{1}{2}(I_2(f) - \text{Trace}(AA^*)) = J,
\]

(3.33)

where \( J \) was introduced in (3.17). In fact from Lemma 3.8 it follows that with \( \hat{\xi}_N^\alpha \) as in (3.25), and \( \phi^{(\alpha)} \) introduced below (3.25), as \( N \to \infty \),

\[
\left( \sum_{\alpha \in K} \hat{\xi}_N^\alpha (\phi^{(\alpha)}), J^N(T) \right) \Rightarrow \left( \sum_{\alpha \in K} I_1(\hat{\phi}^{(\alpha)}), J \right).
\]

(3.34)

We will now use the above convergence result to complete the proof of Theorem 3.4.

3.4.3 Completing the proof of Theorem 3.4

It follows from Lemma 1.2 of [62] (cf. Lemma 2.2) and Lemma 3.3 that \( \mathbb{E}_{\hat{\mathbb{P}}}(\exp(J)) = 1 \), where \( \hat{\mathbb{P}} \) is as introduced above (3.16). Along with (3.33) and the fact that \( \mathbb{E}_{\mathbb{P}^N}(\exp(J^N(T))) = 1 \), we have from Schefle’s theorem that \( \exp(J^N(T)) \) is uniformly integrable. Since

\[
\left| \exp \left( i \sum_{\alpha \in K} \hat{\xi}_N^\alpha (\phi^{(\alpha)}) \right) \right| = 1,
\]

\( \exp(i \sum_{\alpha \in K} \hat{\xi}_N^\alpha (\phi^{(\alpha)}) + J^N(T)) \) is also uniformly integrable. From (3.34) the latter random variable converges in distribution, as \( N \to \infty \), to \( \left( i \sum_{\alpha \in K} I_1(\hat{\phi}^{(\alpha)}) + J \right) \). Hence, using uniform integrability,

\[
\lim_{N \to \infty} \mathbb{E}_{\mathbb{P}^N}(\exp \left( i \sum_{\alpha \in K} \hat{\xi}_N^\alpha (\phi^{(\alpha)}) + J^N(T) \right))
\]

\[
= \mathbb{E}_{\hat{\mathbb{P}}} \left[ \exp \left( i \sum_{\alpha \in K} I_1(\hat{\phi}^{(\alpha)}) + J \right) \right]
\]

\[
= \exp \left( - \frac{1}{2} \left\| (I - A)^{-1} \left( \sum_{\alpha \in K} \hat{\phi}^{(\alpha)} \right) \right\|^2_{L^2(\Omega^K_d, \hat{\nu})} \right),
\]

where the last equality is a consequence of Lemma 1.3 of [62] (cf. Lemma 2.2) and Lemma 3.3. Thus we have proved (3.26), which completes the proof of Theorem 3.4. 

\[ \square \]
3.5 Proof of Theorem 3.5

It was argued in Section 3.4.3 that $E_{\tilde{P}}(\exp(J)) = 1$. Consider now the second statement in the theorem. Recall the definition of $\tilde{\xi}^N(\phi)$ in (3.2). For $\phi \in \mathcal{A}^K$, let

$$\tilde{\xi}^N(\phi) \equiv \frac{1}{\sqrt{N_1 \cdots N_K}} \sum_{i_1 \in N_1} \cdots \sum_{i_K \in N_K} \phi(X_{i_1}, \ldots, X_{i_K}).$$

Then $P \circ (\xi^N(\phi))^{-1} = Q^N \circ (\tilde{\xi}^N(\phi))^{-1}$. Using Lemma 3.8 as for the proof of (3.34), we see that under $P^N$, with $\phi = \phi(X_*(\cdot), \ldots, X_*(\cdot))$, as $N \to \infty$,

$$\exp\left(i\tilde{\xi}^N(\phi) + J^N(T)\right) \Rightarrow \exp\left(iI_K(\hat{\phi}^{sym}) + J\right).$$

As before, $\exp(i\tilde{\xi}^N(\phi) + J^N(T))$ is uniformly integrable. Hence as $N \to \infty$,

$$E_P \exp\left(i\xi^N(\phi)\right) = E_{Q^N} \exp\left(i\tilde{\xi}^N(\phi)\right)$$

$$= E_P \exp\left(i\tilde{\xi}^N(\phi) + J^N(T)\right)$$

$$\to E_P \exp\left(iI_K(\hat{\phi}^{sym}) + J\right)$$

$$= E_{\tilde{Q}} \exp\left(iI_K(\hat{\phi}^{sym})\right),$$

which finishes the proof. \qed
CHAPTER 4
WEAKLY INTERACTING MULTI-TYPE PARTICLE SYSTEMS WITH A COMMON FACTOR

Systems with a common factor arise in many different areas. In Mathematical Finance, they have been used to model correlations between default probabilities of multiple firms [23]. In neuroscience modeling common factor models a systematic noise in the external current input to a neuronal ensemble [30]. For particle approximation schemes for stochastic partial differential equations (SPDE), the common factor corresponds to the underlying driving noise in the SPDE [48, 49]. Central limit theorems for systems of weakly interacting particles with a common factor have previously been studied in [49, 15]. However, both of these papers are limited to exchangeable populations (i.e. $K = 1$).

In this chapter we consider weakly interacting multi-type particle systems with a common factor, which is a variant of the model from Chapter 3. The drift coefficients of the interacting diffusions $\{Z^{i,N}_j\}_{i=1}^N$ will be suitable functions of not only the state of individual particles but also another stochastic process that represents a common source of random input to particle dynamics (see equations (4.1) – (4.3)). Additionally, unlike Section 3.1, we consider a general nonlinear dependence of particle dynamics on empirical measures of particles of different types. However for simplicity we take the time dependent coefficient $f_\alpha$ to be 0.

We will establish an analogous CLT result of Theorem 3.4. Recall that Theorem 3.4 says that the family $\{\xi^N_\alpha(\phi), \phi \in \mathcal{A}_\alpha, \alpha \in K\}$ converges weakly, in the sense of finite dimensional distributions, to a mean zero Gaussian field $\{\xi_\alpha(\phi), \phi \in \mathcal{A}_\alpha, \alpha \in K\}$. Due to the presence of a common factor, the centering term in the definition of $\xi^N_\alpha(\phi)$ (above Theorem 3.4) is in general random (a function of the common factor) and we denote by $\mathcal{V}^N_\alpha(\phi)$ these suitably randomly centered and normalized sums of $\{\phi(Z^{i,N}_j), i \in N_\alpha\}$ (see (4.9)). We prove that under suitable conditions, $\{\mathcal{V}^N_\alpha(\phi), \phi \in \tilde{\mathcal{A}}_\alpha, \alpha \in K\}$, where $\tilde{\mathcal{A}}_\alpha$ is a collection of functions on the path space with suitable integrability, converges in the
sense of finite dimensional distributions to a random field whose distribution is given in terms of a Gaussian mixture.

The chapter is organized as follows. In Section 4.1, we state our main results. Specifically, Section 4.1.1 states a basic condition (Condition 4.1), which will ensure pathwise existence and uniqueness of solutions to both SDE for the N-particle system and a related family of SDE describing the limiting nonlinear Markov process. Main result is Theorem 4.3 which appears in Section 4.1.2. For the sake of the exposition, some of the conditions for this theorem (Conditions 4.4 and 4.5) and related notation appear later in Section 4.2. Finally Section 4.3 contains proofs of Theorem 4.3 and related results.

4.1 Model

For fixed \( N \geq 1 \), consider the system of equations for the \( \mathbb{R}^d \)-valued continuous stochastic processes \( Z_{i,N}^N \), \( i \in \mathbb{N} \), and the \( \mathbb{R}^m \)-valued continuous stochastic processes \( U^N \), given on \( (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}) \). For \( i \in \mathbb{N} \),

\[
Z_{i,N}^t = Z_{i,N}^0 + \int_0^t b_{\alpha}(Z_{i,N}^s, U^N_s, \mu^N_s) \, ds + W^i_t,
\]

\[
U^N_t = U_0 + \int_0^t \bar{b}(U^N_s, \mu^N_s) \, ds + \int_0^t \bar{\sigma}(U^N_s, \mu^N_s) \, d\bar{W}_s,
\]

\[
\mu^N_t = (\mu^1_t, \ldots, \mu^K_t), \quad \mu^N_t = \frac{1}{N} \sum_{j \in \mathbb{N}_\gamma} \delta_{Z_{i,j}^N}.
\]

Here \( (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}) \), \{\( W^i \)\} are as in Section 3.1, \( \bar{W} \) is an \( m \)-dimensional \{\( \mathcal{F}_t \)\}-BM independent of \{\( W^i \)\}. We assume that for \( \alpha \in K \), \{\( Z_{i,N}^0 \)\} \( i \in \mathbb{N}_\alpha \) are i.i.d. with common distribution \( \mu^0_0 \) and are also mutually independent. Moreover, \( U_0 \) is independent of \{\( Z_{i,N}^0 \)\} \( i \in \mathbb{N} \) and has probability distribution \( \tilde{\mu}_0 \). \{\( Z_{i,0}^N \)\} \( i \in \mathbb{N} \) and \( U_0 \) are \( \mathcal{F}_0 \)-measurable.

We note that the model studied in Chapter 3 corresponds to a setting where

\[
b_{\alpha}(z, u, \nu) = \sum_{\gamma=1}^K \langle b_{\alpha \gamma}(z, \cdot), \nu_\gamma \rangle, \quad \nu = (\nu_1, \ldots, \nu_K) \in [\mathcal{P}(\mathbb{R}^d)]^K, \quad (z, u) \in \mathbb{R}^d \times \mathbb{R}^m.
\]

As in Section 3.1, along with above \( N \)-particle equations, we will also consider a related infinite system of equations of McKean-Vlasov type for the \( \mathbb{R}^d \)-valued continuous stochastic processes \( X^i \), \( i \in \mathbb{N} \), and the \( \mathbb{R}^m \)-valued continuous stochastic processes \( Y \), given on \( (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}) \). For \( \alpha \in K \)
and $i \in N_\alpha$,

\begin{align*}
X_i^t &= X_i^0 + \int_0^t b_\alpha(X_i^s, Y_s, \mu_s) \, ds + W_i^t, \quad (4.4) \\
Y_t &= Y_0 + \int_0^t \bar{b}(Y_s, \mu_s) \, ds + \int_0^t \tilde{\sigma}(Y_s, \mu_s) \, d\bar{W}_s, \quad (4.5) \\
\mu_t &= (\mu_t^1, \ldots, \mu_t^K), \quad \mu_t^\gamma = \lim_{N \to \infty} \frac{1}{N_\gamma} \sum_{j \in N_\gamma} \delta_{X_j^t}. \quad (4.6)
\end{align*}

Here $Y_0 = U_0$ and $\{X_0^i\}_{i \in N}$ are independent $\mathcal{F}_0$-measurable random variables with $\mathcal{L}(X_0^i) = \mu_0^\alpha$ for $i \in N_\alpha$ and $\alpha \in K$. Note that $\mu_t^\gamma$ is a random measure for $\gamma \in K$ and the limit in (4.6) is in a.s. sense.

### 4.1.1 Well-posedness

We now give conditions on the coefficient functions under which the systems of equations (4.1) – (4.3) and (4.4) – (4.6) have unique pathwise solutions. A pathwise solution of (4.4) – (4.6) is a collection of continuous processes $(X_i, Y)$, $i \in N$, with values in $\mathbb{R}^d \times \mathbb{R}^m$ such that: (a) $Y$ is $\mathcal{G}_t$-adapted, where $\mathcal{G}_t$ is the $\mathbb{P}$ completion of $\sigma\{Y_0, \bar{W}_s, s \leq t\}$; (b) $X$ is $\mathcal{F}_t$-adapted where $X = (X_i)_{i \in N}$; (c) stochastic integrals on the right sides of (4.4) and (4.5) are well defined; (d) Equations (4.4) – (4.6) hold a.s. Uniqueness of pathwise solutions says that if $(X, Y)$ and $(X', Y')$ are two such solutions with $(X_0, Y_0) = (X'_0, Y'_0)$ then they must be indistinguishable. Existence and uniqueness of solutions to (4.1) – (4.3) are defined in a similar manner. In particular, in this case (a) and (b) are replaced by the requirement that $(Z_{i,N}, U^N)_{i \in N}$ are $\{\mathcal{F}_t\}$-adapted.

Define the metric $d^{(K)}_{BL}$ on $[\mathcal{P}(\mathbb{R}^d)]^K$ as

\[ d_{BL}(\nu, \nu') \triangleq \sum_{\alpha=1}^K d_{BL}(\nu_\alpha, \nu'_\alpha), \quad \nu = (\nu_\alpha)_{\alpha \in K} \in [\mathcal{P}(\mathbb{R}^d)]^K, \nu' = (\nu'_\alpha)_{\alpha \in K} \in [\mathcal{P}(\mathbb{R}^d)]^K. \]

We now introduce conditions on the coefficients that will ensure existence and uniqueness of solutions.

**Condition 4.1.** There exists $L \in (0, \infty)$ such that

(a) For all $x \in \mathbb{R}^d$, $y \in \mathbb{R}^m$, $\nu \in [\mathcal{P}(\mathbb{R}^d)]^K$ and $\alpha \in K$,

\[ \max\{\|b_\alpha(x, y, \nu)\|, \|\bar{b}(y, \nu)\|, \|\tilde{\sigma}(y, \nu)\|\} \leq L. \]
(b) For all $x, x' \in \mathbb{R}^d$, $y, y' \in \mathbb{R}^m$, $\nu, \nu' \in [\mathcal{P}^{(K)}(\mathbb{R}^d)]$, and $\alpha \in K$,

\[
\|b_\alpha(x, y, \nu) - b_\alpha(x', y', \nu')\| \leq L(\|x - x'\| + \|y - y'\| + d_{BL}^{(K)}(\nu, \nu')),
\]

\[
\|\bar{b}(y, \nu) - \bar{b}(y', \nu')\| + \|\bar{\sigma}(y, \nu) - \bar{\sigma}(y', \nu')\| \leq L(\|y - y'\| + d_{BL}^{(K)}(\nu, \nu')).
\]

Under the above condition we can establish the following well-posedness result. The proof follows along the lines of Theorem 2.1 of [15] and is therefore omitted.

**Theorem 4.2.** Suppose that $\int_{\mathbb{R}^m} \|y\|^2 \mu_0(dy) < \infty$, $\int_{\mathbb{R}^d} \|x\|^2 \mu_0^\alpha(dx) < \infty$ for every $\alpha \in K$ and Condition 4.1 holds. Then:

(a) The systems of equations (4.4) – (4.6) has a unique pathwise solution.

(b) The systems of equations (4.1) – (4.3) has a unique pathwise solution.

**Remark 4.1.** (i) We note that the unique pathwise solvability in Theorem 4.2(a) implies from classical results of Yamada–Watanabe [70] (see also Section IV.1 in [41]) that there is a measurable map $U: \mathbb{R}^m \times C_m \to C_m$ such that the process $Y$ that solves (4.4) – (4.6) is given as $Y = U(Y_0, \bar{W})$.

(ii) Recall that $G_s$ is the $P$ completion of $\sigma\{Y_0, \bar{W}_r, r \leq s\}, s \in [0, T]$. Let $G = G_T$. Then similar to Theorem 2.3 of [48] (see also Remark 2.1 in [15]) we can show that if $\{(X^i), Y\}$ is a solution of (4.4) – (4.6) then

\[
\mu_t^\alpha = \mathcal{L}(X^i_t|G) = \mathcal{L}(X^i_t|G_t), \ t \in [0, T], i \in N_\alpha, \alpha \in K.
\]

In particular, there are measurable maps $\Pi^\alpha: \mathbb{R}^m \times C_m \to \mathcal{P}(C_d)$ such that $\Pi^\alpha(Y_0, \bar{W}) = \mu^\alpha$ a.s., where $\mu^\alpha \equiv \mathcal{L}(X^i|G)$ for $\alpha \in K$. Clearly $\mu_t^\alpha$ is identical to the marginal of $\mu^\alpha$ at time instant $t$.

### 4.1.2 Central limit theorem

The conditions and proof for the central limit theorem when there is a common factor require some notations which we prefer to introduce in later sections. In this section we will present the basic limit result while referring the reader to Section 4.2 for precise conditions and definitions.

Recall $\Omega_d = \mathcal{C}_d \times \mathcal{C}_d$ and let $\Omega_m = \mathcal{C}_m \times \mathcal{C}_m$. Define for $N \in \mathbb{N}$ the probability measure $\bar{\mathbb{P}}^N$ on $(\bar{\Omega}^N, \mathcal{B}(\bar{\Omega}^N))$, where $\bar{\Omega}^N \equiv \Omega_m \times \Omega_d^N$, as

\[
\bar{\mathbb{P}}^N \equiv \mathcal{L}((\bar{W}, Y), (W^1, X^1), (W^2, X^2), \ldots, (W^N, X^N)),
\]

where

\[
\bar{\Omega}^N = \Omega_m \times \Omega_d^N,
\]

and

\[
\mathcal{B}(\bar{\Omega}^N) = \mathcal{B}(\Omega_m) \times \mathcal{B}(\Omega_d^N)
\]

with

\[
\mathcal{B}(\Omega_d) = \{E \subseteq \Omega_d : E = \bigcup_{n \in \mathbb{N}} E_n, E_n \text{ open} \},
\]

\[
\mathcal{B}(\Omega_m) = \{E \subseteq \Omega_m : E = \bigcup_{n \in \mathbb{N}} E_n, E_n \text{ open} \},
\]

and

\[
\mathcal{B}(\Omega_d^N) = \{E \subseteq \Omega_d^N : E = \bigcup_{n \in \mathbb{N}} E_n, E_n \text{ open} \}.
\]
where processes on the right side are as introduced below (4.3). Note that \( \bar{P}^N \) can be disintegrated as
\[
\bar{P}^N(d\bar{\omega}d\omega_1 \cdots d\omega_N) = \rho_{p(1)}(\bar{\omega},d\omega_1)\rho_{p(2)}(\bar{\omega},d\omega_2) \cdots \rho_{p(N)}(\bar{\omega},d\omega_N)\bar{P}(d\bar{\omega}),
\]
where \( \bar{P} = \mathcal{L}(\bar{W},Y) \) and for \( \alpha \in K \),
\[
\rho_\alpha(\bar{\omega},d\omega) \doteq \Pi^\alpha((\bar{\omega}_0^2,\bar{\omega}_1^1))(d\omega) \text{ for } \bar{P} \text{ a.e. } \bar{\omega} = (\bar{\omega}_1^1,\bar{\omega}_2^2) \in \Omega_m.
\]

We can now present the main result of this chapter. Recall \( X^* \) introduced in (3.9). For \( \alpha \in K \), denote by \( \bar{A}_\alpha \) the collection of all measurable maps \( \phi: C_d \to \mathbb{R} \) such that \( \phi(X^*) \in L^2(\Omega_d, \rho_\alpha(\bar{\omega}, \cdot)) \) for \( \bar{P} \) a.e. \( \bar{\omega} \in \Omega_m \). For \( \phi \in \bar{A}_\alpha \) and \( \bar{\omega} \in \Omega_m \), let
\[
m_{\phi}^\alpha(\bar{\omega}) \doteq \int_{\Omega_d} \phi(X^*(\omega))\rho_\alpha(\bar{\omega},d\omega), \quad (4.8)
\]
\[
V_N^\alpha(\phi) \doteq \sqrt{N_\alpha} \left( \frac{1}{N_\alpha} \sum_{i \in N_\alpha} \phi(Z_i^N) - m_{\phi}^\alpha(V^0) \right), \quad (4.9)
\]
where \( V^0 \doteq (\bar{W},Y) \). For \( \phi_\alpha \in \bar{A}_\alpha, \alpha \in K \), let \( \pi^N(\phi_1, \ldots, \phi_K) \in \mathcal{P}(\mathbb{R}^K) \) be the probability distribution of \( (V_1^N(\phi_1), \ldots, V_K^N(\phi_K)) \). For \( \bar{\omega} \in \Omega_m \), let \( \pi_{\bar{\omega}}(\phi_1, \ldots, \phi_K) \) be the \( K \)-dimensional multivariate normal distribution with mean zero and covariance matrix \( \Sigma_{\bar{\omega}} \doteq (\Sigma_{\bar{\omega}})_{\alpha,\gamma \in K} \) introduced in (4.16). Let \( \pi(\phi_1, \ldots, \phi_K) \in \mathcal{P}(\mathbb{R}^K) \) be defined as
\[
\pi(\phi_1, \ldots, \phi_K) \doteq \int_{\Omega_m} \pi_{\bar{\omega}}(\phi_1, \ldots, \phi_K)\bar{P}(d\bar{\omega}). \quad (4.10)
\]

The following is the main result of this chapter. The proof is given in Section 4.3.

**Theorem 4.3.** Suppose Conditions 4.1, 4.4 and 4.5 hold. Then for all \( \alpha \in K \) and \( \phi_\alpha \in \bar{A}_\alpha \), \( \pi^N(\phi_1, \ldots, \phi_K) \) converges weakly to \( \pi(\phi_1, \ldots, \phi_K) \) as \( N \to \infty \), where \( \pi(\phi_1, \ldots, \phi_K) \in \mathcal{P}(\mathbb{R}^K) \) is as in (4.10).

**Remark 4.2.** We note that unlike in Theorem 3.4, there is a random centering term \( m_{\phi_\alpha}^\alpha(V^0) \) in the limit theorem (cf. (4.9)). Also, as seen for the definition of \( \pi \) in (4.10), the asymptotic distribution of \( V_N^\alpha(\phi_\alpha) \) is not Gaussian but rather a mixture of Gaussian distributions.
4.2 Conditions and notations for CLT

In this section we will present the main condition that is assumed in Theorem 4.3 and also introduce some functions and operators needed in its proof.

4.2.1 Assumptions for the central limit theorem

Consider the systems of equations given by (4.1) – (4.3) and (4.4) – (4.6). Since, unlike the model considered in Section 3.1, here the dependence of the coefficients on the empirical measure is nonlinear, we will need to impose suitable smoothness conditions. These smoothness conditions can be formulated as follows.

Denote by $\mathcal{J}$ [resp. $\mathcal{J}$] the collection of all real measurable functions $f$ on $\mathbb{R}^{d+m+d}$ [resp. $\mathbb{R}^{m+d}$] that are bounded by 1. Denote by $\bar{\mathcal{J}}$ the class of all $g: \mathbb{R}^m \times [\mathcal{P}(\mathbb{R}^d)]^K \to \mathbb{R}^m$ such that there exist $c_g \in (0, \infty)$; a finite subset $\mathcal{J}_g$ of $\mathcal{J}$; continuous and bounded functions $g(1), g(2)$ from $\mathbb{R}^m \times [\mathcal{P}(\mathbb{R}^d)]^K$ to $\mathbb{R}^{m \times m}$ and from $\mathbb{R}^m \times [\mathcal{P}(\mathbb{R}^d)]^K \times \mathbb{R}^d$ to $\mathbb{R}^m$ respectively; and $\theta_g: \mathbb{R}^m \times [\mathcal{P}(\mathbb{R}^d)]^K \times \mathbb{R}^m \times [\mathcal{P}(\mathbb{R}^d)]^K \to \mathbb{R}^m$ such that for all $r = (y, \nu), r' = (y', \nu') \in \mathbb{R}^m \times [\mathcal{P}(\mathbb{R}^d)]^K$

$$g(r') - g(r) = g(1)(r)(y' - y) + \sum_{\gamma=1}^{K} g(2)_{\gamma}(r, \cdot), (\nu'_{\gamma} - \nu_{\gamma})) + \theta_g(r, r'),$$

and

$$||g(r') - g(r)|| \leq c_g(||y' - y|| + \sum_{\gamma=1}^{K} \max_{f \in \mathcal{J}_g} |(f(y, \cdot), (\nu'_{\gamma} - \nu_{\gamma}))|),$$

where

$$||\theta_g(r, r')|| \leq c_g(||y' - y||^2 + \sum_{\gamma=1}^{K} \max_{f \in \mathcal{J}_g} |(f(y, \cdot), (\nu'_{\gamma} - \nu_{\gamma}))|^2).$$

Write $\tilde{\sigma} = (\tilde{\sigma}_1, \ldots, \tilde{\sigma}_m)$ where $\tilde{\sigma}_k$ is an $\mathbb{R}^m$-valued function for $k = 1, \ldots, m$. The following will be the key assumptions needed in Theorem 4.3.

**Condition 4.4.** $\bar{b}, \tilde{\sigma}_k, k = 1, \ldots, m$ are in class $\bar{\mathcal{J}}$.

We impose analogous smoothness conditions on $b_\alpha$ as follows.

**Condition 4.5.** There exist $c_b \in (0, \infty)$; a finite subset $\mathcal{J}_b$ of $\mathcal{J}$; continuous and bounded functions $b_{\alpha(1)}, b_{\alpha(2)}$ from $\mathbb{R}^{d+m} \times [\mathcal{P}(\mathbb{R}^d)]^K$ to $\mathbb{R}^{d \times m}$ and from $\mathbb{R}^{d+m} \times [\mathcal{P}(\mathbb{R}^d)]^K \times \mathbb{R}^d$ to $\mathbb{R}^d$ respectively; and $\theta_{b_\alpha}: \mathbb{R}^{d+m} \times [\mathcal{P}(\mathbb{R}^d)]^K \times \mathbb{R}^m \times [\mathcal{P}(\mathbb{R}^d)]^K \to \mathbb{R}^d$ such that for all $\alpha \in K$, $x \in \mathbb{R}^d$, $r = (y, \nu)$ and
\[ r' = (y', \nu') \in \mathbb{R}^m \otimes [\mathcal{P}(\mathbb{R}^d)]^K \]

\[ b_\alpha(x, r') - b_\alpha(x, r) = b_{\alpha,(1)}(x, r)(y' - y) + \sum_{\gamma=1}^{K} (b_{\alpha\gamma,(2)}(x, r, \cdot), (\nu'_\gamma - \nu_\gamma)) + \theta_{b_\alpha}(x, r; r') \]

and

\[ \|\theta_{b_\alpha}(x, r; r')\| \leq c_b(\|y' - y\|^2 + \sum_{\gamma=1}^{K} \max_{f \in J_b} |\langle f(x, y, \cdot), (\nu'_\gamma - \nu_\gamma)\rangle|^2). \tag{4.11} \]

The above conditions on \( b_\alpha, \bar{b}, \bar{\sigma} \) are satisfied quite generally. We refer the reader to [15] for details and examples.

### 4.2.2 Canonical processes

Recall the canonical space \( \bar{\Omega}^N = \Omega_m \times \Omega_d^N \) defined in Section 4.1.2. We introduce the following canonical stochastic processes.

For \( \omega = (\bar{\omega}, \omega_1, \omega_2, \ldots, \omega_N) \in \bar{\Omega}^N \), let \( V^i(\omega) = \omega_i, i \in \mathcal{N} \) and \( \bar{V}(\omega) = \bar{\omega} \). As before, abusing notation, we write \( V^i = (W^i, X^i), i \in \mathcal{N}, \bar{V} = (\bar{W}, Y) \). Also with \( \Pi^\alpha \) as in Remark 4.1 let \( \mu^\alpha \doteq \Pi^\alpha(Y_0, \bar{W}) \) for \( \alpha \in \mathcal{K} \) and \( \mu = (\mu^1, \ldots, \mu^K) \). Recall \( \bar{\mathcal{P}}^N \in \mathcal{P}(\bar{\Omega}^N) \) introduced in (4.7). With these definitions, under \( \bar{\mathcal{P}}^N \), (4.4) – (4.6) are satisfied a.s. for \( i \in \mathcal{N} \), where \( \mu^\alpha_i \) is the marginal of \( \mu^\alpha \) at time instant \( t \); also with \( \mathcal{G}_t = \sigma\{Y_0, \bar{W}_s, s \leq t\} \) and \( \mathcal{G} = \mathcal{G}_T, \mu^\alpha_i = \mathcal{L}(X^i|\mathcal{G}) = \mathcal{L}(X^i|\mathcal{G}_t), t \in [0, T], i \in \mathcal{N}_\alpha, \alpha \in \mathcal{K}, \bar{\mathcal{P}}^N \) a.s.; and \( Y \) is \( \{\mathcal{G}_t\} \) adapted. Recall the process \( V_s = (W_s, X_s) \) defined on \( \Omega_d \) in Section 3.2.1. Also define \( \bar{V}_s = (\bar{W}_s, Y_s) \) on \( \Omega_m \) as follows: For \( \bar{\omega} = (\bar{\omega}_1, \bar{\omega}_2) \in \Omega_m, \bar{V}_s(\bar{\omega}) = (\bar{W}_s(\bar{\omega}), Y_s(\bar{\omega})) = (\bar{\omega}_1, \bar{\omega}_2) \). Let \( \mu^\alpha_s : \Omega_m \rightarrow \mathcal{P}(\mathcal{C}_d) \) be defined as \( \mu^\alpha_s(\bar{\omega}) \doteq \Pi^\alpha(Y_{s,0}(\bar{\omega}), \bar{W}_s(\bar{\omega})) \) for \( \bar{\omega} \in \Omega_m \) and \( \alpha \in \mathcal{K} \). Note that \( t \mapsto \mu^\alpha_s(t) \) is a continuous map, i.e. an element of \( \mathcal{C}([0, T] : \mathcal{P}(\mathbb{R}^d)) \), which once more we will denote as \( \mu^\alpha_s \). Finally let \( \mu_s \doteq (\mu^1_s, \ldots, \mu^K_s) \) and \( D_s = (\bar{W}_s, Y_s, \mu_s) \).

### 4.2.3 Some random integral operators

We now introduce some random integral operators, similar to the integral operators introduced in Section 3.2.1, which will be needed to formulate the CLT. Randomness of the integral operators is due to the fact that the kernel function of these operators will depend on the common factor. Recall \( b_{\alpha\gamma,(2)} \) in Condition 4.5. Define function \( b^c_{\alpha\gamma,(2)} \) from \( \mathbb{R}^{d+m} \times [\mathcal{P}(\mathbb{R}^d)]^K \times \mathbb{R}^d \) to \( \mathbb{R}^d \) as follows: For \( (x, r, \bar{x}) \in \mathbb{R}^{d+m} \times [\mathcal{P}(\mathbb{R}^d)]^K \times \mathbb{R}^d \) with \( r = (y, \nu), \)

\[ b^c_{\alpha\gamma,(2)}(x, r, \bar{x}) \doteq b_{\alpha\gamma,(2)}(x, r, \bar{x}) - (b_{\alpha\gamma,(2)}(x, r, \cdot), \nu_\gamma). \]
Recall functions $\bar{b}_{(2),\gamma}$ and $\bar{\sigma}_{k,(2),\gamma}$ introduced in Condition 4.4. Similarly, define functions $\bar{b}_{c,(2),\gamma}$ and $\bar{\sigma}_{c,k,(2),\gamma}$ from $\mathbb{R}^m \times [\mathcal{P}(\mathbb{R}^d)]^K \times \mathbb{R}^d$ to $\mathbb{R}^m$ as follows: For $(\mathbf{r}, \bar{x}) \in \mathbb{R}^m \times [\mathcal{P}(\mathbb{R}^d)]^K \times \mathbb{R}^d$ with $\mathbf{r} = (y, \nu)$,

$$\bar{b}_{c,(2),\gamma}(\mathbf{r}, \bar{x}) \doteq \bar{b}_{(2),\gamma}(\mathbf{r}, \bar{x}) - \langle \bar{b}_{(2),\gamma}(\mathbf{r}, \cdot), \nu_{\gamma} \rangle, \quad (4.12)$$

$$\bar{\sigma}_{c,k,(2),\gamma}(\mathbf{r}, \bar{x}) \doteq \bar{\sigma}_{k,(2),\gamma}(\mathbf{r}, \bar{x}) - \langle \bar{\sigma}_{k,(2),\gamma}(\mathbf{r}, \cdot), \nu_{\gamma} \rangle. \quad (4.13)$$

We now introduce another function given on a suitable path space that will be used to define the kernels in our integral operators. One ingredient in the definition of this function requires additional notational preparation and its precise definition is postponed to Section 4.3. Define for $t \in [0, T]$ and $\alpha, \gamma \in \mathbf{K}$, the function $f_{\alpha,\gamma,t}$ from $\mathbb{R}^d \times \mathcal{C}(\mathbb{C}, [0, t] : \mathbb{R}^{d+2m} \times [\mathcal{P}(\mathbb{R}^d)]^K)$ to $\mathbb{R}^d$ as follows: For $(x^{(1)}, x^{(2)}_{[0,t]}, d_{[0,t]}) \in \mathbb{R}^d \times \mathcal{C}(\mathbb{C}, [0, t] : \mathbb{R}^{d+2m} \times [\mathcal{P}(\mathbb{R}^d)]^K)$ with $d \doteq (w, \mathbf{r}) \doteq (w, y, \nu)$,

$$f_{\alpha,\gamma,t}(x^{(1)}, x^{(2)}_{[0,t]}, d_{[0,t]}) \doteq b_{c,\gamma,t}(x^{(1)}, r, x^{(2)}_t) + b_{\alpha,1}(x^{(1)}, r_t) s_{\gamma,t}(x^{(2)}_{[0,t]}, d_{[0,t]}), \quad (4.14)$$

where the function $s_{\gamma,t}$ from $\mathcal{C}(\mathbb{C}, [0, t] : \mathbb{R}^{d+2m} \times [\mathcal{P}(\mathbb{R}^d)]^K)$ to $\mathbb{R}^m$ will be introduced in Lemma 4.10.

Recall the transition probability kernel $\rho_\alpha$ introduced below (4.7). Fix $\bar{\omega} \in \Omega_m$ and consider the Hilbert space $\mathcal{H}_{\bar{\omega}} = L^2(\Omega_{\alpha_d}^K, \hat{\rho}(\bar{\omega}, \cdot))$, where $\hat{\rho}(\bar{\omega}, d\omega_1, \ldots, d\omega_K) \doteq \rho_1(\bar{\omega}, d\omega_1) \otimes \cdots \otimes \rho_K(\bar{\omega}, d\omega_K)$. Define for $\tilde{P}$ a.e. $\bar{\omega}$, $h_{\alpha,\gamma} \in L^2(\Omega_{\alpha_d} \times \Omega_{\alpha_{\gamma}}, \rho_\alpha(\bar{\omega}, \cdot) \times \rho_{\gamma}(\bar{\omega}, \cdot))$ as

$$h_{\alpha,\gamma}(\omega, \omega') \doteq \sqrt{\frac{\lambda_\alpha}{\lambda_\gamma}} \int_0^T f_{\alpha,\gamma,t}(X_{s,t}(\omega), X_{s,[0,t]}(\omega'), D_{s,[0,t]}(\bar{\omega})) \cdot dW_{s,t}(\omega), \quad (\omega, \omega') \in \Omega_{\alpha_d}^2.$$  

Let $\hat{h}_{\alpha,\gamma} \in L^2(\Omega_{\alpha_d}^K, \rho(\bar{\omega}, \cdot) \otimes \hat{\rho}(\bar{\omega}, \cdot))$ be the lifted version of $h_{\alpha,\gamma}$, namely $\hat{h}_{\alpha,\gamma}(\omega, \omega') \doteq h_{\alpha,\gamma}(\omega_\alpha, \omega'_\gamma)$ for $\omega = (\omega_1, \ldots, \omega_K) \in \Omega_{\alpha_d}^K$ and $\omega' = (\omega'_1, \ldots, \omega'_K) \in \Omega_{\alpha_{\gamma}}^K$. Define the integral operator $A_{\alpha,\gamma}^{\omega}$ on $\mathcal{H}_{\bar{\omega}}$ as follows. For $g \in \mathcal{H}_{\bar{\omega}}$ and $\omega \in \Omega_{\alpha_d}^K$,

$$A_{\alpha,\gamma}^{\omega} g(\omega) \doteq \int_{\Omega_{\alpha_d}^K} g(\omega') \hat{h}_{\alpha,\gamma}(\omega', \omega) \hat{\rho}(\bar{\omega}, d\omega'). \quad (4.15)$$

Let $A_{\omega} \doteq \sum_{\alpha,\gamma=1}^K A_{\alpha,\gamma}^{\omega}$. Then this is the integral operator on $\mathcal{H}_{\bar{\omega}}$ associated with the kernel $\hat{h}_{\omega} \doteq \sum_{\alpha,\gamma=1}^K \hat{h}_{\alpha,\gamma} \in L^2(\Omega_{\alpha_d}^K \times \Omega_{\alpha_{\gamma}}^K, \rho(\bar{\omega}, \cdot) \otimes \hat{\rho}(\bar{\omega}, \cdot))$. Denote by $I$ the identity operator on $\mathcal{H}_{\bar{\omega}}$. The following lemma is proved similarly as Lemma 3.3. Proof is omitted.
Lemma 4.6. For \( \bar{P} \) a.e. \( \bar{\omega} \), (a) Trace\( (A^n_{\bar{\omega}}) = 0 \) for all \( n \geq 2 \), and (b) \( I - A_{\bar{\omega}} \) is invertible.

Recall the collection \( \tilde{A}_\alpha, \alpha \in K \) introduced in Section 4.1.2. For \( \alpha \in K, \phi_\alpha \in \tilde{A}_\alpha \) and \( \bar{\omega} \in \Omega_m \), let

\[
\Phi^\alpha_{\bar{\omega}}(\omega) = \phi_\alpha(X_s(\omega)) - m^\alpha_{\bar{\omega}}(\bar{\omega}), \quad \omega \in \Omega_d
\]

\[
\Sigma^\alpha_{\bar{\omega}} = ((I - A_{\bar{\omega}})^{-1}\hat{\Phi}^\alpha_{\bar{\omega}}, (I - A_{\bar{\omega}})^{-1}\hat{\Phi}^\alpha_{\bar{\omega}})_{L^2(\Omega^K_d, \hat{\rho}(\bar{\omega}, \cdot))}, \tag{4.16}
\]

where \( m^\alpha_{\bar{\omega}} \) is as in (4.8) and \( \hat{\Phi}^\alpha_{\bar{\omega}} \) is the lifted function defined as in (3.14), namely for \( \omega = (\omega_1, \ldots, \omega_K) \in \Omega^K_d \) and \( \alpha \in K \),

\[
\Phi^\alpha_{\bar{\omega}}(\omega) = \Phi^\alpha_{\bar{\omega}}(\omega_\alpha) = \phi_\alpha(X_s(\omega_\alpha)) - m^\alpha_{\bar{\omega}}(\bar{\omega}). \tag{4.17}
\]

The quantities \( \Sigma^\alpha_{\bar{\omega}}, \alpha, \gamma \in K \) were used in Section 4.1.2 to characterize the limit distribution of \( (\nu^N_1(\phi_1), \ldots, \nu^N_K(\phi_K)) \). In particular, recall that \( \pi_{\bar{\omega}}(\phi_1, \ldots, \phi_K) \) is the \( K \)-dimensional multivariate normal distribution with mean zero and variance-covariance matrix \( \Sigma_{\bar{\omega}} = (\Sigma^\alpha_{\bar{\omega}})_{\alpha, \gamma \in K} \), and \( \pi(\phi_1, \ldots, \phi_K) \) is the Gaussian mixture defined by (4.10). Theorem 4.3, which is proved in Section 4.3 below, says that under Conditions 4.1, 4.4 and 4.5, \( (\nu^N_1(\phi_1), \ldots, \nu^N_K(\phi_K)) \) converges in distribution to \( \pi(\phi_1, \ldots, \phi_K) \), where \( \nu^N_i(\phi_i) \) are as in (4.9).

4.3 Proofs of Theorem 4.3 and related results

In this section we will present the proof of Theorem 4.3. With \( \{V^i = (W^i, X^i)\}_{i \in N}, \bar{V} = (\bar{W}, Y) \) as introduced in Section 4.2.2, define \( Y^N \) as the unique solution of the following equation

\[
Y^N_t = Y_0 + \int_0^t \hat{b}(Y^N_s, \mu^N_s) \, ds + \int_0^t \hat{\sigma}(Y^N_s, \mu^N_s) \, dW_s, \tag{4.18}
\]

where \( \mu^N_t = (\mu^N_1, \ldots, \mu^N_K) \) and \( \mu^N_\gamma = \frac{1}{N^\gamma} \sum_{j \in N^\gamma} \delta_{X^j_t} \) for \( \gamma \in K \). We begin in Section 4.3.1 by introducing the Girsanov’s change of measure that is key to the proof. The main additional work required for the proof of CLT in the presence of a common factor is in the estimation of the difference between \( Y^N \) and \( Y \). This is done in Section 4.3.2. These estimates are used in Sections 4.3.3 and 4.3.4 to study the asymptotics of the Radon-Nikodym derivative. Finally Sections 4.3.5 and 4.3.6 combine these asymptotic results to complete the proof of Theorem 4.3.
4.3.1 Girsanov’s change of measure

Let $R^N = (Y^N, \mu^N)$ and $R = (Y, \mu)$. For $t \in [0, T]$, define \{J^N(t)\} as

$$J^N(t) = J^{N,1}(t) - \frac{1}{2} J^{N,2}(t),$$  \hspace{1cm} (4.19)

where

$$J^{N,1}(t) = \sum_{\alpha=1}^{K} \sum_{i \in N_\alpha} \int_0^t \left[ b_\alpha(X^i_s, R^N_s) - b_\alpha(X^i_s, R_s) \right] \cdot dW^i_s,$$  \hspace{1cm} (4.20)

and

$$J^{N,2}(t) = \sum_{\alpha=1}^{K} \sum_{i \in N_\alpha} \int_0^t \left\| b_\alpha(X^i_s, R^N_s) - b_\alpha(X^i_s, R_s) \right\|^2 ds.$$  \hspace{1cm} (4.21)

Letting for $t \in [0, T]$, $\bar{F}_t^N = \sigma\{\bar{V}(s), V^i(s) : 0 \leq s \leq t, i \in N\}$, we see that \{exp($J^N(t)$)\} is an $\bar{F}_t^N$-martingale under $\bar{P}^N$. Define a new probability measure $\bar{Q}^N$ on $\bar{\Omega}^N$ by

$$\frac{d\bar{Q}^N}{d\bar{P}^N} = \exp(J^N(T)).$$  \hspace{1cm} (4.22)

By Girsanov’s theorem, $(X^1, \ldots, X^N, Y^N, \bar{V})$ has the same probability law under $\bar{Q}^N$ as $(Z^1, \ldots, Z^N, U^N, V^0)$ under $P$, where $V^0$ is introduced below (4.9). For $\phi_\alpha \in \bar{A}_\alpha$, $\alpha \in K$, let

$$\tilde{V}_\alpha^N(\phi_\alpha) = \sqrt{N_\alpha} \left( \frac{1}{N_\alpha} \sum_{i \in N_\alpha} \phi_\alpha(X^i) - m^\bar{\alpha}_\phi(\bar{V}) \right).$$  \hspace{1cm} (4.23)

In order to prove the theorem, from the definition of $\Sigma_\omega$ in Section 4.2.3 and $\pi(\phi_1, \ldots, \phi_K)$ in Section 4.1.2, it suffices to show that

$$\lim_{N \to \infty} E_{\bar{Q}^N} \exp \left( i \sum_{\alpha=1}^{K} \tilde{V}_\alpha^N(\phi_\alpha) \right) = \int_{\Omega_m} \exp \left( - \frac{1}{2} \left\| (I - A_\omega)^{-1} \sum_{\alpha=1}^{K} \hat{\phi}_\omega^\alpha \right\|^2_{L^2(\Omega^K_d, \rho(\omega, \cdot))} \right) \bar{P}(d\bar{\omega}),$$

which from (4.22) is equivalent to showing

$$\lim_{N \to \infty} E_{\bar{P}^N} \exp \left( i \sum_{\alpha=1}^{K} \tilde{V}_\alpha^N(\phi_\alpha) + J^{N,1}(T) - \frac{1}{2} J^{N,2}(T) \right)$$

$$= \int_{\Omega_m} \exp \left( - \frac{1}{2} \left\| (I - A_\omega)^{-1} \sum_{\alpha=1}^{K} \hat{\phi}_\omega^\alpha \right\|^2_{L^2(\Omega^K_d, \rho(\omega, \cdot))} \right) \bar{P}(d\bar{\omega}).$$  \hspace{1cm} (4.24)
The above equality will be established in Section 4.3.6.

4.3.2 Studying $Y^N - Y$

The following lemma is an immediate consequence of the fact that for each $\gamma \in K$, conditionally on $G$, $X^j$ are i.i.d. for $j \in N_\gamma$. Proof is omitted.

**Lemma 4.7.** For each $\gamma \in K$ and $r \in \mathbb{N}$, there exists $\tilde{a}_r \in (0, \infty)$ such that for all $N \in \mathbb{N}$

$$\sup_{\|f\|_\infty \leq 1} \mathbb{E}_{\bar{P}^N} \left| \left( f(\cdot), (\bar{\mu}^\gamma - \mu^\gamma_t, N) \right) \right|^r \leq \frac{\tilde{a}_r}{N^r \gamma}.$$

As an immediate consequence of the above lemma we have the following lemma.

**Lemma 4.8.** For each $r \in \mathbb{N}$, there exists $a_r \in (0, \infty)$ such that

$$\sup_{t \in [0, T]} \mathbb{E}_{\bar{P}^N} \|Y^N_t - Y_t\|^r \leq \frac{a_r}{N^{r/2}}.$$

**Proof.** Fix $r \in \mathbb{N}$ and $t \in [0, T]$. By (4.5), (4.18), Burkholder–Davis–Gundy inequality and the fact that $\bar{b}$ and $\bar{\sigma}$ satisfy Condition 4.4, there exists $\kappa \in (0, \infty)$ such that

$$\mathbb{E}_{\bar{P}^N} \|Y^N_t - Y_t\|^r = \mathbb{E}_{\bar{P}^N} \left\| \int_0^t \left( \bar{b}(R^N_s) - \bar{b}(R_s) \right) ds + \int_0^t \left( \bar{\sigma}(R^N_s) - \bar{\sigma}(R_s) \right) d\bar{W}_s \right\|^r \leq \kappa \mathbb{E}_{\bar{P}^N} \int_0^t \|Y^N_s - Y_s\|^r ds + \kappa \mathbb{E}_{\bar{P}^N} \sum_{\gamma = 1}^K \max_{f \in \mathcal{J}_\gamma} \int_0^t \left| \left( f(Y_s, \cdot), (\bar{\mu}^\gamma_t, N, \bar{\mu}_s^\gamma) \right) \right|^r ds.$$ 

The result now follows from Lemma 4.7 and an application of Gronwall’s inequality.

The following lemma follows from standard uniqueness results for stochastic differential equations (see eg. Theorem 5.1.1 of [43]) and straightforward applications of Itô’s formula. Recall the canonical space $\bar{\Omega}^N$ from Section 4.1.2 along with the Borel $\sigma$-field $\bar{\mathcal{F}}^N = \mathcal{B}(\bar{\Omega}^N)$ and probability measures $\bar{P}^N$ (see (4.7)). Let $\{\bar{\mathcal{F}}^N_t\}_{t \in [0, T]}$ denote the canonical filtration on $(\bar{\Omega}^N, \bar{\mathcal{F}}^N)$. Note that $\bar{W}$ introduced in Section 4.2.2 is an $m$-dimensional $\{\bar{\mathcal{F}}^N_t\}$-BM under $\bar{P}^N$.

**Lemma 4.9.** Let $\{A_t\}_{t \in [0, T]}$, $\{F^k_t\}_{t \in [0, T]}$, $k = 1, \ldots, m$ be continuous bounded $\{\bar{\mathcal{F}}^N_t\}$-adapted processes with values in $\mathbb{R}^{m \times m}$, given on $(\bar{\Omega}^N, \bar{\mathcal{F}}^N, \bar{P}^N)$. Also let $\{a_t\}_{t \in [0, T]}$, $\{f^k_t\}_{t \in [0, T]}$, $k = 1, \ldots, m$
be progressively measurable processes with values in $\mathbb{R}^m$ such that

$$
\mathbb{E}_{\mathbb{P}^N} \int_0^T \|a_s\|^2 \, ds + \sum_{k=1}^m \mathbb{E}_{\mathbb{P}^N} \int_0^T \|f_s^k\|^2 \, ds < \infty.
$$

Write $\bar{W} = (\bar{W}^1, \ldots, \bar{W}^m)$. Then

(a) The following $m \times m$ dimensional equation has a unique pathwise solution:

$$
\Phi_t = I_m + \int_0^t A_s \Phi_s \, ds + \sum_{k=1}^m \int_0^t F_s^k \Phi_s \, d\bar{W}^k_s, \quad (4.25)
$$

$$
\Psi_t = I_m - \int_0^t \Psi_s A_s \, ds - \sum_{k=1}^m \int_0^t \Psi_s F_s^k \, d\bar{W}^k_s + \sum_{k=1}^m \int_0^t \Psi_s (F_s^k)^2 \, ds, \quad (4.26)
$$

where $I_m$ is the $m \times m$ identity matrix. Furthermore, $\Phi_t, \Psi_t$ are $m \times m$ invertible matrices a.s. and $\Psi_t = \Phi_t^{-1}$.

(b) Given a square integrable $\mathcal{F}^N_0$-measurable random variable $\hat{Y}_0$, the following $m$-dimensional equation has a unique pathwise solution:

$$
\hat{Y}_t = \hat{Y}_0 + \int_0^t (A_s \hat{Y}_s + a_s) \, ds + \sum_{k=1}^m \int_0^t (F_s^k \hat{Y}_s + f_s^k) \, d\bar{W}^k_s. \quad (4.27)
$$

Furthermore the solution is given as

$$
\hat{Y}_t = \Phi_t \left[ \hat{Y}_0 + \int_0^t \Phi_s^{-1} a_s \, ds + \sum_{k=1}^m \int_0^t \Phi_s^{-1} f_s^k \, d\bar{W}^k_s - \sum_{k=1}^m \int_0^t \Phi_s^{-1} F_s^k f_s^k \, ds \right]. \quad (4.28)
$$

The following lemma will give a useful representation for $Y^N_t - Y_t$, and the function $s_{\gamma,t}$ from $\mathbb{C}([0,t]) : \mathbb{R}^{d+2m} \times \mathbb{P}^{(\mathbb{R}^d)^K}$ to $\mathbb{R}^m$ introduced in this lemma is used to define the integral operator $A^{\mathbb{W}}_\omega$ in Section 4.2.3. Recall the functions $g_{(1)}$, $g_{(2)}$ and $\theta_{\omega}$ introduced above Condition 4.4 and centered functions defined in (4.12) and (4.13). Let $D = (\bar{W}, R) = (\bar{W}, Y, \mu)$.

Lemma 4.10. For $t \in [0, T]$,

$$
Y^N_t - Y_t = \sum_{\gamma=1}^K \frac{1}{N^\gamma} \sum_{j \in \mathcal{N}^\gamma} s_{\gamma,t}(X^j_{[0,t]}, D_{[0,t]}) + T^N_1(t),
$$
where

\[ s_{\gamma,t}(X^j_{[0,t]}, D_{[0,t]}) = \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \hat{b}_{(2),\gamma}(R_s, X^j_s) \, ds + \sum_{k=1}^m \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \hat{\sigma}_{k,(2),\gamma}(R_s, X^j_s) \, d\hat{W}^k_s \]

\[ - \sum_{k=1}^m \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \hat{\sigma}_{k,(1)}(R_s) \hat{\sigma}_{k,(2),\gamma}(R_s, X^j_s) \, ds, \]

(4.29)

\[ \mathcal{T}_N^1(t) = \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \theta_b(R_s, R_s^N) \, ds + \sum_{k=1}^m \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \theta_{\sigma_k}(R_s, R_s^N) \, d\hat{W}^k_s \]

\[ - \sum_{k=1}^m \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \hat{\sigma}_{k,(1)}(R_s) \theta_{\sigma_k}(R_s, R_s^N) \, ds, \]

and \( \mathcal{E}_t = \hat{\mathcal{E}}_t(D_{[0,t]}) \) is the unique solution of the \( m \times m \) dimensional SDE

\[ \mathcal{E}_t = I_m + \int_0^t \hat{b}_{(1)}(R_s) \mathcal{E}_s \, ds + \sum_{k=1}^m \int_0^t \hat{\sigma}_{k,(1)}(R_s) \mathcal{E}_s \, d\hat{W}^k_s. \]

Proof. For \( t \in [0,T] \), we have

\[ Y^N_t - Y_t = \int_0^t \left( \hat{b}(R^N_s) - \hat{b}(R_s) \right) \, ds + \sum_{k=1}^m \int_0^t \left( \hat{\sigma}_k(R^N_s) - \hat{\sigma}_k(R_s) \right) \, d\hat{W}^k_s \]

\[ = \int_0^t \left( \hat{b}_{(1)}(R_s)(Y^N_s - Y_s) + \sum_{\gamma=1}^K \left( \hat{b}_{(2),\gamma}(R_s, \cdot), (\mu^{\gamma,N}_s - \mu^\gamma_s) \right) + \theta_b(R_s, R^N_s) \right) \, ds \]

\[ + \sum_{k=1}^m \int_0^t \left( \hat{\sigma}_{k,(1)}(R_s)(Y^N_s - Y_s) + \sum_{\gamma=1}^K \left( \hat{\sigma}_{k,(2),\gamma}(R_s, \cdot), (\mu^{\gamma,N}_s - \mu^\gamma_s) \right) \right) \]

\[ + \theta_{\sigma_k}(R_s, R^N_s) \right) \, d\hat{W}^k_s. \]

The result is now immediate on applying Lemma 4.9 with \( \hat{Y} = Y^N - Y \), \( \Phi = \mathcal{E} \) and

\[ A_s = \hat{b}_{(1)}(R_s), \quad a_s = \sum_{\gamma=1}^K \left( \hat{b}_{(2),\gamma}(R_s, \cdot), (\mu^{\gamma,N}_s - \mu^\gamma_s) \right) + \theta_b(R_s, R^N_s), \]

\[ F^k_s = \hat{\sigma}_{k,(1)}(R_s), \quad f^k_s = \sum_{\gamma=1}^K \left( \hat{\sigma}_{k,(2),\gamma}(R_s, \cdot), (\mu^{\gamma,N}_s - \mu^\gamma_s) \right) + \theta_{\sigma_k}(R_s, R^N_s), \]

for \( k = 1, \ldots, m \). 

\[ \square \]
Lemma 4.11. For every $r \in \mathbb{N}$, we have $\sup_{t \in [0, T]} E_{\mathbb{P}^N} \| \mathcal{E}_t \|^r < \infty$, $\sup_{t \in [0, T]} E_{\mathbb{P}^N} \| \mathcal{E}_t^{-1} \|^r < \infty$ and for every $\gamma \in \mathbb{K}$

$$\sup_{t \in [0, T]} \sup_{j \in \mathbb{N}_\gamma} E_{\mathbb{P}^N} \| s_{\gamma, t}(X_{[0, t], j}^j, D_{[0, t], j}) \|^r < \infty. \quad (4.30)$$

There exist $a_0 \in (0, \infty)$ such that for all $t \in [0, T],

$$E_{\mathbb{P}^N} \| \mathcal{E}_s \|^2 \leq \frac{a_0}{N}, \quad E_{\mathbb{P}^N} \| T_{1, N}^{\gamma}(t) \|^2 \leq \frac{a_0}{N^2}. \quad (4.31)$$

Proof. For fixed $r \in \mathbb{N}$ and $t \in [0, T]$, it follows by the boundedness of $\bar{b}_{(1)}$ and $\bar{\sigma}_{(1)}$ that

$$E_{\mathbb{P}^N} \| \mathcal{E}_t \|^r \leq \kappa_1 \int_0^t E_{\mathbb{P}^N} \| \mathcal{E}_s \|^r \, ds + \kappa_1. \quad (4.32)$$

By Gronwall’s inequality,

$$\sup_{t \in [0, T]} E_{\mathbb{P}^N} \| \mathcal{E}_t \|^r < \infty. \quad (4.33)$$

Using Lemma 4.9, it follows by a similar argument that, for each $r \in \mathbb{N}$

$$\sup_{t \in [0, T]} E_{\mathbb{P}^N} \| \mathcal{E}_t^{-1} \|^r < \infty. \quad (4.34)$$

The estimate in (4.30) now follows from (4.29), (4.32), (4.33), and boundedness of $\bar{b}_{(2), \cdot}$, $\bar{\sigma}_{(2), \cdot}$, $\bar{\sigma}_{(1), \cdot}$. Once again, by boundedness of $\bar{\sigma}_{(2), \cdot}$ and (4.32), (4.33), we get that

$$E_{\mathbb{P}^N} \left\| \sum_{\gamma=1}^{K} \frac{1}{N_{\gamma}} \sum_{j \in \mathbb{N}_{\gamma}} \sum_{k=1}^{m} \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \mathcal{E}_s^{-1} \bar{\sigma}_{k,(2), \gamma}^c(R_s, X_s^j) dW_s^k \right\|^2
\leq \left( E_{\mathbb{P}^N} \| \mathcal{E}_t \|^4 \right)^{1/2} \left( E_{\mathbb{P}^N} \left\| \sum_{k=1}^{m} \int_0^t \mathcal{E}_s^{-1} \sum_{\gamma=1}^{K} \frac{1}{N_{\gamma}} \sum_{j \in \mathbb{N}_{\gamma}} \bar{\sigma}_{k,(2), \gamma}^c(R_s, X_s^j) dW_s^k \right\|^4 \right)^{1/2}
\leq \kappa_2 \left[ \sum_{k=1}^{m} \int_0^t E_{\mathbb{P}^N} \left( \| \mathcal{E}_s^{-1} \|^4 \right) \left( \sum_{\gamma=1}^{K} \frac{1}{N_{\gamma}} \sum_{j \in \mathbb{N}_{\gamma}} \bar{\sigma}_{k,(2), \gamma}^c(R_s, X_s^j) \right)^4 ds \right]^{1/2}
\leq \kappa_2 \left[ \sum_{k=1}^{m} \int_0^t \left( E_{\mathbb{P}^N} \| \mathcal{E}_s^{-1} \|^8 \right)^{1/2} \left( E_{\mathbb{P}^N} \left( \sum_{\gamma=1}^{K} \frac{1}{N_{\gamma}} \sum_{j \in \mathbb{N}_{\gamma}} \bar{\sigma}_{k,(2), \gamma}^c(R_s, X_s^j) \right)^8 \right)^{1/2} ds \right]^{1/2}
\leq \frac{\kappa_3}{N},$$

51
where the last inequality is from Lemma 4.7. Similarly, by boundedness of $\tilde{b}_{(2),\cdot}, \tilde{\sigma}_{\cdot,(1)}, \tilde{\sigma}_{\cdot,(2)}$, and Lemma 4.7, we have that

$$
\mathbb{E}_{\bar{P}_N} \left\| \sum_{\gamma=1}^{K} \frac{1}{N_{\gamma}} \sum_{j \in N_{\gamma}} \left( \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \tilde{b}_{(2),\gamma}^c (R_s, X_s^j) \, ds \right) - \sum_{k=1}^{m} \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \tilde{\sigma}_{k,(1)}^c (R_s) \tilde{\sigma}_{k,(2),\gamma}^c (R_s, X_s^j) \, ds \right\|^2 \leq \frac{\kappa_4}{N}.
$$

Combining the above two observations and recalling the definition of $s_{\gamma,t}$ from (4.29), we have the first estimate in (4.31). A similar argument using Condition 4.4, (4.32), (4.33), Lemmas 4.7 and 4.8 gives the second estimate in (4.31). The result follows.

\[\Box\]

### 4.3.3 Asymptotics of $J^{N,1}(T)$

In this section we analyze the term $J^{N,1}(T)$ defined in (4.20). Recall $S_{\alpha,\gamma} = \{(i,k) \in N_\alpha \times N_\gamma : i \neq k\}$ defined in Section 3.4.2 for $\alpha, \gamma \in K$.

**Lemma 4.12.**

$$J^{N,1}(T) = \sum_{\alpha,\gamma=1}^{K} \frac{1}{N_{\gamma}} \sum_{(i,j) \in S_{\alpha,\gamma}} \int_0^T f_{\alpha,\gamma,t}(X_i^j, X_{[0,t]}, D_{[0,t]}) \cdot dW_t^i + \mathcal{R}_1^N,$$

where $\mathcal{R}_1^N \rightarrow 0$ in probability under $\bar{P}_N$, as $N \rightarrow \infty$.

**Proof.** Note that for each $\alpha \in K$ and $i \in N_\alpha$,

$$b_\alpha(X_i^j, R_t^N) - b_\alpha(X_i^j, R_t) = b_{\alpha,(1)}(X_i^j, R_t)(Y_t^N - Y_t) + \sum_{\gamma=1}^{K} (b_{\alpha,\gamma,(2)}(X_i^j, R_t, \cdot), (\mu_{\gamma,N}^i - \mu_t^i)) + \theta_{ba}(X_i^j, R_t, R_t^N).$$

For the last term in above display, we have from (4.11) and Lemmas 4.7 and 4.8

$$\mathbb{E}_{\bar{P}_N} \left( \sum_{\alpha=1}^{K} \sum_{i \in N_\alpha} \int_0^T \theta_{ba}(X_i^j, R_t, R_t^N) \cdot dW_t^i \right)^2 \leq \frac{\kappa_1}{N} \rightarrow 0, \quad (4.35)$$

52
Now consider the second term on the right side of (4.34):

\[
\sum_{\alpha=1}^{K} \sum_{i \in N_{\alpha}} \int_{0}^{T} \sum_{\gamma=1}^{K} \langle b_{\alpha \gamma}(X_{t}^{i}, R_{t}, \cdot) (\mu_{t}^{\gamma:N} - \mu_{t}^{\gamma}) \rangle \cdot dW_{t}^{i}
\]

\[
= \sum_{\alpha, \gamma=1}^{K} \frac{1}{N_{\gamma}} \sum_{(i,j) \in S_{\alpha \gamma}} \int_{0}^{T} b_{\alpha \gamma}(X_{t}^{i}, R_{t}, X_{t}^{j}) \cdot dW_{t}^{i}
\]

\[
+ \sum_{\alpha=1}^{K} \frac{1}{N_{\alpha}} \sum_{i \in N_{\alpha}} \int_{0}^{T} b_{\alpha \alpha}(X_{t}^{i}, R_{t}, X_{t}^{i}) \cdot dW_{t}^{i}.
\]

Using the boundedness of \( b_{\cdot, (2)} \) it follows that the second moment of the second term on the right side above converges to 0. Finally consider the first term on the right side of (4.34). It follows from Lemma 4.10 that

\[
\sum_{\alpha=1}^{K} \sum_{i \in N_{\alpha}} \int_{0}^{T} b_{\alpha}(X_{t}^{i}, R_{t}) (Y_{t}^{N} - Y_{t}) \cdot dW_{t}^{i}
\]

\[
= \sum_{\alpha, \gamma=1}^{K} \frac{1}{N_{\gamma}} \sum_{(i,j) \in S_{\alpha \gamma}} \int_{0}^{T} b_{\alpha}(X_{t}^{i}, R_{t}, X_{t}^{j}) s_{\gamma,t}(X_{[0,t]}^{j}, D_{[0,t]}^{j}) \cdot dW_{t}^{i}
\]

\[
+ \sum_{\alpha=1}^{K} \frac{1}{N_{\alpha}} \sum_{i \in N_{\alpha}} \int_{0}^{T} b_{\alpha}(X_{t}^{i}, R_{t}) s_{\alpha,t}(X_{[0,t]}^{i}, D_{[0,t]}^{i}) \cdot dW_{t}^{i}
\]

\[
+ \sum_{\alpha=1}^{K} \sum_{i \in N_{\alpha}} \int_{0}^{T} b_{\alpha}(X_{t}^{i}, R_{t}) T_{N}^{N}(t) \cdot dW_{t}^{i}.
\]

By boundedness of \( b_{\cdot, (1)} \) and Lemma 4.11, we have the second moments of the last two terms on the right side above are bounded by \( \frac{\kappa_{2}}{N} \). Result now follows by combining above observations and recalling \( f_{\alpha \gamma,t} \) defined in (4.14). \(\square\)

### 4.3.4 Asymptotics of \( J_{N,2}(T) \)

In this section we analyze the term \( J_{N,2}(T) \) defined in (4.21). We will need some notations.

Let \( x \in C_{d} \) and \( z = (x^{(1)}, x^{(2)}, d) \in C([0, T] : \mathbb{R}^{2d+2m} \times [\mathcal{P}(\mathbb{R}^{d})]^{K}) \) with \( d = (w, r) = (w, y, v) \).

Define for \( \alpha, \beta, \gamma \in K \), functions \( s_{\alpha \beta \gamma,i} \), \( i = 1, 2, 3 \), and \( s_{\alpha \beta \gamma} \) from \( C([0, T] : \mathbb{R}^{3d+2m} \times [\mathcal{P}(\mathbb{R}^{d})]^{K}) \) to
$\mathbb{R}$ as follows:

$$s_{\alpha\beta\gamma,1}(x, z) = \int_0^T b_{\alpha(1)}(x_t, r_t) s_{\beta, t}(x^{(1)}_t, d^{(0)}_t) \cdot b_{\alpha(1)}(x_t, r_t) s_{\gamma, t}(x^{(2)}_t, d^{(0)}_t) \, dt,$$

$$s_{\alpha\beta\gamma,2}(x, z) = \int_0^T \left( b_{\alpha(1)}(x_t, r_t) s_{\beta, t} (x^{(1)}_t, d^{(0)}_t) \cdot b_{\alpha\gamma(2)}^{(1)}(x_t, r_t, x^{(2)}_t) 
+ b_{\alpha(1)}(x_t, r_t) s_{\gamma, t} (x^{(2)}_t, d^{(0)}_t) \cdot b_{\alpha\beta(2)}^{(1)}(x_t, r_t, x^{(1)}_t) \right) \, dt,$$

$$s_{\alpha\beta\gamma,3}(x, z) = \int_0^T b_{\alpha\beta(2)}^{(1)}(x_t, r_t, x^{(1)}_t) \cdot b_{\alpha\gamma(2)}^{(1)}(x_t, r_t, x^{(2)}_t) \, dt,$$

$$s_{\alpha\beta\gamma}(x, z) = \sum_{i=1}^3 s_{\alpha\beta\gamma,i}(x, z). \quad (4.36)$$

Note that

$$s_{\alpha\beta\gamma}(x, z) = \int_0^T f_{\alpha\beta, t}(x_t, x^{(1)}_t, d^{(0)}_t) \cdot f_{\alpha\gamma, t}(x_t, x^{(2)}_t, d^{(0)}_t) \, dt, \quad (4.37)$$

where $f_{\alpha\beta, t}$ is as in (4.14). Define $m_{\alpha\beta\gamma} : \mathbb{C}([0, T] : \mathbb{R}^{2d+2m} \times \mathcal{P}(\mathbb{R}^d)^K) \to \mathbb{R}$ as

$$m_{\alpha\beta\gamma}(z) = \int_0^T \int_{\mathbb{R}^d} f_{\alpha\beta, t}(x', x^{(1)}_t, d^{(0)}_t) \cdot f_{\alpha\gamma, t}(x', x^{(2)}_t, d^{(0)}_t) \nu_{\alpha, t}(dx') \, dt, \quad (4.38)$$

and let $s_{\alpha\beta\gamma}^c : \mathbb{C}([0, T] : \mathbb{R}^{3d+2m} \times \mathcal{P}(\mathbb{R}^d)^K) \to \mathbb{R}$ be given as

$$s_{\alpha\beta\gamma}^c(x, z) = s_{\alpha\beta\gamma}(x, z) - m_{\alpha\beta\gamma}(z).$$

Recall $D$ introduced above Lemma 4.10. The following lemma gives a useful representation for $J^{N,2}(T)$.

**Lemma 4.13.**

$$J^{N,2}(T) = \sum_{\alpha, \beta, \gamma = 1}^K \frac{N_\alpha}{N_{\beta} N_{\gamma}} \sum_{(j,k) \in S_{\beta\gamma}} m_{\alpha\beta\gamma}(X^j, X^k, D)$$

$$+ \sum_{\alpha, \gamma = 1}^K \frac{N_\alpha}{N_{\gamma}^2} \sum_{j \in N_{\gamma}} m_{\alpha\gamma\gamma}(X^j, X^j, D) + \mathcal{R}_2^N,$$

where $\mathcal{R}_2^N \to 0$ in probability under $\bar{P}^N$, as $N \to \infty$.  

54
Proof. Note that for each $\alpha \in K$ and $i \in N_\alpha$,

$$
\|b_\alpha(X^i_t, R^N_t) - b_\alpha(X^i_t, R_t)\|^2 \\
= \|b_{\alpha,(1)}(X^i_t, R_t)(Y^N_t - Y_t) + \sum_{\gamma=1}^{K} \langle b_{\alpha,(2)}(X^i_t, R_t, \cdot), (\mu^{\gamma,N}_t - \mu^\gamma) \rangle + \theta_{b_\alpha}(X^i_t, R_t, R^N_t)\|^2 \\
= \|b_{\alpha,(1)}(X^i_t, R_t)(Y^N_t - Y_t)\|^2 + \left\| \sum_{\gamma=1}^{K} \langle b_{\alpha,(2)}(X^i_t, R_t, \cdot), (\mu^{\gamma,N}_t - \mu^\gamma) \rangle \right\|^2 + \|\theta_{b_\alpha}(X^i_t, R_t, R^N_t)\|^2 \\
+ 2 \langle b_{\alpha,(1)}(X^i_t, R_t)(Y^N_t - Y_t) \rangle \sum_{\gamma=1}^{K} \langle b_{\alpha,(2)}(X^i_t, R_t, \cdot), (\mu^{\gamma,N}_t - \mu^\gamma) \rangle + T^{N,i}_2(t), \quad (4.39)
$$

where $T^{N,i}_2(t)$ consists of the remaining two crossproduct terms. Using (4.11), Lemma 4.7 and 4.8, as for the proof of (4.35), we see that

$$
E_{\bar{\mathbb{P}}^N} \sum_{\alpha=1}^{K} \sum_{i \in N_\alpha} \int_0^T \|\theta_{b_\alpha}(X^i_t, R_t, R^N_t)\|^2 dt \leq \frac{\kappa_1}{N} \to 0 \text{ as } N \to \infty. \quad (4.40)
$$

Similar estimates together with Cauchy–Schwarz inequality show that

$$
E_{\bar{\mathbb{P}}^N} \sum_{\alpha=1}^{K} \sum_{i \in N_\alpha} \int_0^T |T^{N,i}_2(t)| dt \leq \frac{\kappa_2}{\sqrt{N}} \to 0 \text{ as } N \to \infty. \quad (4.41)
$$

Next we study the first term on the right side of (4.39). Using Lemma 4.10, we have

$$
\|b_{\alpha,(1)}(X^i_t, R_t)(Y^N_t - Y_t)\|^2 \\
= \|b_{\alpha,(1)}(X^i_t, R_t) \left( \sum_{\gamma=1}^{K} \frac{1}{N_\gamma} \sum_{j \in N_\gamma} s_{\gamma,t}(X^j_{[0,t]}, D_{[0,t]}) + T^{N}_1(t) \right) \|^2 \\
= \|b_{\alpha,(1)}(X^i_t, R_t) \sum_{\gamma=1}^{K} \frac{1}{N_\gamma} \sum_{j \in N_\gamma} s_{\gamma,t}(X^j_{[0,t]}, D_{[0,t]}) \|^2 + \|b_{\alpha,(1)}(X^i_t, R_t) T^{N}_1(t)\|^2 + T^{N,i}_3(t),
$$

where $T^{N,i}_3(t)$ is the corresponding crossproduct term. Making use of the boundedness of $b_{\cdot,(1)}$ and Lemma 4.11, we can show that

$$
E_{\bar{\mathbb{P}}^N} \sum_{\alpha=1}^{K} \sum_{i \in N_\alpha} \int_0^T \|b_{\alpha,(1)}(X^i_t, R_t) T^{N}_1(t)\|^2 dt \leq \frac{\kappa_3}{N}, \quad E_{\bar{\mathbb{P}}^N} \sum_{\alpha=1}^{K} \sum_{i \in N_\alpha} \int_0^T |T^{N,i}_3(t)| dt \leq \frac{\kappa_4}{\sqrt{N}}.
$$
Thus recalling the definition of \( s_{\alpha\beta\gamma,1} \), we have

\[
\sum_{\alpha=1}^{K} \sum_{i \in N_{\alpha}} \int_{0}^{T} \| b_{\alpha,(1)}(X_{i}^{t}, R_{t})(Y_{t}^{N} - Y_{t}) \|^2 dt
\]

\[
= \sum_{\alpha=1}^{K} \sum_{i \in N_{\alpha}} \int_{0}^{T} \left\| b_{\alpha,(1)}(X_{i}^{t}, R_{t}) \sum_{\gamma=1}^{K} \frac{1}{N_{\gamma}} \sum_{j \in N_{\gamma}} s_{\gamma,t}(X_{[0,t]}^{j}, D_{[0,t]}) \right\|^2 dt + \mathcal{R}_{1}^{N}
\]

\[
= \sum_{\alpha,\beta,\gamma=1}^{K} \frac{1}{N_{\beta}N_{\gamma}} \sum_{i \in N_{\alpha}, j \in N_{\beta}, k \in N_{\gamma}} s_{\alpha\beta\gamma,1}(X_{i}, X_{j}, X_{k}, D)
\]

where \( E_{\mathbb{P}|N} |\mathcal{R}_{1}^{N}| \to 0 \) as \( N \to \infty \).

We now consider the second term on the right side of (4.39). Recalling the definition of \( s_{\alpha\beta\gamma,3} \), we have

\[
\sum_{\alpha=1}^{K} \sum_{i \in N_{\alpha}} \int_{0}^{T} \left\| \sum_{\gamma=1}^{K} b_{\alpha\gamma,(2)}(X_{i}^{t}, R_{t}, \cdot, \mu_{\gamma,N}^{t} - \mu_{\gamma}^{t}) \right\|^2 dt
\]

\[
= \sum_{\alpha,\beta,\gamma=1}^{K} \frac{1}{N_{\beta}N_{\gamma}} \sum_{i \in N_{\alpha}, j \in N_{\beta}, k \in N_{\gamma}} s_{\alpha\beta\gamma,3}(X_{i}, X_{j}, X_{k}, D).
\]

Finally we consider the crossproduct term on the right side of (4.39). Using Lemma 4.10 we have

\[
\sum_{\alpha=1}^{K} \sum_{i \in N_{\alpha}} 2b_{\alpha,(1)}(X_{i}^{t}, R_{t})(Y_{t}^{N} - Y_{t}) \cdot \sum_{\gamma=1}^{K} b_{\alpha\gamma,(2)}(X_{i}^{t}, R_{t}, \cdot, \mu_{\gamma,N}^{t} - \mu_{\gamma}^{t})
\]

\[
= \sum_{\alpha,\beta,\gamma=1}^{K} \frac{2}{N_{\beta}N_{\gamma}} \sum_{i \in N_{\alpha}, j \in N_{\beta}, k \in N_{\gamma}} b_{\alpha,(1)}(X_{i}^{t}, R_{t}) s_{\beta,t}(X_{j}^{t}, D_{[0,t]}) \cdot b_{\alpha\gamma,(2)}(X_{i}^{t}, R_{t}, X_{k}^{t})
\]

\[
+ \sum_{\alpha,\gamma=1}^{K} \sum_{i \in N_{\alpha}} b_{\alpha,(1)}(X_{i}^{t}, R_{t}) T_{1}^{N}(t) \cdot \frac{2}{N_{\gamma}} \sum_{k \in N_{\gamma}} b_{\alpha\gamma,(2)}(X_{i}^{t}, R_{t}, X_{k}^{t})
\]

\[
= T_{4}^{N}(t) + T_{5}^{N}(t).
\]

Using boundedness of \( b_{\cdot,(1)}, b_{\cdot,(2)} \), Lemma 4.7, 4.11 and Cauchy–Schwarz inequality, we see that

\[
E_{\mathbb{P}|N} \int_{0}^{T} |T_{5}^{N}(t)| dt \leq \frac{\kappa_{5}}{\sqrt{N}} \to 0 \text{ as } N \to \infty.
\]
For the term $T_4^N(t)$, recalling the definition of $s_{\alpha,\beta,\gamma}$ and using elementary symmetry properties, we have

$$
\int_0^T T_4^N(t) \, dt = \sum_{\alpha, \beta, \gamma = 1}^{K} \frac{1}{N_{\beta} N_{\gamma}} \sum_{i \in N_{\alpha}, j \in N_{\beta}, k \in N_{\gamma}} s_{\alpha,\beta,\gamma}(X^i, X^j, X^k, D).
$$

Thus we have

$$
\sum_{\alpha=1}^{K} \sum_{i \in N_{\alpha}} \int_0^T \frac{2b_{\alpha,1}(X^i_t, R_t)(Y^i_t - Y_t) \cdot \sum_{\gamma=1}^{K} b_{\alpha,\gamma,2}(X^i_t, R_t, \cdot, (\mu_{\gamma}^N - \mu_t^*) \cdot dt}
= \sum_{\alpha, \beta, \gamma = 1}^{K} \frac{1}{N_{\beta} N_{\gamma}} \sum_{i \in N_{\alpha}, j \in N_{\beta}, k \in N_{\gamma}} s_{\alpha,\beta,\gamma}(X^i, X^j, X^k, D) + \tilde{R}_2^N, \quad (4.44)
$$

where $\tilde{R}_2^N \to 0$ in probability as $N \to \infty$.

Combining (4.39) – (4.44) and recalling the definition of $s_{\alpha,\beta}$ in (4.36), we have

$$
J^{N,2}(T) = \sum_{\alpha, \beta, \gamma = 1}^{K} \frac{1}{N_{\beta} N_{\gamma}} \sum_{i \in N_{\alpha}, j \in N_{\beta}, k \in N_{\gamma}} s_{\alpha,\beta,\gamma}(X^i, X^j, X^k, D) + \tilde{R}_3^N
$$

\begin{equation}
J^{N,2}(T) = \sum_{\alpha, \beta, \gamma = 1}^{K} \frac{1}{N_{\beta} N_{\gamma}} \sum_{i \in N_{\alpha}, j \in N_{\beta}, k \in N_{\gamma}} s_{\alpha,\beta,\gamma}(X^i, X^j, X^k, D) + \tilde{R}_3^N
+ \sum_{\alpha, \beta, \gamma = 1}^{K} \frac{N_{\alpha}}{N_{\beta} N_{\gamma}} \sum_{(j,k) \in S_{\beta,\gamma}} m_{\alpha,\beta,\gamma}(X^j, X^k, D) + \sum_{\alpha, \gamma = 1}^{K} \frac{N_{\alpha}}{N_{\beta} N_{\gamma}} \sum_{j \in N_{\gamma}} m_{\alpha,\gamma,\gamma}(X^j, X^j, D),
\end{equation}

where $m_{\alpha,\beta}$ is as defined in (4.38) and $\tilde{R}_3^N \to 0$ in probability as $N \to \infty$. From the boundedness of second moment of $s_{\alpha,\beta,\gamma}$ (which follows from Lemma 4.11), conditional independence of $X^i, X^j, X^k$ for distinct indices $i \in N_{\alpha}, j \in N_{\beta}, k \in N_{\gamma}$, and the fact that for all $(x, z) \in \mathbb{C}([0, T] : \mathbb{R}^{3d+2m} \times \mathcal{P}(\mathbb{R}^d)^K)$, $E_{\mathbb{P}_N} s_{\alpha,\beta,\gamma}^c(x^i, x^{(1)}, x^{(2)}, d) = E_{\mathbb{P}_N} s_{\alpha,\beta,\gamma}^c(x, x^{(2)}, d_{[0, t]} = E_{\mathbb{P}_N} s_{\alpha,\beta,\gamma}^c(x, x^{(1)}, X^k, d) = 0,$

it follows that the first term on right side of (4.45) converges to 0 in probability as $N \to \infty$, which completes the proof.

\subsection{4.3.5 Combining contributions from $J^{N,1}(T)$ and $J^{N,2}(T)$}

In this section we will combine Lemmas 4.12 and 4.13 to study the asymptotics of the exponent on the left side of (4.24). Recall $m_{\alpha,\beta}$ defined in (4.38) and canonical maps $X_*, Y_*, \hat{W}_*, \mu_*$ and $D_*$ defined in Section 4.2.2. For fixed $\bar{\omega} \in \Omega_m$, define functions $l_{\omega}^{\alpha,\beta,\gamma} \in L^2(\Omega_d \times \Omega_d, \rho_\alpha(\bar{\omega}, \cdot) \times \rho_\gamma(\bar{\omega}, \cdot))$
as
\[ l^\alpha_{\omega,\beta}(\omega,\omega') \doteq \frac{\lambda_\alpha}{\sqrt{\lambda_\beta \lambda_\gamma}} m_{\alpha\beta\gamma}(X_\star(\omega), X_\star(\omega'), D_\star(\bar{\omega})), \quad (\omega, \omega') \in \Omega_d \times \Omega_d. \] (4.46)

Let \( \hat{\rho}_{\omega,\beta}^{\alpha\beta\gamma} \in L^2(\Omega_d^K \times \Omega_d^K, \bar{\rho}(\omega, \cdot) \times \bar{\rho}(\bar{\omega}, \cdot)) \) be lifted versions of \( l^\alpha_{\omega,\beta} \), namely
\[ \hat{\rho}_{\omega,\beta}^{\alpha\beta\gamma}(\omega,\omega') \doteq l^\alpha_{\omega,\beta}(\omega,\omega'), \quad \omega = (\omega_1, \ldots, \omega_K) \in \Omega_d^K, \omega' = (\omega_1', \ldots, \omega_K') \in \Omega_d^K, \]
and let \( \hat{l}_{\omega} = \sum_{\alpha,\beta,\gamma=1}^{K} \hat{\rho}_{\omega,\beta}^{\alpha\beta\gamma} \). Recall \( \hat{\rho} \) and functions \( \hat{h}_{\omega}^{\alpha\beta\gamma} \) introduced in Section 4.2.3. It follows from (4.37) and (4.38) that for \( \alpha, \alpha', \beta, \gamma \in K \) and \( \omega, \omega' \in \Omega_d^K \),
\[ \int_{\Omega_d^K} \hat{h}_{\omega}^{\alpha\beta}(\omega'', \omega) \hat{h}_{\omega}^{\alpha'\gamma}(\omega'', \omega') \hat{\rho}(\bar{\omega}, d\omega'') = 1_{\{\alpha = \alpha'\}} \hat{\rho}_{\omega}^{\alpha\beta\gamma}(\omega, \omega'). \] (4.47)

Thus, with \( \hat{h}_{\omega} \) as in Section 4.2.3,
\[ \hat{l}_{\omega}(\omega, \omega') = \int_{\Omega_d^K} \hat{h}_{\omega}(\omega'', \omega) \hat{h}_{\omega}(\omega'', \omega') \hat{\rho}(\bar{\omega}, d\omega''). \] (4.48)

Recall the integral operators \( A_{\omega}^{\alpha\beta\gamma} \) defined on \( \mathcal{H}_\omega \) introduced in Section 4.2.3. Then for \( \alpha, \alpha', \beta, \gamma \in K \), the operators \( A_{\omega}^{\alpha\beta}(A_{\omega}^{\alpha'\gamma})^*: \mathcal{H}_\omega \to \mathcal{H}_\omega \) are given as follows: For \( g \in \mathcal{H}_\omega \) and \( \omega \in \Omega_d^K \),
\[ A_{\omega}^{\alpha\beta}(A_{\omega}^{\alpha'\gamma})^* g(\omega) = \int_{\Omega_d^K} \left( \int_{\Omega_d^K} \hat{h}_{\omega}^{\alpha\beta}(\omega'', \omega) \hat{h}_{\omega}^{\alpha'\gamma}(\omega'', \omega') \hat{\rho}(\bar{\omega}, d\omega') \right) g(\omega'') \hat{\rho}(\bar{\omega}, d\omega'') \]
\[ = 1_{\{\alpha = \alpha'\}} \int_{\Omega_d^K} \hat{h}_{\omega}^{\alpha\beta\gamma}(\omega, \omega'') g(\omega'') \hat{\rho}(\bar{\omega}, d\omega''), \]
where the last equality is from (4.47). In particular, we have \( A_{\omega}^{\alpha\beta}(A_{\omega}^{\alpha'\gamma})^* = 0 \) if \( \alpha \neq \alpha' \). Moreover, it follows from the display in (4.47) that for \( \alpha, \beta, \gamma \in K \),
\[ \text{Trace}(A_{\omega}^{\alpha\beta}(A_{\omega}^{\alpha'\gamma})^*) = \int_{\Omega_d^K \times \Omega_d^K} \hat{h}_{\omega}^{\alpha\beta}(\omega, \omega') \hat{h}_{\omega}^{\alpha'\gamma}(\omega, \omega') \hat{\rho}(\bar{\omega}, d\omega') \hat{\rho}(\bar{\omega}, d\omega') \]
\[ = \frac{\lambda_\alpha}{\sqrt{\lambda_\beta \lambda_\gamma}} \int_{\Omega_d^K} m_{\alpha\beta\gamma}(X_\star(\omega_\beta'), X_\star(\omega_\gamma'), D_\star(\bar{\omega})) \hat{\rho}(\bar{\omega}, d\omega') \]
\[ = 1_{\{\beta = \gamma\}} \frac{\lambda_\alpha}{\lambda_\gamma} \int_{\Omega_d} m_{\alpha\gamma\gamma}(X_\star(\omega_\gamma'), X_\star(\omega_\gamma'), D_\star(\bar{\omega})) \hat{\rho}(\bar{\omega}, d\omega'), \] (4.49)
where the last equality holds because of the centered terms in the definition of $s_{\gamma,t}$ in (4.29) and the definition of $s_{\alpha\beta\gamma}$ in (4.36). Thus $\text{Trace}(A_{\alpha\gamma}^\alpha(A_{\alpha\gamma}^\alpha)^*) = 0$ if $\beta \neq \gamma$. Define $\tau: \Omega_m \to \mathbb{R}$ as $\tau(\bar{\omega}) = \text{Trace}(A_{\bar{\omega}}^\alpha A_{\bar{\omega}}^\alpha)$. Thus $\text{Trace}(A_{\bar{\omega}}^\alpha A_{\bar{\omega}}^\alpha) = 0$ if $\beta \neq \gamma$. Define $\tau(\bar{\omega}) = \text{Trace}(A_{\bar{\omega}}^\alpha A_{\bar{\omega}}^\alpha)$, where $A_{\bar{\omega}}$ is the operator introduced below (4.15). The following lemma is immediate from the above calculations.

**Lemma 4.14.**

$$
\sum_{\alpha,\gamma=1}^{K} \frac{N_{\alpha}}{N_{\gamma}^2} \sum_{j \in N_{\gamma}} m_{\alpha\gamma\gamma}(X^j, X^j, D) - \tau(\bar{V}) \to 0
$$

in probability under $\bar{P}^N$ as $N \to \infty$.

**Proof.** Note that for fixed $\bar{\omega} \in \Omega_m$,

$$
\tau(\bar{\omega}) = \text{Trace}(A_{\bar{\omega}}^\alpha A_{\bar{\omega}}^\alpha)
$$

$$
= \text{Trace}\left( \left( \sum_{\alpha=1}^{K} \sum_{\gamma=1}^{K} A_{\alpha\gamma}^\alpha \right) \left( \sum_{\alpha=1}^{K} \sum_{\gamma=1}^{K} A_{\alpha\gamma}^\alpha \right)^* \right) = \sum_{\alpha=1}^{K} \sum_{\gamma=1}^{K} \text{Trace}(A_{\alpha\gamma}^\alpha(A_{\alpha\gamma}^\alpha)^*).
$$

It suffices to show for each pair of $\alpha, \gamma \in K$,

$$
\frac{N_{\alpha}}{N_{\gamma}^2} \sum_{j \in N_{\gamma}} m_{\alpha\gamma\gamma}(X^j, X^j, D) - \text{Trace}(A_{\bar{V}}^{\alpha\gamma}(A_{\bar{V}}^{\alpha\gamma})^*)
$$

converges to 0 in probability as $N \to \infty$. However, this property is immediate from (4.49) and the law of large numbers, since $m_{\alpha\gamma\gamma}(X^j, X^j, D)$ is square integrable and conditional on $\mathcal{G}$, $\{X^j, j \in N_{\gamma}\}$ are i.i.d. with common distribution $\rho_{\gamma}(\bar{V}, \cdot) \circ X^{-1}$.

We will now use the results from Section 2.1 with $S = \Omega_d^K$ and $\nu = \hat{\rho}(\bar{\omega}, \cdot), \bar{\omega} \in \Omega_m$. For each $\bar{\omega} \in \Omega_m, k \geq 1$ and $f \in L_{sym}^2(\hat{\rho}(\bar{\omega}, \cdot) \otimes k)$ the MWI $I_k^\alpha(f)$ is defined as in Section 2.1. More precisely, let $\mathcal{A}^k$ be the collection of all measurable $f: \Omega_m \times (\Omega_d^K)^k \to \mathbb{R}$ such that

$$
\int_{(\Omega_d^K)^k} |f(\bar{\omega}, \omega_1, \ldots, \omega_k)|^2 \rho(\bar{\omega}, 0, \ldots, \rho(\bar{\omega}, d\omega_k) < \infty, \quad \hat{P}\text{ a.e. } \bar{\omega}
$$

and for every permutation $\pi$ on $\{1, \ldots, k\}$,

$$
f(\bar{\omega}, \omega_1, \ldots, \omega_k) = f(\bar{\omega}, \omega_{\pi(1)}, \ldots, \omega_{\pi(k)}) , \quad \hat{P} \otimes \hat{\rho}^\otimes k \text{ a.s.}
$$

59
where
\[ \bar{P} \otimes \hat{\rho}^\otimes k(d\bar{\omega}, d\omega_1, \ldots, d\omega_k) = \bar{P}(d\bar{\omega}) \prod_{i=1}^k \hat{\rho}(\bar{\omega}, d\omega_i). \]

Then there is a measurable space \((\Omega^*, \mathcal{F}^*)\) and a regular conditional probability distribution \(\lambda^*: \Omega_m \times \mathcal{F}^* \rightarrow [0, 1]\) such that on the probability space \((\Omega_m \times \Omega^*, \mathcal{B}(\Omega_m) \otimes \mathcal{F}^*, \bar{P} \otimes \lambda^*)\), where
\[ \bar{P} \otimes \lambda^*(A \times B) = \int_A \lambda^*(\bar{\omega}, B) \bar{P}(d\bar{\omega}), \quad A \times B \in \mathcal{B}(\Omega_m) \otimes \mathcal{F}^*, \]
there is a collection of real valued random variables \(\{I_k(f) : f \in \mathcal{A}^k, k \geq 1\}\) with the properties that

(a) For all \(f \in \mathcal{A}^1\) the conditional distribution of \(I_1(f)\) given \(G^* = \mathcal{B}(\Omega_m) \otimes \{\emptyset, \Omega^*\}\) is Normal with mean zero and variance \(\int_{\Omega^*_k} f^2(\bar{\omega}, \omega) \hat{\rho}(\bar{\omega}, d\omega)\).

(b) \(I_k\) is (a.s.) linear map on \(\mathcal{A}^k\).

(c) For \(f \in \mathcal{A}^k\) of the form
\[ f(\bar{\omega}, \omega_1, \ldots, \omega_k) = \prod_{i=1}^k h(\bar{\omega}, \omega_i) \] such that \(\int_{\Omega^*_k} h^2(\bar{\omega}, \omega) \hat{\rho}(\bar{\omega}, d\omega) < \infty, \bar{P} \text{ a.e. } \bar{\omega},\)
we have for \(\bar{P} \otimes \lambda^*\) a.e. \((\bar{\omega}, \omega^*)\)
\[ I_k(f)(\bar{\omega}, \omega^*) = \sum_{j=0}^{[k/2]} (-1)^j C_{k,j} \left( \int_{\Omega^*_k} h^2(\bar{\omega}, \omega) \hat{\rho}(\bar{\omega}, d\omega) \right)^j (I_1(h)(\bar{\omega}, \omega^*))^{k-2j}, \]
and for \(\bar{P}\) a.e. \(\bar{\omega}\)
\[ \int_{\Omega^*} (I_k(f)(\bar{\omega}, \omega^*))^2 \lambda^*(\bar{\omega}, d\omega^*) = k! \left( \int_{\Omega^*_k} h^2(\bar{\omega}, \omega) \hat{\rho}(\bar{\omega}, d\omega) \right)^k, \]
where \(C_{k,j}\) are as in (2.2). We write \(I_k(f)(\bar{\omega}, \cdot)\) as \(I_k^2(f)\). With an abuse of notation, we will denote once more by \(V_\ast\) the canonical process on \(\Omega_m \times \Omega^*\), i.e. \(V_\ast(\bar{\omega}, \omega^*) = \bar{\omega}\), for \((\bar{\omega}, \omega^*) \in \Omega_m \times \Omega^*\).

From Lemmas 4.12, 4.13 and 4.14 it follows that
\[ J^N(T) = J^{N,1}(T) - \frac{1}{2} J^{N,2}(T) = \tilde{S}^N - \frac{1}{2} \tau(\bar{V}) + \tilde{R}^N, \quad (4.50) \]
where

\[
S^N = \sum_{\alpha, \gamma = 1}^{K} \frac{1}{N_{\gamma}} \sum_{(i,j) \in S_{\alpha \gamma}} \int_0^T f_{\alpha \gamma, i}(X^i_t, X^j_{[0,t]}, D_{[0,t]}) \cdot dW^i_t
\]

\[
- \frac{1}{2} \sum_{\alpha, \beta, \gamma = 1}^{K} \frac{N_{\alpha}}{N_{\beta} N_{\gamma}} \sum_{(j,k) \in S_{\beta \gamma}} m_{\alpha \beta \gamma}(X^j, X^k, D),
\]

and \( \mathcal{R}^N \to 0 \) in probability as \( N \to \infty \) under \( \mathbb{P}^N \).

Define \( F: \Omega_m \times \Omega_d^K \times \Omega_d^K \to \mathbb{R} \) as follows: For \((\bar{\omega}, \omega) \in \Omega_m \times \Omega_d^K \times \Omega_d^K, \)

\[
F(\bar{\omega}, \omega) = \bar{h}_\omega(\omega, \omega) + \hat{h}_\omega(\omega, \omega) - \hat{I}_\omega(\omega, \omega)
\]

where the second equality is from (4.48). Note that \( F \in \mathcal{A}^2 \) and so \( I_2(F) \) is a well defined random variable on \((\Omega_m \times \Omega^*, \mathcal{B}(\Omega_m) \otimes \mathcal{F}^*, \bar{P} \otimes \lambda^*)\). Recall the collection \( \bar{\mathcal{A}}_\alpha, \alpha \in K \), introduced in Section 4.1.2. For \( \phi_\alpha \in \bar{\mathcal{A}}_\alpha, m^\alpha_{\phi_\alpha} \) is defined as in (4.8). For such a \( \phi_\alpha \in \bar{\mathcal{A}}_\alpha, \mathcal{F}^\alpha_\omega(\omega) \) is as defined in (4.17). We denote \( \bar{\Phi}^\alpha: \Omega_m \times \Omega_d^K \to \mathbb{R} \) as \( \bar{\Phi}^\alpha(\bar{\omega}, \omega) \equiv \mathcal{F}^\alpha_\omega(\omega) \), namely for \( \bar{\omega} \in \Omega_m \) and \( \omega = (\omega_1, \ldots, \omega_K) \in \Omega_d^K \),

\[
\bar{\Phi}^\alpha(\bar{\omega}, \omega) = \bar{\Phi}^\alpha_\omega(\omega) \equiv \phi_\alpha(X_\omega(\omega_\alpha)) - m^\alpha_{\phi_\alpha}(\bar{\omega})
\]

Note that \( \Phi^\alpha \in \mathcal{A}^1 \) and so \( I_1(\Phi^\alpha) \) is well defined. Let for \( \phi_\alpha \in \bar{\mathcal{A}}_\alpha, \mathcal{V}^N_\alpha(\phi_\alpha) \) be as in (4.23). From the definition of \( \mathcal{G} \) and \( \mathcal{G}^* \) it follows that there are maps \( L_N \) and \( L \) from \( \Omega_m \) to \( \mathcal{P}(\mathbb{R}^{K+1}) \) such that

\[
\mathcal{L}\left(\left(\mathcal{V}^1_N(\phi_1), \ldots, \mathcal{V}^N_K(\phi_K), \Sigma^N\right) \bigg| \mathcal{G}\right) = L_N(\mathcal{V}), \quad \mathbb{P}^N \text{ a.s.,}
\]

\[
\mathcal{L}\left(\left(I_1(\bar{\Phi}^1), \ldots, I_1(\bar{\Phi}^K), \frac{1}{2}I_2(F)\right) \bigg| \mathcal{G}^*\right) = L(\mathcal{V}_*), \quad \bar{P} \otimes \lambda^* \text{ a.s.}
\]

From conditional independence of \( \{X^i\} \) it follows using Lemma 3.8 that

\[
L_N(\bar{\omega}) \to L(\bar{\omega}) \text{ weakly, for } \bar{P} \text{ a.e. } \bar{\omega}.
\]  

(4.51)

Define \( \bar{\tau}: \Omega_m \times \Omega^* \to \mathbb{R} \) as \( \bar{\tau}(\bar{\omega}, \omega^*) \equiv \text{Trace}(A_\omega A^*_{\omega}) \). The following lemma is the key step.
Lemma 4.15. As $N \to \infty$, $i \sum_{\alpha=1}^{K} \tilde{V}_{\alpha}^{N}(\phi_{\alpha}) + J^{N,1}(T) - \frac{1}{2}J^{N,2}(T)$ converges in distribution to $i \sum_{\alpha=1}^{K} I_{1}(\bar{\phi}_{\alpha}) + \frac{1}{2} I_{2}(F) - \frac{1}{2} \bar{\tau}$.

Proof. Note that from (4.50),

$$i \sum_{\alpha=1}^{K} \tilde{V}_{\alpha}^{N}(\phi_{\alpha}) + J^{N,1}(T) - \frac{1}{2}J^{N,2}(T) = i \sum_{\alpha=1}^{K} \tilde{V}_{\alpha}^{N}(\phi_{\alpha}) + \tilde{S}^{N} - \frac{1}{2} \tau(\bar{V}) + \mathcal{R}^{N},$$

where $\mathcal{R}^{N} \to 0$ in probability as $N \to \infty$. Let $l_{N}$ and $l$ from $\Omega_{m}$ to $\mathcal{P}(\mathbb{C})$, where $\mathbb{C}$ is the space of complex numbers, be such that

$$l_{N}(V) = \mathcal{L}(i \sum_{\alpha=1}^{K} \tilde{V}_{\alpha}^{N}(\phi_{\alpha}) + \tilde{S}^{N} - \frac{1}{2} \tau(\bar{V}) \middle| G), \quad \bar{P}^{N} \text{ a.s.},$$

$$l(V) = \mathcal{L}(i \sum_{\alpha=1}^{K} I_{1}(\bar{\phi}_{\alpha}) + \frac{1}{2} I_{2}(F) - \frac{1}{2} \bar{\tau} \middle| G^{*}), \quad \bar{P} \otimes \lambda^{*} \text{ a.s.}$$

It follows from (4.51) and definition of $\tau$, $\bar{\tau}$ that

$$l_{N}(\bar{\omega}) \to l(\bar{\omega}) \text{ weakly, for } \bar{P} \text{ a.e. } \bar{\omega}.$$ 

The desired convergence is now immediate on combining above observations.

4.3.6 Completing the proof of Theorem 4.3

It follows from Lemma 1.2 of [62] (cf. Lemma 2.2) and Lemma 4.6 that $\bar{P}$ a.s.

$$\mathbb{E}_{\bar{P} \otimes \lambda^{*}} \left[ \exp \left( \frac{1}{2} I_{2}(F) \right) \middle| G^{*} \right] = \exp \left( \frac{1}{2} \text{Trace}(A_{V_{*}}(A_{V_{*}})^{*}) \right).$$

Recalling the definition of $\bar{\tau}$ below (4.51) it follows that

$$\mathbb{E}_{\bar{P} \otimes \lambda^{*}} \left[ \exp \left( \frac{1}{2} I_{2}(F) - \frac{1}{2} \bar{\tau} \right) \right] = 1.$$ 

Also, recall that

$$\mathbb{E}_{\bar{P}^{N}} \left[ \exp \left( J^{N,1}(T) - \frac{1}{2} J^{N,2}(T) \right) \right] = 1.$$
Using Lemma 4.15 along with Scheffé’s theorem we now have as in Section 3.4.3 that

\[
\lim_{N \to \infty} E_{\bar{\Phi}^N} \left[ \exp \left( i \sum_{\alpha=1}^{K} \tilde{V}_{\alpha}^{N}(\phi_{\alpha}) + J_{T}^{N,1}(T) - \frac{1}{2} J_{T}^{N,2}(T) \right) \right]
\]

\[
= E_{\bar{P} \otimes \lambda^*} \left[ \exp \left( i \sum_{\alpha=1}^{K} I_1(\Phi^{\alpha}) + \frac{1}{2} I_2(F) - \frac{1}{2} \bar{\tau} \right) \right]
\]

\[
= E_{\bar{P} \otimes \lambda^*} \left[ E_{P \otimes \lambda^*} \left( \exp \left( i \sum_{\alpha=1}^{K} I_1(\Phi^{\alpha}) + \frac{1}{2} I_2(F) - \frac{1}{2} \bar{\tau} \right) \right) | G_* \right]
\]

\[
= \int_{\Omega_m} \exp \left( - \frac{1}{2} \| (I - A_{\bar{\omega}})^{-1} \sum_{\alpha=1}^{K} \Phi^{\alpha}_{\bar{\omega}} \|_{L^2(\Omega_0^K, \hat{\rho}(\bar{\omega}, \cdot))}^2 \right) \bar{P}(\bar{\omega}),
\]

where the last equality is a consequence of Lemma 1.3 of [62] (cf. Lemma 2.2) and Lemma 4.6.

Thus we have proved (4.24), which completes the proof of Theorem 4.3.
CHAPTER 5

WEAKLY INTERACTING PARTICLE SYSTEMS ON RANDOM GRAPHS

In the example of weakly interacting diffusions considered in (1.1) in Chapter 1, each particle interacts with every other particle and the magnitude of the interaction is of the order $\frac{1}{N}$. As noted in Chapter 1, under suitable conditions, one has propagation of chaos and central limit theorem phenomenon for such a particle system. We can associate the system \{\(Z^{1,N}, \ldots, Z^{N,N}\)\} in (1.1) to a complete graph on \(N\) vertexes which represents the fact that an interaction occurs between every pair of particles. In this chapter, we study the case where the associated interaction graph is not necessarily complete but rather is given in terms of an Erdős–Rényi type random graph and magnitude of the interaction that the neighbors of a vertex have on the evolution of the state at this vertex is inversely proportional to the number of neighbors. We will show that under suitable conditions analogous LLN, POC and CLT results as in Chapter 1 are valid. The precise model we consider is as follows.

Consider for \(N \in \mathbb{N}\), a collection of \(\mathbb{R}^d\)-valued interacting diffusions \(\{Z^{1,N}, \ldots, Z^{N,N}\}\) given on a filtered probability space \((\Omega, \mathcal{F}, P, \{\mathcal{F}_t\})\) as follows: For \(i \in \mathbb{N} \ni \{1, \ldots, N\}\),

\[
Z^{i,N}_t = X^0_i + \int_0^t b(Z^{i,N}_s, \mu^{i,N}_s) \, ds + W^i_t, \quad \mu^{i,N}_t = \frac{1}{N_i(t)} \sum_{j=1}^N \xi^{N}_ij(t) \delta_{Z^j,N_t},
\]

where for \(x \in \mathbb{R}^d\) and \(\theta \in \mathcal{P}(\mathbb{R}^d)\), \(b(x, \theta) \equiv \int_{\mathbb{R}^d} \beta(x, y) \, \theta(dy)\). Here \(\{X^0_i : i \in \mathbb{N}\}\) are i.i.d. \(\mathcal{F}_0\)-measurable \(\mathbb{R}^d\)-valued random variables with distribution \(\mu_0\); \(\{W^i : i \in \mathbb{N}\}\) are i.i.d. \(d\)-dimensional \(\mathcal{F}_t\)-Brownian Motions; \(N_i(t) \equiv \sum_{j=1}^N \xi^{N}_{ij}(t)\), where \(\xi^{N}_{ii}(t) \equiv 1\) for \(i \in \mathbb{N}\) and \(\{\xi^{N}_{ij}(t) = \xi^{N}_{ji}(t) : 1 \leq i < j \leq N\}\) are i.i.d. \(\mathcal{F}_t\)-adapted RCLL processes with \(p_N(t) \equiv \mathbb{P}(\xi^{N}_{ii}(t) = 1) = 1 - \mathbb{P}(\xi^{N}_{ij}(t) = 0)\). Thus \(N_i(t)\) represents the number of neighbors of vertex \(i\) at time \(t\). Note that \(N_i(t) \geq 1\) for all \(t \geq 0\) and \(i \in \mathbb{N}\). We assume that \(\{\xi^{N}_{ij}\}\) are independent of \(\{W_i : i \in \mathbb{N}\}\) and let

\[
\bar{p}_N \equiv \inf_{t \in [0, T]} p_N(t).
\]
Along with the $N$-particle system (5.1) we will also consider a related infinite system of equations for $\mathbb{R}^d$-valued continuous stochastic processes $X^i$ governed by the equation

$$X^i_t = X^i_0 + \int_0^t b(X^i_s, \mu_s) \, ds + W^i_t, \quad \mu_t = \mathcal{L}(X^i_t), \quad i \in \mathbb{N}. \quad (5.2)$$

The existence and uniqueness of pathwise solutions of (5.1) and (5.2) can be shown under the bounded and Lipschitz assumption on the coefficients $\beta$ (Condition 5.1) (cf. [67]). We will show in Theorem 5.2 and its corollaries that under Condition 5.1 and an assumption on sparsity of the interaction graph in terms of $\bar{p}_N$ (Condition 5.3), one has LLN and POC results.

We are also interested in the fluctuations of $\{Z^{i,N}\}$ from $\{X^i\}$ by establishing a suitable CLT. Specifically, let

$$\eta^N(\phi) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \phi(Z^{i,N}), \quad \phi \in L^2_{\mathcal{C}}(\mathcal{C}_d, \mu),$$

where $\mathcal{C}_d = C([0, T] : \mathbb{R}^d)$, $\mu = \mathcal{L}(X^i)$ with $X^i$ as in (5.2) and $L^2_{\mathcal{C}}(\mathcal{C}_d, \mu)$ is the collection of all $\phi \in L^2(\mathcal{C}_d, \mu)$ such that $\langle \phi, \mu \rangle = 0$. We show in Theorem 5.8 that under Condition 5.1 and a stronger assumption on edge probability $p_N$ (Condition 5.7) the family $\{\eta^N(\phi) : \phi \in L^2_{\mathcal{C}}(\mathcal{C}_d, \mu)\}$ converges weakly to a mean zero Gaussian field $\{\eta(\phi) : \phi \in L^2_{\mathcal{C}}(\mathcal{C}_d, \mu)\}$ in the sense of convergence of finite dimensional distributions. The variance-covariance structure of this Gaussian field is the same as that in the setting where the interaction graph is complete (cf. Theorem 1.4), which indicates that as long as the interaction graph is not too sparse (see Condition 5.7), the asymptotic behavior of such general system is similar to the one with a complete interaction graph.

The chapter is organized as follows. In Section 5.1 we begin by introducing our model of weakly interacting diffusions on random graphs. Two basic conditions (Conditions 5.1 and 5.3) on coefficients in the model and sparsity of the interaction graph are stated, under which a law of large numbers and propagation of chaos property are established in Theorem 5.2 and its corollaries. Then we present a central limit theorem (Theorem 5.8) in Section 5.1.3 under a stronger condition (Condition 5.7) on sparsity of the interaction graph. Rest of this chapter gives proofs of the main results, namely Theorems 5.2 and 5.8. Section 5.2 collects some preliminary result. Theorem 5.2 is proved in Section 5.3. In Section 5.4 we prove Theorem 5.8. Finally Section 5.5 collects some auxiliary results.
5.1 Model assumptions and results

Recall the assumptions on the stochastic processes \( \{ \xi_{ij}^N \} \) made below (5.1). The following are two natural families of examples where these assumptions are satisfied.

**Example 5.1.** Let \( \{ \xi_{ij}^N(t) \} \) be time-independent, i.e. \( \xi_{ij}^N(t) \equiv \xi_{ij}^N(0) \). In this case \( \bar{p}_N = p_N(0) \) and \( N_i(t) \equiv N_i(0) \) for all \( t \in [0,T] \) and \( i,j \in \mathbf{N} \).

**Example 5.2.** Let \( \{ \xi_{ij}^N(t) \} \) change values independently according to exponential clocks, i.e. \( \{ \xi_{ij}^N(t) = \xi_{ji}^N(t) : 1 \leq i < j \leq N \} \) is a collection of i.i.d. \( \{0,1\} \)-valued jump processes independent of \( \{ W_i, X_i : i \in \mathbf{N} \} \), with rate matrix

\[
\Gamma_N = \begin{bmatrix} -\lambda_N & \lambda_N \\ \mu_N & -\mu_N \end{bmatrix}
\]

for some positive \( \lambda_N \) and \( \mu_N \). In this setting

\[
p_N(t) = p_N(0)e^{-(\lambda_N+\mu_N)t} + \frac{\lambda_N}{\lambda_N + \mu_N} \left( 1 - e^{-(\lambda_N+\mu_N)t} \right)
\]

and hence

\[
\bar{p}_N = \min\{p_N(0), p_N(T)\} \geq \min\left\{ p_N(0), \frac{\lambda_N}{\lambda_N + \mu_N} \right\}.
\]

Consider the collection of weakly interacting diffusions \( \{ Z_{i,N}, i \in \mathbf{N} \} \) described by (5.1). We make the following assumption on the coefficient \( \beta \).

**Condition 5.1.** There exists some \( K \in (0, \infty) \) such that for all \( x, y, x', y' \in \mathbb{R}^d \),

\[
\| \beta(x,y) \| \leq K, \quad \| \beta(x,y) - \beta(x',y') \| \leq K(\| x - x' \| + \| y - y' \|). \]

It is easy to show that under Condition 5.1 there is a unique pathwise solution to both (5.1) and (5.2). Furthermore, we have the following moment estimate. The proof is given in Section 5.3.

**Theorem 5.2.** Under Condition 5.1,

\[
\sup_{N \geq 1} \sqrt{N p_N} \mathbb{E} \| Z_{i,N} - X_{i} \|_{s,T} < \infty. \quad (5.3)
\]
We will make the following assumption on $p_N$.

**Condition 5.3.** $Np_N \to \infty$ as $N \to \infty$.

Recall that $\mu$ denotes the probability law of $X^i$ on $\mathcal{C}_d$. Theorem 5.2 together with a standard argument implies that, under Conditions 5.1 and 5.3, the following propagation of chaos result holds. We omit the proof.

**Corollary 5.4.** Suppose Conditions 5.1 and 5.3 hold. Then for any $k$-tuple $(i_1, \ldots, i_k)$ with distinct coordinates, as $N \to \infty$,

$$\mathcal{L}(Z^{i_1,N}, \ldots, Z^{i_k,N}) \to \mu^\otimes k.$$ 

Using above results and an argument similar to [67] one can further show the following law of large numbers result. Proof is included in Section 5.5 for completeness.

**Corollary 5.5.** Suppose Conditions 5.1 and 5.3 hold. Then as $N \to \infty$,

(a) $\mu^N \doteq \frac{1}{N} \sum_{i=1}^N \delta_{Z^{i,N}} \Rightarrow \mu$,

(b) Suppose in addition that we are in the setting of Example 5.1, namely $\xi_{ij}^N(t) \equiv \xi_{ij}^N(0)$ for all $t \geq 0$ and $i,j \in N$. Then for each $i \in N$, $\mu^i,N \doteq \frac{1}{N} \sum_{j=1}^N \xi_{ij}^N(0) \delta_{Z^{j,N}} \Rightarrow \mu$.

Next we will show the fluctuations of $\langle f, \mu^N \rangle$ about its law of large numbers limit $\langle f, \mu \rangle$, for $f \in L^2(\mu)$. To simplify the notation, we will abbreviate $\xi_{ij}^N(t)$ and $\xi_{ij}^N$ as $\xi_{ij}(t)$ and $\xi_{ij}$ in the rest of this chapter.

### 5.1.1 Canonical processes

We first introduce the following canonical spaces and stochastic processes. Let $\Omega_d \doteq \mathcal{C}_d \times \mathcal{C}_d$, $\Omega_e \doteq \mathbb{D}([0,T] : \{0,1\})$ and $\Omega_N \doteq \Omega_d^N \times \Omega_e^N$. Denote by $\nu \in \mathcal{P}(\Omega_d)$ the common law of $(W^i, X^i)$ where $i \in N$ and $X^i$ is given by (5.2). Also denote by $\nu_{e,N} \in \mathcal{P}(\Omega_e)^{N \times N}$ the law of the random adjacency matrix process $\{\xi_{ij}(t) : i,j \in N, t \in [0,T]\}$. Define for $N \in \mathbb{N}$ the probability measure $\mathbb{P}^N$ on $\Omega_N$ as

$$\mathbb{P}^N \doteq \mathcal{L}((W^1, X^1), (W^2, X^2), \ldots, (W^N, X^N), \{\xi_{ij} : i,j \in N\}) = \nu^\otimes N \otimes \nu_{e,N}.$$ 

For $\omega = (\omega_1, \omega_2, \ldots, \omega_N, \bar{\omega}) \in \Omega_N$ with $\bar{\omega} = (\bar{\omega}_{ij})_{1 \leq i,j \leq N}$, let $V^i(\omega) \doteq \omega_i, i \in N$ and abusing notation, write

$$V^i \doteq (W^i, X^i), \quad \xi_{ij}(\omega) \doteq \bar{\omega}_{ij}, \quad i,j \in N.$$
Also define the canonical processes $V_\ast \triangleq (W_\ast, X_\ast)$ on $\Omega_d$ as

$$V_\ast(\omega) = (W_\ast(\omega), X_\ast(\omega)) \triangleq (\omega_1, \omega_2), \quad \omega = (\omega_1, \omega_2) \in \Omega_d.$$  

### 5.1.2 Some integral operators

We will need the following functions for stating our central limit theorem. Define for $t \in [0, T]$, function $\beta_t$ from $\mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^d$ as

$$\beta_t(x, y) \triangleq \beta(x, y) - \int_{\mathbb{R}^d} \beta(x, z) \mu_t(dz), \quad (x, y) \in \mathbb{R}^d \times \mathbb{R}^d. \quad (5.4)$$

Define function $h$ from $\Omega_d \times \Omega_d$ to $\mathbb{R}$ (\nu \otimes \nu \text{ a.s.}) as

$$h(\omega, \omega') \triangleq \int_0^T \beta_s(X_s(\omega), X_s(\omega')) \cdot dW_s(\omega), \quad (\omega, \omega') \in \Omega_d \times \Omega_d. \quad (5.5)$$

Now consider the Hilbert space $L^2(\Omega_d, \nu)$. Define integral operator $A$ on $L^2(\Omega_d, \nu)$ as

$$Af(\omega) \triangleq \int_{\Omega_d} h(\omega', \omega) f(\omega') \nu(d\omega'), \quad f \in L^2(\Omega_d, \nu), \quad \omega \in \Omega_d. \quad (5.6)$$

Denote by $I$ the identity operator on $L^2(\Omega_d, \nu)$. For $t \in [0, T]$, let

$$\lambda_t \triangleq \int_{\mathbb{R}^d \times \mathbb{R}^d} \|\beta_t(x, y)\|^2 \mu_t(dx) \mu_t(dy). \quad (5.7)$$

The following lemma is a special case of Lemma 3.3 and proof is omitted.

**Lemma 5.6.** (a) $\text{Trace}(AA^*) = \int_{\Omega_d^2} h^2(\omega, \omega') \nu(d\omega) \nu(d\omega') = \int_0^T \lambda_t \, dt$. (b) $\text{Trace}(A^n) = 0$ for all $n \geq 2$. (c) $I - A$ is invertible.

### 5.1.3 Central limit theorem

For the central limit theorem we need the following strengthened version of Condition 5.3.

**Condition 5.7.** $\|p_N(\cdot) - p(\cdot)\|_{*,T} \to 0$ as $N \to \infty$, and $\bar{p} \triangleq \inf_{t \in [0, T]} p(t) > 0$.

We can now present the following central limit theorem. For $\phi \in L^2_c(\mathcal{C}_d, \mu)$, let $\eta^N(\phi) \triangleq \frac{1}{\sqrt{N}} \sum_{i=1}^N \phi(Z^{i,N})$. 

68
Theorem 5.8. Under Conditions 5.1 and 5.7, \( \{ \eta^N(\phi) : \phi \in L^2_c(C_d, \mu) \} \) converges as \( N \to \infty \) to a mean zero Gaussian field \( \{ \eta(\phi) : \phi \in L^2_c(C_d, \mu) \} \) in the sense of convergence of finite dimensional distributions, where for \( \phi, \psi \in L^2_c(C_d, \mu) \),

\[
E[\eta(\phi)\eta(\psi)] = \langle (I - A)^{-1} \phi, (I - A)^{-1} \psi \rangle_{L^2(\Omega_d, \nu)},
\]

where \( \phi \sim \phi(X) \) and \( \psi \sim \psi(X) \).

Proof of the theorem is given in Section 5.4.

5.2 Preparatory results

In this section we present several elementary results for a binomial distribution, which will be used for the proof of Theorems 5.2 and 5.8. Proofs to these results are provided in Section 5.5 for completeness. Denote by \( Bin(n, p) \) the Binomial distribution with number of trials \( n \) and probability of success \( p \), and \( Bern(p) \) the Bernoulli distribution with probability of success \( p \).

Lemma 5.9. Let \( X \sim Bin(n, p) \) and \( q \equiv 1 - p \), then

\[
E \frac{1}{X + 1} = \frac{1 - q^{n+1}}{(n + 1)p} \leq \frac{1}{(n + 1)p}.
\]

Also for each \( m = 2, 3, \ldots \),

\[
E \frac{1}{X + m} \leq \frac{1 - q^{n+1}}{(n + m)p} \leq \frac{1}{(n + m)p},
\]

\[
E \frac{1}{(X + 1)^m} \leq \frac{m^m}{(n + 1)^m p^m}.
\]

For the following two lemmas, let \( \zeta_{ij} \equiv 1 \) for \( i \in \mathbb{N} \) and \( \{ \zeta_{ij} : 1 \leq i < j \leq N \} \) be i.i.d. \( Bern(p_N) \) random variables. Let \( q_N \equiv 1 - p_N \). For \( i \in \mathbb{N} \), let \( N_i \equiv \sum_{j=1}^{N} \zeta_{ij} \).

Lemma 5.10.

\[
E \left( \sum_{j=1}^{N} \frac{\zeta_{ij}}{N_j} - 1 \right)^2 \leq \frac{3}{Np_N}, \quad i \in \mathbb{N}.
\]

We have the following tail bound on the random variable \( N_1 \).
Lemma 5.11. For $k > 0$, let $\alpha_N(k) = \sqrt{k(N-1) \log N}$, then

$$P(\left| N_1 - N p_N \right| > \alpha_N(k) + 1) \leq \frac{2}{N^{2k}}.$$ 

5.3 Proof of Theorem 5.2

For fixed $t \in [0, T]$, we have

$$E \left\| Z^{i,N} - X_i \right\|_{s,t} = E \left\| \int_0^t (b(Z_s^{i,N}, \mu_s^{i,N}) - b(X_s^i, \mu_s)) \, ds \right\|_{s,t}$$

$$\leq \int_0^t E \left\| \sum_{j=1}^N \frac{\xi_{ij}(s)}{N_i(s)} \beta(Z_s^{i,N}, Z_s^{j,N}) - b(X_s^i, \mu_s) \right\| \, ds. \quad (5.8)$$

Adding and subtracting terms gives us

$$\left\| \sum_{j=1}^N \frac{\xi_{ij}(s)}{N_i(s)} \beta(Z_s^{i,N}, Z_s^{j,N}) - b(X_s^i, \mu_s) \right\|$$

$$\leq \left\| \sum_{j=1}^N \frac{\xi_{ij}(s)}{N_i(s)} (\beta(Z_s^{i,N}, Z_s^{j,N}) - \beta(X_s^i, Z_s^{j,N})) \right\| + \left\| \sum_{j=1}^N \frac{\xi_{ij}(s)}{N_i(s)} (\beta(X_s^i, Z_s^{j,N}) - \beta(X_s^i, X_s^j)) \right\|$$

$$+ \left\| \sum_{j=1}^N \frac{\xi_{ij}(s)}{N_i(s)} (\beta(X_s^i, X_s^j) - b(X_s^i, \mu_s)) \right\|$$

$$= \sum_{n=1}^3 T^{N,n}(s). \quad (5.9)$$

For $T^{N,1}$, we have

$$E T^{N,1}(s) = E \left\| \sum_{j=1}^N \frac{\xi_{ij}(s)}{N_i(s)} (\beta(Z_s^{i,N}, Z_s^{j,N}) - \beta(X_s^i, Z_s^{j,N})) \right\|$$

$$\leq E \sum_{j=1}^N \frac{\xi_{ij}(s)}{N_i(s)} \left\| \beta(Z_s^{i,N}, Z_s^{j,N}) - \beta(X_s^i, Z_s^{j,N}) \right\|$$

$$\leq K E \sum_{j=1}^N \frac{\xi_{ij}(s)}{N_i(s)} \left\| Z_s^{i,N} - X_s^i \right\|$$

$$= K E \left\| Z_s^{i,N} - X_s^i \right\|. \quad (5.10)$$
For $\mathcal{T}^{N,2}$, since $\mathcal{L}(\xi_{ij}(s), N_i(s), Z_{s,a}^{i,N}, X_s^i) = \mathcal{L}(\xi_{ji}(s), N_j(s), Z_{s,a}^{i,N}, X_s^i)$, we have

$$
\mathbb{E}\mathcal{T}^{N,2}(s) = \mathbb{E} \left\| \sum_{j=1}^{N} \frac{\xi_{ij}(s)}{N_i(s)} \left( \beta(X_s^i, Z_{s,a}^{i,N}) - \beta(X_s^i, X_s^i) \right) \right\|
\leq \mathbb{E} \sum_{j=1}^{N} \frac{\xi_{ij}(s)}{N_i(s)} \left\| \beta(X_s^i, Z_{s,a}^{i,N}) - \beta(X_s^i, X_s^i) \right\|
\leq K \sum_{j=1}^{N} \mathbb{E} \left( \frac{\xi_{ij}(s)}{N_i(s)} \left\| Z_{s,a}^{i,N} - X_s^i \right\| \right)
= K \sum_{j=1}^{N} \mathbb{E} \left( \frac{\xi_{ij}(s)}{N_j(s)} \left\| Z_{s,a}^{i,N} - X_s^i \right\| \right)
= K \mathbb{E} \left( \left( \sum_{j=1}^{N} \frac{\xi_{ij}(s)}{N_j(s)} - 1 \right) \left\| Z_{s,a}^{i,N} - X_s^i \right\| \right) + K \mathbb{E} \left\| Z_{s,a}^{i,N} - X_s^i \right\|
\leq K \sqrt{\mathbb{E} \left( \sum_{j=1}^{N} \frac{\xi_{ij}(s)}{N_j(s)} - 1 \right)^2 \left\| Z_{s,a}^{i,N} - X_s^i \right\|^2} + K \mathbb{E} \left\| Z_{s,a}^{i,N} - X_s^i \right\|. \tag{5.11}
$$

Note that Condition 5.1 implies for all $i \in \mathbb{N}$,

$$
\left\| Z_{s,a}^{i,N} - X_s^i \right\| = \left\| \int_0^t \left( b(Z_{s,a}^{i,N}, \mu_{s,a}^{i,N}) - b(X_s^i, \mu_s) \right) \, ds \right\| \leq 2Kt. \tag{5.12}
$$

Applying (5.12) and Lemma 5.10 to (5.11) gives us

$$
\mathbb{E}\mathcal{T}^{N,2}(s) \leq \frac{2\sqrt{3}K^2s}{\sqrt{Np_N(s)}} + K \mathbb{E} \left\| Z_{s,a}^{i,N} - X_s^i \right\|. \tag{5.13}
$$

For $\mathcal{T}^{N,3}$, since $\{X^i : i \in \mathbb{N}\}$ are independent of $\{\xi_{ij}(s), N_i(s) : i, j \in \mathbb{N}\}$, we have

$$
\mathbb{E} \left[ \mathcal{T}^{N,3}(s) \right]^2 = \mathbb{E} \left\| \sum_{j=1}^{N} \frac{\xi_{ij}(s)}{N_i(s)} \left( \beta(X_s^i, X_s^j) - b(X_s^i, \mu_s) \right) \right\|^2
= \mathbb{E} \sum_{j=1}^{N} \sum_{k=1}^{N} \frac{\xi_{ij}(s)}{N_i(s)} \frac{\xi_{ik}(s)}{N_i(s)} \left( \beta(X_s^i, X_s^j) - b(X_s^i, \mu_s) \right) \left( \beta(X_s^i, X_s^k) - b(X_s^i, \mu_s) \right)
= \sum_{j=1}^{N} \mathbb{E} \frac{\xi_{ij}(s)}{N_i(s)} \mathbb{E} \left( \beta(X_s^i, X_s^j) - b(X_s^i, \mu_s) \right)^2.
$$
From Condition 5.1 and Lemma 5.9 it follows that the above display can be bounded by

\[
4K^2 \mathbb{E} \frac{2}{N} \sum_{j=1}^{N} \frac{\xi_{ij}(s)}{N_i(s)}^2 \leq \frac{4K^2}{N \mathbb{P}(s)}.
\]

Hence we have

\[
\mathbb{E} T_{N,3}(s) \leq \frac{2K}{\sqrt{N \mathbb{P}(s)}}.
\] (5.14)

Combining (5.8)–(5.10) and (5.13)–(5.14) gives us

\[
\mathbb{E} \|Z^{i,N} - X^i\|_{s,t} \leq \int_0^t 2K \mathbb{E} \|Z^{i,N} - X^i\|_{s,s} ds + 2\sqrt{3K^2 s + 2K} \sqrt{\frac{3K^2 T^2 + 2K}{N \mathbb{P}(s)}}.
\]

The result now follows from Gronwall’s lemma.

\[\square\]

5.4 Proof of Theorem 5.8

For \(N \in \mathbb{N}\), let \(\Omega, \mathbb{P}, V, V^i, i \in \mathbb{N}, \nu\) be as in Section 5.1.1. For \(t \in [0, T]\), define

\[
J^N(t) = J^{N,1}(t) - \frac{1}{2} J^{N,2}(t), \quad \bar{\mu}^N_i = \frac{1}{N} \sum_{i=1}^{N} \delta_{X^i},
\]

where

\[
J^{N,1}(t) = \sum_{i=1}^{N} \int_0^t \left( b(X^i, \bar{\mu}^N) - b(X^i, \nu) \right) \cdot dW^i_s.
\]

(5.15)

and

\[
J^{N,2}(t) = \sum_{i=1}^{N} \int_0^t \left\| b(X^i, \bar{\mu}^N) - b(X^i, \nu) \right\|^2 ds.
\]

(5.16)

Let \(\mathcal{F}^N_t = \sigma\{V^i(s), 0 \leq s \leq t, i \in \mathbb{N}\}\). Note that \(\{\exp(J^N(t))\}\) is an \(\mathcal{F}^N_t\)-martingale under \(\mathbb{P}^N\).

Define a new probability measure \(\mathbb{Q}^N\) on \(\Omega_N\) by

\[
\frac{d\mathbb{Q}^N}{d\mathbb{P}^N} = \exp(J^N(T)).
\]

By Girsanov’s Theorem, \((X^1, \ldots, X^N, \{\xi_{ij} : i, j \in \mathbb{N}\})\) has the same probability distribution under \(\mathbb{Q}^N\) as \((Z^{1,N}, \ldots, Z^{N,N}, \{\xi_{ij} : i, j \in \mathbb{N}\})\) under \(\mathbb{P}\). For \(\phi \in L^2(\mathcal{C}_d, \mu)\), let \(\bar{\eta}^N(\phi) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \phi(X^i)\). Thus in order to prove the theorem it suffices to show that for any
φ ∈ L^2_c(\mathcal{C}_d, \mu),

\lim_{N \to \infty} \mathbb{E}_{\mathbb{Q}^N} \exp(i\tilde{\eta}^N(\phi)) = \exp \left( -\frac{1}{2} \left\| (I - A)^{-1} \phi \right\|^2_{L^2(\Omega, \nu)} \right),

which is equivalent to showing

\lim_{N \to \infty} \mathbb{E}_{\mathbb{P}^N} \exp \left( i\tilde{\eta}^N(\phi) + J^{N,1}(T) - \frac{1}{2} J^{N,2}(T) \right) = \exp \left( -\frac{1}{2} \left\| (I - A)^{-1} \phi \right\|^2_{L^2(\Omega, \nu)} \right). \quad (5.17)

For this we will need to study the asymptotics of J^{N,1} and J^{N,2} as N → ∞.

### 5.4.1 Asymptotics of J^{N,1}

First we consider the term J^{N,1} in (5.15):

\[ J^{N,1}(T) = \sum_{i=1}^{N} \int_0^T \left( b(X^i_s, \mu^i_s) - b(X^i_s, \mu_s) \right) \cdot dW^i_s \]

\[ = \sum_{i=1}^{N} \int_0^T \left( \frac{1}{N_i(s)} \sum_{j=1}^{N} \xi_{ij}(s) \beta(X^i_s, X^j_s) - b(X^i_s, \mu_s) \right) \cdot dW^i_s \]

\[ = \sum_{i=1}^{N} \sum_{j=1}^{N} \int_0^T \frac{\xi_{ij}(s)}{N_i(s)} \beta_s(X^i_s, X^j_s) \cdot dW^i_s. \]

Let

\[ \tilde{J}^{N,1}(T) = \sum_{i=1}^{N} \sum_{j=1}^{N} \int_0^T \frac{\xi_{ij}(s)}{N_i(s)} \beta_s(X^i_s, X^j_s) \cdot dW^i_s. \quad (5.18) \]

We will argue in Lemma 5.13 that the asymptotic behavior of J^{N,1} is the same as that of \tilde{J}^{N,1}, the proof of which relies on the following lemma.

**Lemma 5.12.** As N → ∞,

\[ \sup_{s \in [0,T]} \mathbb{E}_{\mathbb{P}^N} \left[ \frac{N^2}{N_i^2(s)} - \frac{1}{p_N^2(s)} \right]^2 \to 0, \quad \sup_{s \in [0,T]} \mathbb{E}_{\mathbb{P}^N} \left[ \frac{N}{N_i(s)} - \frac{1}{p_N(s)} \right]^2 \to 0. \]

**Proof.** It suffices to prove the first convergence, since the second one follows from the inequality

\[ \left| \frac{N}{N_i(s)} - \frac{1}{p_N(s)} \right| \leq \left| \frac{N^2}{N_i^2(s)} - \frac{1}{p_N^2(s)} \right| \leq \left| \frac{N^2}{N_i^2(s)} - \frac{1}{p_N^2(s)} \right| . \]
Proof. First note that as Lemma 5.13.

\[ \left( N^2 \left( \frac{N^2}{N^2(s) - \frac{1}{p_N^2(s)}} \right)^2 \right) = \left( \frac{N^2}{N^4(s)p_N^4(s)} \right) 1_{G_N(s)} + \left( \frac{N^2}{N^4(s)p_N^4(s)} \right) 1_{G_N(s)}. \]  

(5.19)

Noting that \( |N^2(s) - N^2p_N^2(s)| \leq N^2 \) and \( N_1(s) \geq 1 \), we have as \( N \to \infty \),

\[ \mathbb{E}_{p_N} \left[ \frac{(N^2(s) - N^2p_N^2(s))^2}{N^4(s)p_N^4(s)} 1_{G_N(s)} \right] \leq \frac{N^4}{p_N^4(s)} \mathbb{P}_{N}(G_N(s)) \leq \frac{2}{N^2p_N^4} \to 0, \]  

(5.20)

where the convergence follows from Condition 5.7. Now consider the second term on the right side of (5.19). Condition 5.7 implies that \( Np_N - \alpha_N(3) - 1 > 0 \) for large enough \( N \). Hence

\[ \mathbb{E}_{p_N} \left[ \frac{(N^2(s) - N^2p_N^2(s))^2}{N^4(s)p_N^4(s)} 1_{G_N(s)} \right] \leq \frac{N^4(N(3) + 1)^2}{Np_N^4(s) - \alpha_N(3)-1} 1_{G_N(s)} \]

\[ \leq \frac{4N^2(\sqrt{3(N-1)\log N} + 1)^2}{(Np_N - \sqrt{3(N-1)\log N} - 1)^4} \to 0 \]  

(5.21)

as \( N \to \infty \). The result follows by combining (5.20) and (5.21). \( \square \)

The following lemma says that to study the asymptotics of \( J^{N,1}(T) \), it suffices to study the asymptotic behavior of \( \tilde{J}^{N,1}(T) \).

Lemma 5.13.

\[ \lim_{N \to \infty} \mathbb{E}_{p_N} \left[ J^{N,1}(T) - \tilde{J}^{N,1}(T) \right]^2 = 0. \]

Proof. First note that as \( N \to \infty \)

\[ \mathbb{E}_{p_N} \left[ \sum_{1 \leq i < j \leq N} \int_0^T \frac{\xi_{ij}(s)}{N_i(s)} \beta_s(X_s^i, X_s^j) \cdot dW_s^i - \sum_{1 \leq i < j \leq N} \int_0^T \frac{\xi_{ij}(s)}{N_p_N(s)} \beta_s(X_s^i, X_s^j) \cdot dW_s^i \right]^2 \]

\[ = \mathbb{E}_{p_N} \left[ \sum_{1 \leq i < j \leq N} \int_0^T \left( \frac{1}{N_i(s)} - \frac{1}{N_p_N(s)} \right) \xi_{ij}(s) \beta_s(X_s^i, X_s^j) \cdot dW_s^i \right]^2 \]

\[ = \sum_{1 \leq i < j \leq N} \int_0^T \left[ \mathbb{E}_{p_N} \left[ \frac{1}{N_i(s)} - \frac{1}{N_p_N(s)} \right] \xi_{ij}(s) \right]^2 \mathbb{E}_{p_N} \beta_s^2(X_s^i, X_s^j) ds \]

\[ \leq \kappa N^2 \int_0^T \mathbb{E}_{p_N} \left[ \frac{1}{N_i(s)} - \frac{1}{N_p_N(s)} \right]^2 ds \to 0, \]
where the convergence follows from Lemma 5.12. Similarly one can show that as \( N \to \infty \),

\[
\mathbb{E}_{P_N} \left| \sum_{1 \leq j < i \leq N} \int_0^T \frac{\xi_{ij}(s)}{N_i(s)} \beta_s(X^i_s, X^j_s) \cdot dW^i_s \right|^2 \to 0,
\]

\[
\mathbb{E}_{P_N} \left| \sum_{i=1}^N \int_0^T \frac{1}{N_i(s)} \beta_s(X^i_s, X^i_s) \cdot dW^i_s \right|^2 \to 0.
\]

Combining above results completes the proof.

Recall \( \lambda_t \) defined in (5.7) and let \( \sigma^2 \).

\[
\sigma^2 = \int_0^T \frac{1 - p(s)}{p(s)} \lambda_s \, ds. \tag{5.22}
\]

Note that Condition 5.7 implies \( \sigma^2 < \infty \). Recall function \( h \) defined in Section 5.1.2 and let \( h^{sym}(\omega, \omega') \equiv \frac{1}{2} (h(\omega, \omega') + h(\omega', \omega)) \) for \((\omega, \omega') \in \Omega_d \times \Omega_d\). The following result in the case where \( \xi_{ij}(s) = \xi_{ij}(0) \) for all \( s \in [0, T] \) and \( i, j \in \mathbb{N} \) was obtained by Janson in the study of incomplete \( U \)-statistics [42]. Proof of the case where \( \{\xi_{ij}\} \) are time dependent is similar, however we provide the argument for completeness. This will be key to studying the asymptotics of \( \tilde{J}^N \).

**Lemma 5.14.**

\[
\sum_{1 \leq j < i \leq N} \int_0^T \frac{\xi_{ij}(s)}{N P_N(s)} \beta_s(X^i_s, X^j_s) \cdot dW^i_s \Rightarrow Z + I_2(h^{sym})
\]

as \( N \to \infty \). Here \( Z \) is a normal random variable with mean zero variance \( \sigma^2 \), and \( I_2(\cdot) \) is as defined in Section 2.1 with \( S = C_d \) and \( \theta = \nu \). Moreover, \( Z \) is independent of \( I_2(f) \) for each \( f \in L^2_{c,sym}(\nu^{\otimes 2}) \) and hence \( Z \) is also independent of \( I_1(g) \) for each \( g \in L^2_\nu \).

**Proof.** Let for \( \zeta \in \mathbb{D}([0, T] : \{0, 1\}) \) and \( \omega, \omega' \in \Omega_d \),

\[
u_N(\zeta, \omega, \omega') = \int_0^T \frac{\zeta(s) - p_N(s)}{p_N(s)} \beta_s(X^*, s(\omega), X^*, s(\omega')) \cdot dW^*_{s}(\omega) \]

\[
+ \int_0^T \frac{\zeta(s) - p_N(s)}{p_N(s)} \beta_s(X^*, s(\omega'), X^*, s(\omega)) \cdot dW^*_{s}(\omega').
\]
Then we have

\[
\sum_{1 \leq i \neq j \leq N} \int_0^T \frac{\xi_{ij}(s)}{N^{p_N(s)}} \beta_s(X^i_s, X^j_s) \cdot dW^i_s
= \sum_{1 \leq i < j \leq N} \left( \int_0^T \frac{\xi_{ij}(s) - p_N(s)}{N^{p_N(s)}} \beta_s(X^i_s, X^j_s) \cdot dW^i_s \right)
\]

\[
\quad + \frac{1}{N} \sum_{1 \leq i < j \leq N} \left( \int_0^T \beta_s(X^i_s, X^j_s) \cdot dW^i_s + \int_0^T \beta_s(X^j_s, X^i_s) \cdot dW^j_s \right)
\]

\[
= \frac{1}{N} \sum_{1 \leq i < j \leq N} u_N(\xi_{ij}, V^i, V^j) + \frac{2}{N} \sum_{1 \leq i < j \leq N} h^{sym}(V^i, V^j)
\]

\[
= U_N + \tilde{U}_N(h^{sym}).
\]

We claim that for each \( f \in L^2_{c,sym}(\nu^{\otimes 2}) \), as \( N \to \infty \)

\[
\left( U_N, \tilde{U}_N(f) \right) \Rightarrow (Z, I_2(f)),
\]

where \( Z \) and \( I_2(f) \) are independent random variables as mentioned in the statement of the lemma.

Once this claim is proved, taking \( f = g \otimes g \) for each \( g \in L^2_c(\nu) \) and noting that \( I_2(f) = [I_1(g)]^2 - \|g\|^2_{L^2(\nu)} \) by (2.1), we have the independence between \( Z \) and \( I_1(g) \).

Now we proceed to verify the above claim. The proof is analogous to the conditioning arguments of Janson [42] (see Lemma 2 and Theorem 1 therein). Denote by \( \mathbb{E}_{\mathbb{P}^N, V} \) the conditional expectation under \( \mathbb{P}^N \) given \( (V^i)_{i=1}^N \). Since \( \{u_N(\xi_{ij}, V^i, V^j) : 1 \leq i < j \leq N\} \) is conditionally independent given \( (V^i)_{i=1}^N \), and \( \mathbb{E}_{\mathbb{P}^N, V}[u_N(\xi_{ij}, V^i, V^j)] = 0 \) for each \( i < j \), we have

\[
\sigma_N^2 = \mathbb{E}_{\mathbb{P}^N, V}[U_N^2] = \frac{1}{N^2} \sum_{1 \leq i < j \leq N} \mathbb{E}_{\mathbb{P}^N, V}[u_N^2(\xi_{ij}, V^i, V^j)].
\]

It follows from Condition 5.7 that as \( N \to \infty \),

\[
\mathbb{E}_{\mathbb{P}^N}[\sigma_N^2] = \frac{1}{N^2} \sum_{1 \leq i < j \leq N} \mathbb{E}_{\mathbb{P}^N}[u_N^2(\xi_{ij}, V^i, V^j)] = \frac{2}{N^2} \sum_{1 \leq i < j \leq N} \int_0^T \frac{1 - p_N(s)}{p_N(s)} \lambda_s ds \to \sigma^2
\]

76
and

\[
E_{\mathbb{P}^N} \left( \sigma_N^2 - E_{\mathbb{P}^N} [\sigma_N^2] \right)^2
\]

\[
= \frac{1}{N^4} \sum_{1 \leq i < j \leq N} E_{\mathbb{P}^N} \left[ \left( E_{\mathbb{P}^N} \left[ u_N^2 (\xi_{ij}, V^i, V^j) \right] - E_{\mathbb{P}^N} [u_N^2 (\xi_{ij}, V^i, V^j)] \right)^2 \right]
\]

\[
\leq \frac{N(N-1)}{2N^4} E_{\mathbb{P}^N} [u_N^4 (\xi_{12}, V^1, V^2)]
\]

\[
\to 0.
\]

So \( \sigma_N^2 \to \sigma^2 \) in probability. Also note that as \( N \to \infty \),

\[
E_{\mathbb{P}^N} \sum_{1 \leq i < j \leq N} E_{\mathbb{P}^N, V} \left| \frac{u_N (\xi_{ij}, V^i, V^j)}{N} \right|^4 = \frac{N(N-1)}{2N^4} E_{\mathbb{P}^N} [u_N^4 (\xi_{12}, V^1, V^2)] \to 0.
\]

Hence the Lyapunov's condition for CLT (see [57], Theorem 27.3) holds with \( \delta = 2 \):

\[
\lim_{N \to \infty} \frac{1}{\sigma_N^{2+\delta}} \sum_{1 \leq i < j \leq N} E_{\mathbb{P}^N, V} \left| \frac{u_N (\xi_{ij}, V^i, V^j)}{N} \right|^{2+\delta} = 0,
\]

where the convergence is in probability. It then follows from standard proofs of CLT and a subsequence argument that for each \( t \in \mathbb{R} \),

\[
E_{\mathbb{P}^N, V} \left[ e^{itU_N} \right] - e^{-\frac{1}{2}t^2 \sigma_N^2} \to 0
\]

in probability as \( N \to \infty \), which together with the convergence of \( \sigma_N^2 \to \sigma^2 \) implies that

\[
E_{\mathbb{P}^N, V} \left[ e^{itU_N} \right] \to e^{-t^2 \sigma^2/2}
\]

in probability as \( N \to \infty \). Now let \( \varphi_N(t, s) \) be the characteristic function of \( (U_N, \tilde{U}_N(f)) \), and \( \varphi(t, s) = e^{-\frac{1}{2}t^2 \sigma^2} \psi(s) \) be that of \( (Z, I_2(f)) \). It follows from Theorem 2.1 that as \( N \to \infty \),

\[
E_{\mathbb{P}^N} [e^{i t \tilde{U}_N(f)}] \to \psi(s).
\]
So we have

\[ \varphi_N(t, s) - \varphi(t, s) = E_{\mathbb{P}^N} \left[ e^{itU_N + is\tilde{U}_N} - e^{-\frac{1}{2}t^2\sigma^2} \psi(s) \right] \]

\[ = E_{\mathbb{P}^N} \left[ (E_{\mathbb{P}^N, V} [e^{itU_N}] - e^{-\frac{1}{2}t^2\sigma^2}) e^{is\tilde{U}_N} \right] + \left( E_{\mathbb{P}^N} [e^{is\tilde{U}_N} - \psi(s)] \right) e^{-\frac{1}{2}t^2\sigma^2}, \]

which converges to 0 as \( N \to \infty \). This completes the proof.

It follows from Condition 5.7 and law of large numbers that as \( N \to \infty \)

\[ \sum_{i=1}^{N} \int_{0}^{T} \frac{1}{Np_N(s)} \beta_s(X^i_s, X^j_s) \cdot dW^i_s \Rightarrow 0. \]

Combining above display with Lemmas 5.13 and 5.14, we have as \( N \to \infty \)

\[ J^{N,1}(T) \Rightarrow Z + I_2(h^{sym}). \]

\[ \text{(5.23)} \]

5.4.2 Asymptotics of \( J^{N,2} \)

Recall the definition of \( J^{N,2} \) in (5.16). We split \( J^{N,2} \) as follows:

\[ J^{N,2}(T) = \sum_{i,j,k=1}^{N} \int_{0}^{T} \frac{\xi_{ij}(s)\xi_{ik}(s)}{N^2(s)} \beta_s(X^i_s, X^j_s) \cdot \beta_s(X^i_s, X^k_s) \, ds \]

\[ = \sum_{n=1}^{5} \sum_{(i,j,k) \in S_n} \int_{0}^{T} \frac{\xi_{ij}(s)\xi_{ik}(s)}{N^2(s)} \beta_s(X^i_s, X^j_s) \cdot \beta_s(X^i_s, X^k_s) \, ds \]

\[ = \sum_{n=1}^{5} \tilde{T}^{N,n}, \]

where \( S_1, S_2, S_3, S_4 \) and \( S_5 \) are collections of \( (i,j,k) \in \mathbb{N}^3 \) such that \( \{i = j = k\}, \{i = j \neq k\}, \{i = k \neq j\}, \{j = k \neq i\} \) and \( \{i,j,k \text{ distinct}\} \), respectively. For \( \tilde{T}^{N,1} \), we have

\[ E_{\mathbb{P}^N} |\tilde{T}^{N,1}| = E_{\mathbb{P}^N} \sum_{i=1}^{N} \int_{0}^{T} \frac{1}{N^2_i(s)} \|\beta_s(X^i_s, X^j_s)\|^2 \, ds \leq \kappa N \int_{0}^{T} E_{\mathbb{P}^N} \frac{1}{N^2_i(s)} \, ds \]

\[ \leq \kappa N \int_{0}^{T} \frac{1}{N^2 p_N(s)} \, ds \leq \frac{\kappa N}{N^2 p_N} \to 0 \]

as \( N \to \infty \), where the last inequality follows from Lemma 5.9.
For studying the asymptotics of $\tilde{T}^{N,2}$, $\tilde{T}^{N,3}$, $\tilde{T}^{N,4}$ and $\tilde{T}^{N,5}$, we will need the following lemma.

**Lemma 5.15.** Suppose functions $\gamma_s \in L^2_c(\mu_s^{\otimes 3})$, $\vartheta_s \in L^2_c(\mu_s^{\otimes 2})$ and $\rho_s \in L^2_c(\mu_s)$ are uniformly bounded in $s \in [0, T]$. Then as $N \to \infty$

\[
\begin{align*}
\mathbb{E}_P^N \left| \sum_{(i,j,k) \in S_5} \int_0^T \xi_{ij}(s)\xi_{ik}(s) \frac{\gamma_s(X^i_s, X^j_s, X^k_s)}{N^2_i(s)} \, ds \right| & \to 0, \quad (5.24) \\
\mathbb{E}_P^N \left| \sum_{1 \leq i < k \leq N} \int_0^T \xi_{ik}(s) \vartheta_s(X^i_s, X^k_s) \, ds \right| & \to 0, \quad (5.25) \\
\mathbb{E}_P^N \left| \sum_{1 \leq i < k \leq N} \int_0^T \xi_{ik}(s) \rho_s(X^k_s) \, ds \right| & \to 0, \quad (5.26) \\
\mathbb{E}_P^N \left| \sum_{1 \leq i < k \leq N} \int_0^T \xi_{ik}(s) \rho_s(X^i_s) \, ds \right| & \to 0. \quad (5.27)
\end{align*}
\]

**Proof.** To prove (5.24), it is enough to prove the convergence with the summation over $S_5$ replaced by the ordered sum $i < j < k$. Now note that

\[
\begin{align*}
\mathbb{E}_P^N \left( \sum_{1 \leq i < j < k \leq N} \int_0^T \xi_{ij}(s)\xi_{ik}(s) \frac{\gamma_s(X^i_s, X^j_s, X^k_s)}{N^2_i(s)} \, ds \right)^2 \\
= \mathbb{E}_P^N \sum_{1 \leq i < j < k \leq N} \left( \int_0^T \xi_{ij}(s)\xi_{ik}(s) \frac{\gamma_s(X^i_s, X^j_s, X^k_s)}{N^2_i(s)} \, ds \right)^2 \\
\leq \kappa N^3 \int_0^T \mathbb{E}_P^N \frac{1}{N^4_i(s)} \, ds \\
\leq \kappa N^3 \int_0^T \frac{1}{N^{4p_N^2}(s)} \, ds \\
\leq \frac{\kappa}{Np_N^4} \\
\to 0,
\end{align*}
\]

where the second inequality follows from Lemma 5.9. Thus (5.24) holds. Proofs for (5.25), (5.26) and (5.27) are similar and hence omitted. \qed
For $\hat{T}^{N,2}$, note that

$$\hat{T}^{N,2} = \sum_{1 \leq i \neq k \leq N} \int_0^T \frac{\xi_{ik}(s)}{N_i^2(s)} \beta_s(X^i_s, X^k_s) \beta_s(X^i_s, X^k_s) ds$$

$$= \sum_{1 \leq i \neq k \leq N} \int_0^T \frac{\xi_{ik}(s)}{N_i^2(s)} \left( \beta_s(X^i_s, X^i_s) \beta_s(X^i_s, X^k_s) - \int_{\mathbb{R}^d} \beta_s(y, y) \beta_s(y, X^k_s) \mu_s(dy) \right) ds$$

$$+ \sum_{1 \leq i \neq k \leq N} \int_0^T \frac{\xi_{ik}(s)}{N_i^2(s)} \int_{\mathbb{R}^d} \beta_s(y, y) \beta_s(y, X^k_s) \mu_s(dy) ds.$$

It then follows from Lemma 5.15 that $E_{P_N} |\hat{T}^{N,2}| \to 0$ as $N \to \infty$, and similarly $E_{P_N} |\hat{T}^{N,3}| \to 0$.

Consider now the fourth term $\hat{T}^{N,4}$. Note that

$$\hat{T}^{N,4} = \sum_{1 \leq i \neq j \leq N} \int_0^T \frac{\xi_{ij}(s)}{N_i^2(s)} \left( \beta_s(X^i_s, X^j_s) \right)^2 ds$$

$$= \sum_{1 \leq i \neq j \leq N} \int_0^T \frac{\xi_{ij}(s)}{N_i^2(s)} \left( \beta_s(x^i_s, x^j_s) \right)^2 ds$$

$$+ \sum_{1 \leq i \neq j \leq N} \int_0^T \frac{\xi_{ij}(s)}{N_i^2(s)} \left( \beta_s(x^i_s, \cdot) \right)^2 ds$$

$$+ \sum_{1 \leq i \neq j \leq N} \int_0^T \frac{\xi_{ij}(s)}{N_i^2(s)} \left( \beta_s(\cdot, x^j_s) \right)^2 ds + \sum_{1 \leq i \neq j \leq N} \int_0^T \frac{\xi_{ij}(s)}{N_i^2(s)} \lambda_s ds$$

$$= \sum_{n=1}^4 \hat{T}^{N,4}_n.$$

It follows from Lemma 5.15 that as $N \to \infty$, $E_{P_N} |\hat{T}^{N,4}_n| \to 0$ for $n = 1, 2, 3, 4$. We claim that

$$E_{P_N} \left| \int_0^T \frac{1}{p(s)} \lambda_s ds \right| \to 0$$

as $N \to \infty$. To see this, first write

$$\hat{T}^{N,4}_4 = \sum_{i=1}^N \int_0^T \frac{N_i(s) - 1}{N_i^2(s)} \lambda_s ds = \sum_{i=1}^N \int_0^T \frac{1}{N_i^2(s)} \lambda_s ds - \sum_{i=1}^N \int_0^T \frac{1}{N_i^2(s)} \lambda_s ds.$$

Next, by Lemma 5.12

$$\lim_{N \to \infty} E_{P_N} \sum_{i=1}^N \int_0^T \frac{1}{N_i^2(s)} \lambda_s ds = \lim_{N \to \infty} \int_0^T E_{P_N} \frac{N}{N_i^2(s)} \lambda_s ds = 0.$$
Also as \( N \to \infty \),
\[
\mathbb{E}_{\mathbb{P}^N} \left| \sum_{i=1}^{N} \int_0^T \frac{1}{N_i(s)} \lambda_s \, ds - \int_0^T \frac{1}{p(s)} \lambda_s \, ds \right|^2 \leq \kappa N \mathbb{E}_{\mathbb{P}^N} \sum_{i=1}^{N} \int_0^T \left( \frac{1}{N_i(s)} - \frac{1}{N p(s)} \right)^2 \lambda_s \, ds
\]
\[
\leq \kappa N^2 \int_0^T \mathbb{E}_{\mathbb{P}^N} \left( \frac{1}{N_i(s)} - \frac{1}{N p(s)} \right)^2 \, ds
\]
\[
\to 0,
\]
where the convergence follows from Lemma 5.12 and Condition 5.7. This proves the claim (5.28) and hence as \( N \to \infty \),
\[
\tilde{T}^{N,4} \Rightarrow \int_0^T \frac{1}{p(s)} \lambda_s \, ds.
\]

Finally consider the last term \( \tilde{T}^{N,5} \). Let \( m_t(x,y) = \int_{\mathbb{R}^d} \beta_t(z,x) \cdot \beta_t(z,y) \mu_t(dz) \) for \( x,y \in \mathbb{R}^d \) and \( t \in [0,T] \). Then
\[
\tilde{T}^{N,5} = \sum_{(i,j,k) \in S_5} \int_0^T \frac{\xi_{ij}(s) \xi_{ik}(s)}{N_i^2(s)} \beta_s(X^i_s, X^j_s) \cdot \beta_s(X^i_s, X^k_s) \, ds
\]
\[
= \sum_{(i,j,k) \in S_5} \int_0^T \frac{\xi_{ij}(s) \xi_{ik}(s)}{N_i^2(s)} \left( \beta_s(X^i_s, X^j_s) \cdot \beta_s(X^i_s, X^k_s) - m_s(X^j_s, X^k_s) \right) \, ds
\]
\[
+ \sum_{(i,j,k) \in S_5} \int_0^T \frac{\xi_{ij}(s) \xi_{ik}(s)}{N_i^2(s)} m_s(X^j_s, X^k_s) \, ds
\]
\[
= \tilde{T}^{N,5}_1 + \tilde{T}^{N,5}_2.
\]

It follows from Lemma 5.15 that \( \mathbb{E}_{\mathbb{P}^N} \left| \tilde{T}^{N,5}_1 \right| \to 0 \) as \( N \to \infty \). Let
\[
\tilde{T}^{N,5} = \sum_{(i,j,k) \in S_5} \int_0^T \frac{\xi_{ij}(s) \xi_{ik}(s)}{N^2 p_N(s)} m_s(X^j_s, X^k_s) \, ds.
\]

We claim that as \( N \to \infty \),
\[
\mathbb{E}_{\mathbb{P}^N} \left| \tilde{T}^{N,5}_2 - \tilde{T}^{N,5} \right|^2 \to 0. \tag{5.29}
\]
To see this, as before, it suffices to consider the summation over ordered indices \( i < j < k \). Note that

\[
E_{p_N} \left( \sum_{1 \leq i < j < k \leq N} \int_0^T \xi_{ij}(s) \xi_{ik}(s) \left( \frac{1}{N_i^2(s)} - \frac{1}{N_i^2 p_N^2(s)} \right) m_s(X^j_s, X^k_s) \, ds \right)^2
\]

\[
\leq N E_{p_N} \sum_{i=1}^N \left( \sum_{1 \leq j < k \leq N} \int_0^T \xi_{ij}(s) \xi_{ik}(s) \left( \frac{1}{N_i^2(s)} - \frac{1}{N_i^2 p_N^2(s)} \right) m_s(X^j_s, X^k_s) \, ds \right)^2
\]

\[
= N E_{p_N} \sum_{i=1}^N \sum_{1 \leq j < k \leq N} \left( \int_0^T \xi_{ij}(s) \xi_{ik}(s) \left( \frac{1}{N_i^2(s)} - \frac{1}{N_i^2 p_N^2(s)} \right) m_s(X^j_s, X^k_s) \, ds \right)^2
\]

\[
\leq \kappa N^4 E_{p_N} \int_0^T \left( \frac{1}{N_i^2(s)} - \frac{1}{N_i^2 p_N^2(s)} \right)^2 \, ds
\]

\[
\to 0
\]

as \( N \to \infty \), where the convergence follows from Lemma 5.12. So the claim (5.29) holds. Next split \( \hat{T}^{N,5} \) as

\[
\hat{T}^{N,5} = \sum_{(i,j,k) \in S_{N,j<k}} \int_0^T \xi_{ij}(s) \xi_{ik}(s) - p_N^2(s) \frac{m_s(X^j_s, X^k_s)}{N^2 p_N^2(s)} \, ds
\]

\[
+ \sum_{(i,j,k) \in S_{N,j>k}} \int_0^T \xi_{ij}(s) \xi_{ik}(s) - p_N^2(s) \frac{m_s(X^j_s, X^k_s)}{N^2 p_N^2(s)} \, ds
\]

\[
+ \sum_{(i,j,k) \in S_5} \frac{1}{N^2} \int_0^T m_s(X^j_s, X^k_s) \, ds
\]

\[
= \hat{T}^{N,5}_1 + \hat{T}^{N,5}_2 + \hat{T}^{N,5}_3.
\]

It follows from Condition 5.7 that as \( N \to \infty \)

\[
E_{p_N} |\hat{T}^{N,5}_1|^2 = \sum_{(i,j,k) \in S_{N,j<k}} E_{p_N} \left( \int_0^T \xi_{ij}(s) \xi_{ik}(s) - p_N^2(s) \frac{m_s(X^j_s, X^k_s)}{N^2 p_N^2(s)} \, ds \right)^2
\]

\[
\leq \kappa \sum_{(i,j,k) \in S_{N,j<k}} \int_0^T E_{p_N} \left( \frac{\xi_{ij}(s) \xi_{ik}(s) - p_N^2(s)}{N^2 p_N^2(s)} \right)^2 \, ds
\]

\[
\leq \frac{\kappa}{N p_N^2}
\]

\[
\to 0.
\]
Similarly $\mathbb{E}_{\mathbb{G}_N} |\hat{T}_2^{N,5}|^2 \to 0$ as $N \to \infty$. It follows from Theorem 2.1 that as $N \to \infty$,

$$\hat{T}_3^{N,5} = \sum_{1 \leq j \neq k \leq N} \frac{N-2}{N^2} \int_0^T m_s(X^j_s, X^k_s) \, ds \Rightarrow I_2(l),$$

where $l$ is defined as

$$l(\omega, \omega') = \int_0^T m_s(X^s, X^s(\omega, \omega')) \, ds, \quad (\omega, \omega') \in \Omega_d^2.$$

Hence $\hat{T}^{N,5} \Rightarrow I_2(l)$ as $N \to \infty$.

Combining above displays gives us as $N \to \infty$,

$$J^{N,2}(T) \Rightarrow I_2(l) + \int_0^T \frac{1}{p(s)} \lambda_s \, ds.$$

In fact by Theorem 2.1 and Lemma 5.14 we have as $N \to \infty$,

$$(J^{N,1}(T), J^{N,2}(T)) \Rightarrow \left( Z + I_2(h^{sym}), I_2(l) + \int_0^T \frac{1}{p(s)} \lambda_s \, ds \right), \quad (5.30)$$

and $Z$ is independent of $(I_2(h^{sym}), I_2(l))$.

**5.4.3 Combining $J^{N,1}$ and $J^{N,2}$ and completing the proof**

It follows from (5.30) that as $N \to \infty$,

$$J^N(T) \Rightarrow \frac{1}{2} I_2(f) + Z - \int_0^T \frac{1}{2p(s)} \lambda_s \, ds \equiv J,$$

where $f$ is defined as

$$f(\omega, \omega') = h(\omega, \omega') + h(\omega', \omega) - l(\omega, \omega'), \quad (\omega, \omega') \in \Omega_d^2$$

and $Z$ is independent of $I_2(f)$. Thanks to such independence, recalling $\text{Trace}(AA^*)$ introduced in Lemma 5.6 and $\sigma^2$ defined in (5.22), we have

$$\mathbb{E} \exp(J) = \mathbb{E} \exp \left( \frac{1}{2} I_2(f) - \frac{1}{2} \text{Trace}(AA^*) \right) \mathbb{E} \exp \left( Z - \frac{1}{2} \sigma^2 \right) = 1,$$

83
where the last equality follows from Lemma 2.2. So \( \{\exp(J^N_T)\} \) is uniformly integrable and consequently so is \( \{\exp(i\tilde{\eta}^N_N(\phi)+J^N_T)\} \). Hence, using the independence between \( Z \) and \((I_2(f), I_1(\phi))\) we have as \( N \to \infty \)

\[
\mathbb{E}_N \exp \left( i\tilde{\eta}^N_N(\phi) + J^N_T \right) \to \mathbb{E} \exp \left( i I_1(\phi) + J \right)
= \mathbb{E} \exp \left( i I_1(\phi) + \frac{1}{2} I_2(f) - \frac{1}{2} \text{Trace}(AA^*) \right) \mathbb{E} \exp \left( Z - \frac{1}{2} \sigma^2 \right)
= \exp \left( -\frac{1}{2} \| (I - A)^{-1} \phi \|_{L^2(\Omega_{d,v})}^2 \right),
\]

where the last equality again follows from Lemma 2.2. Hence we have (5.17), which completes the proof of Theorem 5.8. \( \square \)

5.5 Proof of Corollary 5.5 and preparatory results

(a) For each fixed \( g \in C_b(C_d) \) and \( x \in C_d \), let \( g^c(x) = g(x) - \langle g, \mu \rangle \). Then

\[
\mathbb{E} \left( \langle g, \mu^N \rangle - \langle g, \mu \rangle \right)^2 = \mathbb{E} \left( \frac{1}{N} \sum_{i=1}^N g^c(Z^i,N) \right)^2
= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \mathbb{E} \left( g^c(Z^i,N) g^c(Z^j,N) \right)
= \frac{N-1}{N} \mathbb{E} \left( g^c(Z^1,N) g^c(Z^2,N) \right) + \frac{1}{N} \mathbb{E} \left( g^c(Z^1,N) \right)^2,
\]

which converges to 0 as \( N \to \infty \) by Corollary 5.4. This gives part (a).

(b) Fix \( i \in \mathbb{N} \). Write \( \xi^N_i(0) \) and \( N_i(0) \) as \( \zeta_{ij} \) and \( N_i \) for simplicity. Let \( \tilde{\mu}^{i,N} = \frac{1}{N_i} \sum_{j=1}^N \zeta_{ij} \delta_{X^j} \).
It suffices to show that \( d_{BL}(\mu^{i,N}, \tilde{\mu}^{i,N}) \to 0 \) and \( \tilde{\mu}^{i,N} \Rightarrow \mu \) as \( N \to \infty \) (cf. Theorem 3.1 in Billingsley [4]). Note that

\[
\mathbb{E} d_{BL}(\mu^{i,N}, \tilde{\mu}^{i,N}) = \mathbb{E} \sup_{\|g\|_{BL} \leq 1} \left| \langle g, \mu^{i,N} \rangle - \langle f, \tilde{\mu}^{i,N} \rangle \right|
= \mathbb{E} \sup_{\|g\|_{BL} \leq 1} \left| \sum_{j=1}^N \frac{\zeta_{ij}}{N_i} \left( g(Z^{j,N}) - g(X^j) \right) \right|
\leq \sum_{j=1}^N \frac{\zeta_{ij}}{N_i} \left\| Z^{j,N} - X^j \right\|_{*,T}.
\]

84
Since $L(\zeta_{ij}, N_i, Z_j, X_j) = L(\zeta_{ji}, N_j, Z_i, X_i)$, we have from the above display that

$$E d_{BL}(\mu_{i,N}, \tilde{\mu}_{i,N}) \leq \sum_{j=1}^{N} E \left( \frac{\zeta_{ij}}{N_j} - 1 \right) \|Z_{i,N} - X_i\|_{s,T} + E \|Z_{i,N} - X_i\|_{s,T}$$

$$\leq \sqrt{\frac{3K T}{N_p}} E \|Z_{i,N} - X_i\|_{s,T} + E \|Z_{i,N} - X_i\|_{s,T}$$

where the last inequality follows from Lemma 5.10 and (5.12). It then follows from Theorem 5.2 that $d_{BL}(\mu_{i,N}, \tilde{\mu}_{i,N}) \to 0$ as $N \to \infty$. So $d_{BL}(\mu_{i,N}, \tilde{\mu}_{i,N}) \Rightarrow 0$ as $N \to \infty$.

Next we show that $\tilde{\mu}_{i,N} \Rightarrow \mu$ as $N \to \infty$. For each fixed $g \in C_b(C_d)$ and $x \in C_d$, let $g^c(x) = g(x) - \langle g, \mu \rangle$. Then

$$E \left( \langle g, \tilde{\mu}_{i,N} \rangle - \langle g, \mu \rangle \right)^2 = E \left( \sum_{j=1}^{N} \frac{\zeta_{ij}}{N_i} g(X_j) - \langle g, \mu \rangle \right)^2$$

$$= E \left( \sum_{j=1}^{N} \frac{\zeta_{ij}}{N_i} g^c(X_j) \right)^2$$

$$= \sum_{j=1}^{N} E \frac{\zeta_{ij}}{N_i^2} (g^c(X_j))^2$$

$$\leq 4\|g\|^2 \sum_{j=1}^{N} E \frac{\zeta_{ij}}{N_i^2}$$

$$= 4\|g\|^2 \frac{1}{N_i}$$

$$\leq \frac{4\|g\|^2}{N_p N}$$

$$\to 0$$

as $N \to \infty$, where the last inequality follows from Lemma 5.9. Hence $\tilde{\mu}_{i,N} \Rightarrow \mu$ as $N \to \infty$ and part (b) follows. \qed
Proof of Lemma 5.9

First note that
\[
\mathbb{E} \frac{1}{X + 1} = \sum_{k=0}^{n} \frac{1}{k+1} \binom{n}{k} p^k q^{n-k} = \frac{1}{(n+1)p} \sum_{k=0}^{n} \binom{n+1}{k+1} p^{k+1} q^{n-k} = \frac{1-q^{n+1}}{(n+1)p}.
\]

Hence for each \(m \in \mathbb{N}\),
\[
\mathbb{E} \frac{1}{X + m} = \mathbb{E} \frac{1}{X + 1} \frac{1}{X + m} \leq \mathbb{E} \frac{1}{X + 1} \frac{n+1}{n+m} = \frac{1-q^{n+1}}{(n+1)p} \frac{n+1}{n+m} \leq \frac{1}{(n+m)p}.
\]

Similarly,
\[
\mathbb{E} \frac{1}{(X+1)^m} \leq \mathbb{E} \frac{m^m}{(X+1)(X+2) \cdots (X+m)} = \sum_{k=0}^{m} \frac{m^m}{(k+1)(k+2) \cdots (k+m)} \binom{n}{k} p^k q^{n-k} \\
\leq \frac{m^m}{(n+1)(n+2) \cdots (n+m)p^m} \leq \frac{m^m}{(n+1)^m p^m},
\]

which completes the proof.

Proof of Lemma 5.10

For \(i \in \mathbb{N}\), let \(\tilde{N}_{i,-i} = \sum_{j \neq i} \zeta_{ij} = N_i - 1\). For different \(i,k \in \mathbb{N}\), let \(\tilde{N}_{i,-ik} = \sum_{j \notin \{i,k\}} \zeta_{ij}\). For distinct \(i,j,k \in \mathbb{N}\), let \(\tilde{N}_{i,-ijk} = \sum_{l \notin \{i,j,k\}} \zeta_{il}\). Fixing \(i \in \mathbb{N}\), write
\[
\mathbb{E} \left( \frac{\sum_{j=1}^{N} \zeta_{ij}}{N_j} - 1 \right)^2 = \mathbb{E} \left( \sum_{j=1}^{N} \frac{\zeta_{ij}}{N_j} \right)^2 - 2 \mathbb{E} \sum_{j=1}^{N} \frac{\zeta_{ij}}{N_j} + 1 = \mathbb{E} \sum_{j,k=1}^{N} \frac{\zeta_{ij} \zeta_{ik}}{N_j N_k} - 1 \\
= \sum_{n=1}^{4} \mathbb{E} \sum_{(j,k) \in S_{i}^{N,n}} \frac{\zeta_{ij} \zeta_{ik}}{N_j N_k} - 1 = \sum_{n=1}^{4} \mathbb{E} T_{i}^{N,n} - 1,
\]

(5.31)
where $S_{i}^{N,1}$, $S_{i}^{N,2}$, $S_{i}^{N,3}$ and $S_{i}^{N,4}$ are collections of $(j, k) \in \mathbf{N} \times \mathbf{N}$ such that \{j = k\}, \{j \neq k, j = i\}, \{j \neq k, k = i\} and \{i, j, k\} distinct}, respectively. For $T_{i}^{N,1}$, since $L(\zeta_{ij}, N) = L(\zeta_{ji}, N)$,

$$T_{i}^{N,1} = \sum_{j=1}^{N} E \frac{\zeta_{ij}}{N^{2}} = \sum_{j=1}^{N} E \frac{\zeta_{ij}}{N^{2}} = E \frac{1}{N^{2}} \leq \frac{1}{NpN}, \quad (5.32)$$

where the inequality follows from Lemma 5.9. For $T_{i}^{N,2}$, using independence of $\tilde{N}_{i, -ik}$ and $\tilde{N}_{k, -ki}$ for $i \neq k$, we have

$$T_{i}^{N,2} = \sum_{k=1, k \neq i}^{N} E \frac{\zeta_{ik}}{N_{i}N_{k}}$$

$$= \sum_{k=1, k \neq i}^{N} E \frac{\zeta_{ik}}{(1 + \zeta_{ik} + \tilde{N}_{i, -ik})(1 + \zeta_{ki} + \tilde{N}_{k, -ki})}$$

$$= \sum_{k=1, k \neq i}^{N} p_{N} E \frac{1}{(2 + \tilde{N}_{i, -ik})(2 + \tilde{N}_{k, -ki})}$$

$$= \sum_{k=1, k \neq i}^{N} p_{N} \left( E \frac{1}{2 + \tilde{N}_{i, -ik}} \right)^{2}$$

$$\leq (N - 1)p_{N} \frac{1}{(NP_{N})^{2}}$$

$$\leq \frac{1}{NP_{N}}, \quad (5.33)$$

where the first inequality again follows from Lemma 5.9. Similarly for $T_{i}^{N,3}$, we have

$$T_{i}^{N,3} = \sum_{j=1, j \neq i}^{N} E \frac{\zeta_{ij}}{N_{i}N_{j}} \leq \frac{1}{NP_{N}}. \quad (5.34)$$

Finally for $T_{i}^{N,4}$, note that

$$T_{i}^{N,4} = \sum_{(j, k) \in S_{i}^{N,4}} E \frac{\zeta_{ij}\zeta_{ik}}{N_{j}N_{k}}$$

$$= \sum_{(j, k) \in S_{i}^{N,4}} p_{N}^{2} E \frac{1}{(2 + \tilde{N}_{j, -ji})(2 + \tilde{N}_{k, -ki})}$$

$$\leq \sum_{(j, k) \in S_{i}^{N,4}} p_{N}^{2} E \frac{1}{(2 + \tilde{N}_{j, -jik})(2 + \tilde{N}_{k, -kij})}.$$
Using independence of $\tilde{N}_{j,-jik}$ and $\tilde{N}_{j,-jik}$ for distinct $i,j,k$, we have from the above display that

$$T_i^{N,4} \leq \sum_{(j,k) \in S_i^{N,4}} p_N^2 \left( E \frac{1}{2 + \tilde{N}_{j,-jik}} \right) \left( E \frac{1}{2 + \tilde{N}_{k,-kij}} \right)$$

$$\leq (N - 1)(N - 2)p_N^2 \left( \frac{1}{(N - 1)p_N} \right)^2$$

$$\leq 1,$$

(5.35)

where the second inequality once more follows from Lemma 5.9. Plugging (5.32) – (5.35) into (5.31) completes the proof.

Proof of Lemma 5.11

First note that the result holds trivially when $p_N = 0$ or $p_N = 1$. Now consider the case $p_N \in (0, 1)$. For $t \geq 0$, it follows from Hoeffding’s inequality that

$$P \left( \left| N_1 - Np_N \right| > t + 1 \right) \leq P \left( \left| \sum_{j=2}^{N} (\zeta_{ij} - p_N) \right| > t \right) \leq 2e^{-\frac{2t^2}{N-1}}.$$
CHAPTER 6
MODERATE DEVIATION PRINCIPLES FOR WEAKLY INTERACTING DIFFUSIONS

In Chapters 3, 4 and 5 we studied central limit fluctuations about the law of large numbers limit for certain weakly interacting particle systems, while in this and later chapters we are interested in studying large and moderate deviations.

Large deviation principles (LDP) for weakly interacting particle systems have been well studied in many works. A classical reference is [25] which considers a collection of diffusing particles with non-degenerate diffusion coefficients that interact through the drift terms. Proofs are based on discretization arguments together with careful exponential probability estimates. Alternative methods using weak convergence and certain variational representation formulas have recently been introduced in [11]. Large deviations for pure jump finite state weakly interacting particle systems have been studied in [51, 6, 28]. Large deviations for certain weakly interacting jump-diffusion models with a common factor have been studied in [63].

In this and the next chapters, our focus is on the study of deviations, from the law of large numbers limit for weakly interacting systems, that are of smaller order than those captured by a large deviation principle. Results that give asymptotics of such lower order deviations are usually referred to as moderate deviation principles (MDP). The object of our interest is the empirical measure process \( \mu^m(t) = \frac{1}{m} \sum_{i=1}^{m} \delta_{X_i^m(t)} \). Denoting the state space of particles by \( S \), \( \mu^m(t) \) is a random measure, with values in \( \mathcal{P}(S) \). In order to motivate the problem of interest, we consider as an illustration the setting where the particle distributions are i.i.d. Let \( \{Y_i\}_{i \in \mathbb{N}} \) be an i.i.d. sequence of \( \mathbb{R}^d \)-valued random variables with distribution \( \mu \). Sanov’s theorem that gives a LDP for \( L_m = \frac{1}{m} \sum_{i=1}^{m} \delta_{Y_i} \) formally says that for any measurable set \( U \) in \( \mathcal{P}(\mathbb{R}^d) \),

\[
P(L_m \in U) \approx \exp\{-m \inf_{\nu \in U} I(\nu)\},
\]

89
where for $\nu \in \mathcal{P}(\mathbb{R}^d)$, $I(\nu) = R(\nu \| \mu) \doteq \int_{\mathbb{R}^d} \frac{d\nu}{d\mu} \left( \log \frac{d\nu}{d\mu} \right) d\mu$ is the relative entropy ($R(\nu \| \mu)$ is taken to be $\infty$ if $\nu$ is not absolutely continuous with respect to $\mu$). Now let $\{a(m)\}_{m \in \mathbb{N}}$ be a positive sequence such that as $m \to \infty$,

$$a(m) \to 0 \text{ and } a(m)\sqrt{m} \to \infty \quad (6.1)$$

(e.g. $a(m) = m^{-\theta}$ for some $\theta \in (0, 1/2)$). A moderate deviation principle for $\{L_m\}$, associated with deviations of order $\frac{1}{a(m)\sqrt{m}}$, gives a result of the following form (see e.g. [7]): for any measurable set $U$ in $\mathcal{P}_0(\mathbb{R}^d)$ (the space of signed measures on $\mathbb{R}^d$ such that $\nu(\mathbb{R}^d) = 0$),

$$\mathbb{P}(a(m)\sqrt{m}(L_m - \mu) \in U) \approx \exp \left\{ -\frac{1}{a^2(m)} \inf_{\nu \in U} I^0(\nu) \right\},$$

where for each $\nu \in \mathcal{P}_0(\mathbb{R}^d)$, $I^0(\nu) \doteq \frac{1}{2} \int_{\mathbb{R}^d} \left( \frac{d\nu}{d\mu} \right)^2 d\mu$ (once again we take $I^0(\nu) \doteq \infty$ if $\nu$ is not absolutely continuous with respect to $\mu$). Note that CLT and LDP provide asymptotics for the probabilities on the left side when $a(m) = m^{-\theta}$ with $\theta = 0$ and $\frac{1}{2}$ respectively, whereas a MDP studies an asymptotic regime where $\theta \in (0, \frac{1}{2})$ (a MDP also treats more general scale functions $a(m)$). There is an extensive literature on moderate deviation results in mathematical statistics, including results for i.i.d. sequences and arrays, empirical processes in general topological spaces, weakly dependent sequences, and occupation measures of Markov chains together with general additive functionals of Markov chains (see [12] for many such references). MDP for small noise finite and infinite dimensional SDE with jumps have been studied in [12]. References to other MDP results for SDE in the context of stochastic averaging and multi-scale systems can be found in [12]. For weakly interacting particle systems in discrete time, MDP based on semigroup analysis and projective large deviation methods have been established in [26].

In this and the next chapters we will study moderate deviation principles for continuous time weakly interacting Markov processes. The models we consider will allow for both Brownian and Poisson type noises in the dynamics. Our approach is based on certain variational representations for exponential functionals of such noise processes developed in [8, 10, 13]. In order to keep the presentation simple we consider two types of models: one that corresponds to pure jump interacting Markov processes (in Chapter 7) and the other that considers interacting Markov processes with
continuous sample paths and Brownian noise (in this chapter). Although not treated here, one can use similar techniques to develop moderate deviation results for settings that have both Brownian and Poisson type noises.

For the diffusion model considered here, we allow for state dependence in both drift and diffusion coefficients and the interaction through the empirical measure appears in both coefficients as well. Coefficients are assumed to be bounded with suitable smoothness properties but non-degeneracy of the diffusion term is not required. This is in contrast to the classical results on large deviation principles for such systems (e.g. [25]) which only allow interaction in the drift and require the diffusion coefficient to be uniformly non-degenerate (however see [11] and [63] for results that relax some of these conditions). In order to highlight the main ideas we restrict attention to a one dimensional setting, i.e. the case where the state space of the particles is $\mathbb{R}$. The general multidimensional case can be treated in a similar manner. Specifically, we consider a collection of one dimensional weakly interacting diffusions $\{X^m_i\}_{i=1}^m$ given by the system of equations:

$$X^m_i(t) = x_0 + \int_0^t \sigma(X^m_i(s), \mu^m(s)) \, dW_i(s) + \int_0^t b(X^m_i(s), \mu^m(s)) \, ds, \quad i = 1, \ldots, m,$$

where $\{W_i\}$ is a sequence of i.i.d. one-dimensional standard $\{F_t\}$-Brownian motions given on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ and $\mu^m(t) \doteq \frac{1}{m} \sum_{i=1}^m \delta_{X^m_i(t)}$. Here for $\theta \in \mathcal{P}(\mathbb{R})$, $\sigma(x, \theta) \doteq \int_{\mathbb{R}} \alpha(x,y) \theta(dy)$ and $b(x, \theta) \doteq \int_{\mathbb{R}} \beta(x,y) \theta(dy)$, where $\alpha$ and $\beta$ are bounded and Lipschitz. From [67] it follows that as $m \to \infty$, $\mu^m$ converges in $C([0,T] : \mathcal{P}(\mathbb{R}))$ (the space of $\mathcal{P}(\mathbb{R})$-valued continuous functions, equipped with the usual uniform topology and any metric on $\mathcal{P}(\mathbb{R})$ metrizing the weak topology and making it a Polish space), in probability, to $\mu$ where $\mu(t)$ is the common probability distribution of the i.i.d. collection $\{\bar{X}_i(t)\}$ governed by the equation

$$\bar{X}_i(t) = x_0 + \int_0^t \sigma(\bar{X}_i(s), \mu(s)) \, dW_i(s) + \int_0^t b(\bar{X}_i(s), \mu(s)) \, ds, \quad i \in \mathbb{N}.$$

A central limit theorem studying the asymptotics of the fluctuation process of signed measures $S^m(t) \doteq \sqrt{m}(\mu^m(t) - \mu(t))$ has been established in [39, 54]. As is well understood (cf. [39]), $S^m$ is very irregular as a signed measure-valued process as $m$ becomes large and one cannot expect the limit in general to be a measure-valued process. A common approach is to regard $S^m(t, \cdot)$ as an
element of a suitable distribution space. For example, it is shown in [39] that, under conditions, $S^m$ converges in distribution as a sequence of $\mathbb{C}([0, T]: S')$-valued random variables, where $S'$ is the dual of the Schwartz space $S$, namely the space of rapidly decaying infinitely smooth functions on $\mathbb{R}$. This space is equipped with the usual topology given in terms of a countable collection of Hilbertian seminorms $\{\| \cdot \|_n\}_{n \in \mathbb{Z}}$ with associated Hilbert spaces $\{S_n\}_{n \in \mathbb{Z}}$. We refer the reader to Section 6.1 for some basic background on the Schwartz space, but for now it suffices to note the following properties of the collection of Hilbert spaces $\{S_n\}_{n \in \mathbb{Z}}$: $S^w \subset S^v$ for $w \geq v$, $S' = \bigcup_{n \in \mathbb{Z}} S_n$ and $S = \bigcap_{n \in \mathbb{Z}} S_n$. The paper [39] shows that (see Theorem 1 therein), under suitable conditions, for some $v \in \mathbb{N}$, $S^m$ converges in $\mathbb{C}([0, T]: S^{-v})$, in distribution, as $m \to \infty$, to $S$ given as the solution of

$$S(t) = Z(t) + \int_0^t L^*(s)S(s) \, ds. \tag{6.4}$$

Here $Z$ is an $S_{-(v+2)}$-valued Gaussian process with an explicit covariance operator (see (4.2) in [39]) and $L^*(s)$ is the adjoint of the operator $L(s)$ defined as (see also (1.4) in [39])

$$\langle L(s) \phi(x) \rangle = \phi'(x)b(x, \mu(s)) + \frac{1}{2} \phi''(x)\sigma^2(x, \mu(s))$$

$$+ \int_{\mathbb{R}} \phi'(y)\beta(y, x) \mu_s(dy) + \int_{\mathbb{R}} \phi''(y)\sigma(y, \mu(s))\alpha(y, x) \mu_s(dy), \phi \in \mathcal{S}. \tag{6.5}$$

Under suitable smoothness conditions on the coefficients, $L(s)$ can be regarded as a bounded linear operator from $S_{v+2}$ to $S_v$ and thus $L^*(s)$ is a bounded linear operator from $S_{-v}$ to $S_{-(v+2)}$. Equation (6.4) is interpreted as follows: For all $\phi \in \mathcal{S}$,

$$\langle S(t), \phi \rangle = \langle Z(t), \phi \rangle + \int_0^t \langle S(s), L(s)\phi \rangle \, ds.$$

We remark that [39] actually considers a modified version of the Schwartz distribution space which allows one to use unbounded test functions as well, however their results hold for the classical Schwartz space as presented above. Results of [39, 54] study deviations of $\mu^m$ from $\mu$ that are of order $1/\sqrt{m}$. In this chapter we will be concerned with deviations of $\mu^m$ from $\mu$ that are of higher order than $1/\sqrt{m}$ (but of lower order than $1/m$). Let $\{a(m)\}_{m \in \mathbb{N}}$ be as in (6.1) and

$$Y^m(t) = a(m)S^m(t) = a(m)\sqrt{m}(\mu^m(t) - \mu(t)). \tag{6.6}$$
We will show in Theorem 6.4 that under Conditions 6.1 and 6.3, $Y^m$ satisfies a large deviation principle with speed $a^2(m)$ (see Section 2.2 for a precise definition) in $C([0,T] : S_{-\rho})$ with a suitable value of $\rho > v$. Roughly speaking, this result says that for any Borel set $U$ in $C([0,T] : S_{-\rho})$ one has $P(Y^m \in U) \approx \exp\left\{-\frac{1}{a^2(m)} \inf_{\eta \in U} I(\eta)\right\}$, where $I$ is the associated rate function that will be introduced in (6.16). Since it provides asymptotics for probabilities of deviations of $\mu^m$ from $\mu$ that are of order $\frac{1}{a(m)\sqrt{m}}$, this LDP for $Y^m$ can be viewed as a moderate deviation principle for the empirical measure process $\mu^m$.

The chapter is organized as follows. In Section 6.1 we begin by describing our model of weakly interacting diffusions. Centered and normalized empirical measures are regarded as elements of a suitable distribution space. We introduce this space and note some of its basic properties. We then introduce the two main conditions (Conditions 6.1 and 6.3) for the MDP which is given in Theorem 6.4. Proof of this theorem is provided in Section 6.2.

6.1 Model and main results

In this section we consider the collection of weakly interacting diffusions $\{X^m_i\}_{i=1}^m$ described by (6.2). We are interested in the asymptotic behavior of $Y^m$ defined by (6.6). As noted in the introduction, we will regard $Y^m$ as a stochastic process with values in a suitable space of distributions. The natural space to consider is the standard Schwartz distribution space that is described as follows.

Let $S$ denote the space of functions $\phi: \mathbb{R} \to \mathbb{R}$ such that $\phi$ is infinitely differentiable and $|x|^m|\phi^{(k)}(x)| \to 0$ as $|x| \to \infty$ for every $m, k \in \mathbb{N}_0$, where $\phi^{(k)}$ denotes the $k$-th derivative of $\phi$. On $S$, define a sequence of inner product $\langle \cdot, \cdot \rangle_n$ and seminorms $\|\cdot\|_n$, $n \in \mathbb{N}_0$, as

$$\langle \phi, \psi \rangle_n = \sum_{0 \leq k \leq n} \int_{\mathbb{R}} (1 + x^2)^{2n} \phi^{(k)}(x) \psi^{(k)}(x) \, dx, \quad \|\phi\|^2_n = \langle \phi, \phi \rangle_n, \quad \phi, \psi \in S. \quad (6.7)$$

This sequence of seminorms introduces a nuclear Fréchet topology on $S$ (see e.g. Gel’fand and Vilenkin [32]). Let $S_n$ be the completion of $S$ with respect to $\|\cdot\|_n$. Let $S'$ and $S'_n$ be the dual space of $S$ and $S_n$, respectively. Then $S' = \bigcup_{n \in \mathbb{N}_0} S'_n$. Denote by $\|\cdot\|_{-n}$ the dual norm on $S_{-n} = S'_n$, with corresponding inner product $\langle \cdot, \cdot \rangle_{-n}$. The collection $\{S_n\}_{n \in \mathbb{Z}}$ defines a sequence of nested Hilbert spaces with $S_w \subset S_v$ if $w \geq v$. The main result of this section shows that for a suitable $\rho \in \mathbb{N}$, $\{Y^m\}$ satisfies a LDP in $C([0,T] : S_{-\rho})$ with speed $a^2(m)$ as introduced in (6.1),
namely, for all $F \in \mathcal{C}_b(\mathcal{C}([0,T] : \mathcal{S}_{-\rho}))$

$$\lim_{m \to \infty} -a^2(m) \log \mathbb{E}\exp \left\{ -\frac{1}{a^2(m)} F(Y^m) \right\} = \inf_{\zeta \in \mathcal{C}([0,T] : \mathcal{S}_{-\rho})} \{ I(\zeta) + F(\zeta) \}$$

(6.8)

for a suitable rate function $I$. The form of the rate function will be identified in (6.16).

We make the following assumption on the coefficients $\alpha$ and $\beta$.

**Condition 6.1.** $\alpha,\beta \in \mathcal{C}_b^2(\mathbb{R}^2)$.

It is easy to show that under Condition 6.1 there is a unique pathwise solution to (6.2). In fact under this condition one also has unique solvability of certain controlled analogues of (6.2) that will be used in our proofs. We now introduce these controlled processes.

Let for each $m \in \mathbb{N}$, $\{u_i^m : i = 1, \ldots, m\}$ be a collection of real-valued $\{\mathcal{F}_t\}$-progressively measurable processes such that $\mathbb{E} \sum_{i=1}^m \int_0^T |u_i^m(s)|^2 ds < \infty$. We will refer to $\{u_i^m\}$ as control processes. Define for $t \in [0,T]$,

$$\tilde{\mu}^m(t) = \frac{1}{m} \sum_{i=1}^m \delta_{\tilde{X}_i^m(t)},$$

(6.9)

where

$$\tilde{X}_i^m(t) = x_0 + \int_0^t \sigma(\tilde{X}_i^m(s),\tilde{\mu}^m(s)) \, dW_i(s) + \int_0^t b(\tilde{X}_i^m(s),\tilde{\mu}^m(s)) \, ds$$

$$+ \int_0^t \sigma(\tilde{X}_i^m(s),\tilde{\mu}^m(s))u_i^m(s) \, ds, \quad i = 1, \ldots, m.$$  

(6.10)

It is easy to check that under Condition 6.1 there is a unique pathwise solution to the system of equations in (6.10). Define for $t \in [0,T]$,

$$\tilde{Y}^m(t) = a(m) \sqrt{m}(\tilde{\mu}^m(t) - \mu(t)).$$

(6.11)

In Section 6.2 (see Theorem 6.11), we will show that under Condition 6.1, for every control sequence $\{u_i^m ; i = 1, \ldots, m\}_{m \in \mathbb{N}}$ such that

$$\sup_{m \in \mathbb{N}} a^2(m) \mathbb{E} \sum_{i=1}^m \int_0^T |u_i^m(s)|^2 \, ds < \infty,$$

94
\{\tilde{\gamma}^m\}_{m \in \mathbb{N}} is tight in \mathcal{C}([0, T]: \mathcal{S}_{-v}) for some \ v > 4. Specifically, one can take any \ v > 4 for which there is an \ r \in \mathbb{N}, 4 < r < v, such that \sum_{j=1}^{\infty} \|\phi^r_j\|^2 < \infty and \sum_{j=1}^{\infty} \|\phi^r_j\|^2 < \infty, where for \ n \in \mathbb{Z}, \ \{\phi^r_j\} is a complete orthonormal system of \mathcal{S}_n (see proof of Theorem 6.11, above (6.37)). We remark that the convergence of the above two series is equivalent to the property that the embedding maps \mathcal{S}_{-4} \to \mathcal{S}_{-r} and \mathcal{S}_{-r} \to \mathcal{S}_{-v} are Hilbert-Schmidt.

It will be convenient to introduce another system of seminorms \|\cdot\|_n on \mathcal{S} as

\[|\phi|_n \doteq \sum_{0 \leq k \leq n} \sup_{x \in \mathbb{R}} |\phi^{(k)}(x)|.\]

It is easy to check that, for each \ n \in \mathbb{N}_0, there is a \ \gamma_0(n) \in (0, \infty) such that

\[|\phi|_n \leq \gamma_0(n)\|\phi\|_{n+1}.\] (6.12)

We make the following additional assumption on the coefficients \(\alpha\) and \(\beta\).

**Condition 6.2.** \(\alpha, \beta\) are \(w\)-times continuously differentiable, where \(w = v + 2\), and

(a) \(\sup_y |\alpha(\cdot, y)|_w < \infty\) and \(\sup_y |\beta(\cdot, y)|_w < \infty\).

(b) \(\sup_x \|\alpha(x, \cdot)\|_w < \infty\) and \(\sup_x \|\beta(x, \cdot)\|_w < \infty\).

**Remark 6.1.** Conditions 6.1 and 6.2 are satisfied for \(w\)-times continuously differentiable functions \(\alpha\) and \(\beta\) if the functions along with their derivatives decay rapidly at \(\infty\).

We can now state our main MDP result of this section. We begin by introducing the associated rate function.

Let \(\mathcal{M}_T(\mathbb{R}^d \times [0, T])\) be the space of all measures \(\nu\) on \((\mathbb{R}^d \times [0, T], \mathcal{B}(\mathbb{R}^d \times [0, T]))\) such that \(\nu(\mathbb{R}^d \times [0, t]) = t\) for all \(t \in [0, T]\), equipped with the usual weak convergence topology. With \(\mu_s \equiv \mu(s)\) as in (6.3), define \(\bar{\nu} \in \mathcal{M}_T(\mathbb{R} \times [0, T])\) as

\[\bar{\nu}(A \times [0, t]) \doteq \int_0^t \mu_s(A) \, ds, \quad A \in \mathcal{B}(\mathbb{R}), \quad t \in [0, T].\] (6.13)

For a measure \(\theta \in \mathcal{M}_T(\mathbb{R}^d \times [0, T])\), denote by \(\theta(i)\) [resp. \(\theta(i,j)\)] the \(i\)-th [resp. \((i, j)\)-th joint] marginal. Let

\[\mathcal{P}_\infty \doteq \left\{ \nu \in \mathcal{M}_T(\mathbb{R}^2 \times [0, T]) \mid \nu_{(2,3)} = \bar{\nu}, \int_{\mathbb{R}^2 \times [0, T]} y^2 \nu(dy \, dx \, ds) < \infty \right\}.\] (6.14)
The space $P_\infty$ will be used to formulate the rate function and will play a key role in our weak convergence analysis. Roughly speaking, for a $\nu \in P_\infty$, the first marginal $\nu(1)$ corresponds to the “control variable”, $\nu(2)$ to the “state variable” and $\nu(3)$ to the “time variable” (see (6.15)).

Given $\eta \in C([0,T] : S_{-\nu})$, let $\mathcal{T}(\eta)$ be the collection of all $\nu \in P_\infty$ such that, for all $\phi \in S$ and $t \in [0,T]$

$$
\langle \eta(t), \phi \rangle = \int_0^t \langle \eta(s), L(s)\phi \rangle ds \left + \int_{R^2 \times [0,t]} \phi'(x)\sigma(x,\mu(s))y \nu(dydxds). \quad (6.15)
$$

Note that since $L(s)$ maps $S_w$ to $S_\nu$ (see Lemma 6.12) and $S \subset S_n$ for all $n \in N$, $\langle \eta(s), L(s)\phi \rangle$ is well defined for all $s \in [0,T]$. In Section 6.2.4 (see Lemma 6.14), we will show that under Conditions 6.1 and 6.2, for every $\nu \in P_\infty$, there exists a unique $\eta \in C([0,T] : S_{-\nu})$ that solves (6.15). Define $I: C([0,T] : S_{-\nu}) \to [0,\infty]$ as

$$
I(\eta) \doteq \inf_{\nu \in \mathcal{T}(\eta)} \left \{ \frac{1}{2} \int_{R^2 \times [0,T]} y^2 \nu(dydxds) \right \}, \quad (6.16)
$$

where the infimum over an empty set is taken to be $\infty$. In Sections 6.2.3 and 6.2.4 we will see that under Conditions 6.1 and 6.2 for every $\tau \geq \nu$ Laplace upper and lower bounds (see (6.17) and (6.18)) hold for every $F \in C_b(C([0,T] : S_{-\nu}))$ with $I$ defined as above. Although the Laplace upper and lower bound hold in particular with $\tau = \nu$, the function $I$ need not have relatively compact level sets in $C([0,T] : S_{-\nu})$ (see comments in Section 6.2.5 below (6.51)) and one needs to strengthen Condition 6.2 and enlarge the space in order to obtain the compactness property of $I$. Specifically, we take $\rho > w$ such that $\sum_{j=1}^{\infty} \| \phi_j^\rho \|_w^2 < \infty$ and we strengthen Condition 6.2 as follows.

**Condition 6.3.** $\alpha, \beta$ are $(\rho + 2)$-times continuously differentiable and,

(a) $\sup_y |\alpha(\cdot, y)|_{\rho+2} < \infty$ and $\sup_y |\beta(\cdot, y)|_{\rho+2} < \infty$.

(b) $\sup_x |\alpha(x, \cdot)|_{\rho+2} < \infty$ and $\sup_x |\beta(x, \cdot)|_{\rho+2} < \infty$.

Under Conditions 6.1 and 6.3 we will establish a LDP for $Y^m$ in $C([0,T] : S_{-\rho})$ with rate function $I$. We thus regard $I$ as a function from $C([0,T] : S_{-\rho})$ to $[0,\infty]$, with the convention that $I(\eta) \doteq \infty$ for all $\eta \in C([0,T] : S_{-\rho}) \setminus C([0,T] : S_{-\nu})$.

The following is the main result of this section. The proof will be given in Section 6.2.

**Theorem 6.4.** Under Conditions 6.1 and 6.3, $\{Y^m\}$ satisfies a LDP in $C([0,T] : S_{-\rho})$ with speed $a^2(m)$ and rate function $I$, where $\rho \in N$ is as introduced above.
Outline of the proof: The proof of Theorem 6.4 will be completed in three steps.

- Laplace principle upper bound: In Section 6.2.3 we show that under Conditions 6.1 and 6.2, for all $\tau \geq v$ and $F \in \mathbb{C}_b(\mathbb{C}([0,T] : S_{-\tau}))$,
  \[ \liminf_{m \to \infty} -a^2(m) \log \mathbb{E} \exp \left\{ -\frac{1}{a^2(m)} F(Y^m) \right\} \geq \inf_{\zeta \in \mathbb{C}([0,T] : S_{-v})} \{ I(\zeta) + F(\zeta) \}. \]  
  (6.17)

- Laplace principle lower bound: In Section 6.2.4 we show that under Conditions 6.1 and 6.2, for all $\tau \geq v$ and $F \in \mathbb{C}_b(\mathbb{C}([0,T] : S_{-\tau}))$,
  \[ \limsup_{m \to \infty} -a^2(m) \log \mathbb{E} \exp \left\{ -\frac{1}{a^2(m)} F(Y^m) \right\} \leq \inf_{\zeta \in \mathbb{C}([0,T] : S_{-v})} \{ I(\zeta) + F(\zeta) \}. \]  
  (6.18)

- $I$ is a rate function on $\mathbb{C}([0,T] : S_{-\rho})$: In Section 6.2.5 we show that under Conditions 6.1 and 6.3, for each $K < \infty$, \{ $\eta \in \mathbb{C}([0,T] : S_{-\rho}) : I(\eta) \leq K$ \} is a compact subset of $\mathbb{C}([0,T] : S_{-\rho})$.

Note that since $I(\eta) = \infty$ for $\eta \notin \mathbb{C}([0,T] : S_{-v})$, we can replace $v$ by any $\tau \geq v$ on the right sides of (6.17) and (6.18). Theorem 6.4 follows on combining these results.

Remark 6.2. The rate function $I$ has the following alternative representation. Given $\eta \in \mathbb{C}([0,T] : S_{-v})$, let $T^*(\eta)$ be the collection of $g \in L^2(\tilde{\nu})$ such that for all $\phi \in \mathcal{S}$ and $t \in [0,T]$,
  \[ \langle \eta(t), \phi \rangle = \int_0^t \langle \eta(s), L(s)\phi \rangle \, ds + \int_{\mathbb{R} \times [0,t]} \phi'(x)\sigma(x,\mu(s))g(x,s) \mu_s(dx) \, ds. \]  
  (6.19)

As for (6.15), under Conditions 6.1 and 6.2, for every $g \in L^2(\tilde{\nu})$, there is a unique $\eta \in \mathbb{C}([0,T] : S_{-v})$ that solves (6.19). We take $T^*(\eta)$ to be the empty set if $\eta \notin \mathbb{C}([0,T] : S_{-\rho}) \setminus \mathbb{C}([0,T] : S_{-\rho})$. Define $I^* : \mathbb{C}([0,T] : S_{-\rho}) \to [0, \infty]$ as
  \[ I^*(\eta) \doteq \inf_{g \in T^*(\eta)} \left\{ \frac{1}{2} \int_{\mathbb{R} \times [0,T]} g^2(x,s) \mu_s(dx) \, ds \right\}. \]

It is easy to check that $I^* = I$. Indeed every $g \in T^*(\eta)$ corresponds to a $\nu_g \in \mathcal{P}_\infty$ given as
  \[ \nu_g(dy \, dx \, ds) \doteq \delta_{g(x,s)}(dy) \tilde{\nu}(dx) \, ds \]
and every $\nu \in \mathcal{P}_\infty$ corresponds to a $g_\nu \in L^2(\bar{\nu})$ given as

$$g_\nu(x, s) = \int_{\mathbb{R}} y \vartheta(x, s, dy),$$

where $\vartheta(x, s, dy)$ is obtained by disintegrating $\nu$ as $\nu(dy dx ds) = \vartheta(x, s, dy) \bar{\nu}(dx ds)$.

### 6.2 Proof of Theorem 6.4

The proof of the MDP in Theorem 6.4 will proceed by first establishing the Laplace upper bound (6.17) and then the Laplace lower bound (6.18). The variational representation given in Theorem 2.7 will be a key ingredient in both proofs. Rest of this section is organized as follows. We first analyze certain controlled process in Sections 6.2.1 and 6.2.2. Proof of the Laplace upper bound is given in Section 6.2.3 whereas the lower bound is established in Section 6.2.4. In order to argue that $\{Y^m\}$ satisfies a LDP it then remains to establish that $I$ defined in (6.16) is a rate function. This is proved in Section 6.2.5.

#### 6.2.1 Controlled processes

Throughout this section we assume Condition 6.1. Let $v, w, \rho$ be as introduced in Section 6.1. Fix $\tau \geq v$ and let $F \in \mathcal{C}_b(\mathcal{C}([0, T] : \mathcal{S}_\tau))$. Using the variational representation in Theorem 2.7 on the filtered probability space introduced below (6.2), we have

$$-a^2(m) \log E \exp \left\{-\frac{1}{a^2(m)} F(Y^m)\right\} = \inf_{u^m = \{u^m_i\}_{i=1}^m} E \left\{\frac{1}{2} a^2(m) \sum_{i=1}^m \int_0^T |u^m_i(s)|^2 ds + F(\tilde{Y}^m)\right\},$$

where the infimum is taken over all $\{F_t\}$-progressively measurable $u^m$ such that $\tilde{Y}^m$ is as in (6.11) with $X^m$ given by (6.10) and $E \sum_{i=1}^m \int_0^T |u^m_i(s)|^2 ds < \infty$. We will view $\{u^m_i\}_{i=1}^m$ as a sequence of controls and $\{X^m_i\}_{i=1}^m$ as the controlled analogue of the original interacting particle system (6.2). Letting $\tilde{u}^m_i = a(m) \sqrt{mu^m_i}$, one can write

$$-a^2(m) \log E \exp \left\{-\frac{1}{a^2(m)} F(Y^m)\right\} = \inf_{\tilde{u}^m = \{\tilde{u}^m_i\}_{i=1}^m} E \left\{\frac{1}{2} \frac{1}{m} \sum_{i=1}^m \int_0^T |\tilde{u}^m_i(s)|^2 ds + F(\tilde{Y}^m)\right\},$$

(6.20)
Lemma 6.5. Suppose the control sequence \( \{\tilde{u}_i^m\}_{i=1}^m \) satisfies

\[
\sup_{m \in \mathbb{N}} \mathbf{E} \left( \frac{1}{m} \sum_{i=1}^m \int_0^T |\tilde{u}_i^m(s)|^2 \, ds \right) < \infty. \tag{6.22}
\]

Then there exists \( \gamma_1 \in (0, \infty) \) such that for all \( m \in \mathbb{N} \),

\[
\mathbf{E} \left( \frac{1}{m} \sum_{i=1}^m |\tilde{X}_i^m - \bar{X}_i|^2_{s,T} \right) \leq \frac{\gamma_1}{a^2(m)m}. \tag{6.23}
\]

In particular,

\[
\sup_{m \in \mathbb{N}} \mathbf{E} \left( \frac{1}{m} \sum_{i=1}^m \tilde{X}_i^m \right)^2_{s,T} < \infty. \tag{6.24}
\]

Proof. Using the bounded Lipschitz property of the coefficients \( \alpha \) and \( \beta \),

\[
\mathbf{E} |\tilde{X}_i^m - \bar{X}_i|^2_{s,t} \leq \kappa_1 \int_0^t \mathbf{E} \left( |\sigma(\tilde{X}_i^m(s), \tilde{\mu}_m(s)) - \sigma(\bar{X}_i(s), \tilde{\mu}(s))|^2 + |\sigma(\tilde{X}_i(s), \tilde{\mu}_m(s)) - \sigma(\bar{X}_i(s), \mu(s))|^2 \
+ |b(\tilde{X}_i^m(s), \tilde{\mu}_m(s)) - b(\bar{X}_i(s), \tilde{\mu}(s))|^2 + |b(\tilde{X}_i(s), \tilde{\mu}_m(s)) - b(\bar{X}_i(s), \mu(s))|^2 \
+ \frac{1}{a^2(m)m} |\sigma(\tilde{X}_i^m(s), \tilde{\mu}_m(s))\tilde{u}_i^m(s)|^2 \right) \, ds
\leq \kappa_2 \int_0^t \mathbf{E} \left( |\tilde{X}_i^m(s) - \bar{X}_i(s)|^2 + \frac{1}{m} \sum_{j=1}^m |\tilde{X}_j^m(s) - \bar{X}_j(s)|^2 + \frac{1}{a^2(m)m} |\tilde{u}_i^m(s)|^2 + \frac{1}{m} \right) \, ds,
\]

where the contribution of \( \frac{1}{m} \) is obtained from the second and the fourth terms on the right side upon using the independence of \( \bar{X}_i \) and \( \tilde{X}_j \) for \( i \neq j \). Taking the average over \( i = 1, \ldots, m \) on both
sides of above inequality and using (6.22)

\[
\mathbb{E} \frac{1}{m} \sum_{i=1}^{m} |\tilde{X}_i^m - \bar{X}_i|^2 \leq \kappa_3 \int_0^t \mathbb{E} \frac{1}{m} \sum_{j=1}^{m} |\tilde{X}_j^m - \bar{X}_j|^2 ds + \frac{\kappa_3}{a^2(m)m} + \frac{\kappa_3}{m}.
\]

The estimate in (6.23) is now immediate by Gronwall’s lemma. Using (6.23) and the fact that

\[
\sup_{m \in \mathbb{N}} \mathbb{E} \frac{1}{m} \sum_{i=1}^{m} |\bar{X}_i|^2 |_{s,T} = \mathbb{E} |\bar{X}_1|^2 |_{s,T} < \infty,
\]

we have the estimate in (6.24).

Our main result of this section is the following representation for the controlled processes \(\tilde{Y}^m\). Recall the operator \(L(s)\) defined in (6.5).

**Proposition 6.6.** Suppose that the control sequence \(\{\tilde{u}_i^m\}_{i=1}^{m}\) satisfies (6.22). Then, for every \(t \in [0,T]\) and \(\phi \in \mathcal{S}\),

\[
\langle \tilde{Y}^m(t), \phi \rangle = \int_0^t \langle \tilde{Y}^m(s), L(s) \phi \rangle ds + \int_{\mathbb{R}^2 \times [0,t]} \phi'(x) \sigma(x, \mu(s)) y \tilde{v}^m(dy dx ds) + \mathcal{R}^m(t),
\]

where \(\mathbb{E} |\mathcal{R}^m| |_{s,T} \leq \gamma(m) \| \phi \|_4\) and \(\gamma(m) \to 0\) as \(m \to \infty\).

Rest of this section is devoted to the proof of Proposition 6.6 and so we will assume throughout the remaining section that (6.22) holds. Note that by an application of Ito’s formula, for \(\phi \in \mathcal{S}\),

\[
\phi(\bar{X}_i(t)) = \phi(x_0) + \int_0^t \phi'(\bar{X}_i^m(s)) \sigma(\bar{X}_i^m(s), \bar{\mu}^m(s)) dW_i(s) + \int_0^t \phi'(\bar{X}_i^m(s)) b(\bar{X}_i^m(s), \bar{\mu}^m(s)) ds + \frac{1}{a(m) \sqrt{m}} \int_0^t \phi'(\bar{X}_i^m(s)) \sigma(\bar{X}_i^m(s), \bar{\mu}^m(s)) \tilde{u}_i^m(s) ds + \frac{1}{2} \int_0^t \phi''(\bar{X}_i^m(s)) \sigma^2(\bar{X}_i^m(s), \bar{\mu}^m(s)) ds.
\]

Similarly applying Ito’s formula to \(\phi(\bar{X}_i(t))\) and taking expectations, we have

\[
\langle \mu(t), \phi \rangle = \phi(x_0) + \int_0^t \langle \mu(s), \phi'(\cdot) b(\cdot, \mu(s)) \rangle ds + \frac{1}{2} \int_0^t \langle \mu(s), \phi''(\cdot) \sigma^2(\cdot, \mu(s)) \rangle ds.
\]
Combining the above observations

\[
\langle \tilde{Y}^m(t), \phi \rangle = a(m) \sqrt{m} \left( \langle \tilde{\mu}^m(t), \phi \rangle - \langle \mu(t), \phi \rangle \right)
\]

\[
= \frac{a(m)}{\sqrt{m}} \sum_{i=1}^{m} \int_{0}^{t} \phi'(\tilde{X}_i^m(s)) \sigma(\tilde{X}_i^m(s), \tilde{\mu}^m(s)) dW_i(s)
\]

\[
+ \frac{a(m)}{\sqrt{m}} \sum_{i=1}^{m} \int_{0}^{t} \left( \phi'(\tilde{X}_i^m(s)) b(\tilde{X}_i^m(s), \tilde{\mu}^m(s)) - \langle \mu(s), \phi'(\cdot) b(\cdot, \mu(s)) \rangle \right) ds
\]

\[
+ \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{t} \phi'(\tilde{X}_i^m(s)) \sigma(\tilde{X}_i^m(s), \tilde{\mu}^m(s)) \tilde{a}_i^m(s) ds
\]

\[
+ \frac{1}{2} \frac{a(m)}{\sqrt{m}} \sum_{i=1}^{m} \int_{0}^{t} \left( \phi''(\tilde{X}_i^m(s)) \sigma^2(\tilde{X}_i^m(s), \tilde{\mu}^m(s)) - \langle \mu(s), \phi''(\cdot) \sigma^2(\cdot, \mu(s)) \rangle \right) ds
\]

\[
\approx \sum_{k=1}^{\frac{4}{T_k^m}} T_k^m(t).
\] (6.27)

We will now separately estimate each $T_k^m$, $k = 1, 2, 3, 4$. We begin with $T_1^m$.

**Lemma 6.7.** There exists $\gamma_2 \in (0, \infty)$ such that $\mathbb{E} |T_1^m|_{*, T} \leq \gamma_2 a(m) \|\phi\|_2$.

**Proof.** Using Doob’s maximal inequality, the boundedness of $\alpha$ and (6.12), we have $\mathbb{E} |T_1^m|_{*, T} \leq \kappa_1 a^2(m) |\phi|_1^2 \leq \kappa_2 a^2(m) \|\phi\|_2^2$. The result follows. \(\square\)

Next we estimate the term $T_2^m$. Note that for $t \in [0, T]$

\[
T_2^m(t) = \frac{a(m)}{\sqrt{m}} \sum_{i=1}^{m} \int_{0}^{t} \left( \phi'(\tilde{X}_i^m(s)) b(\tilde{X}_i^m(s), \tilde{\mu}^m(s)) - \langle \mu(s), \phi'(\cdot) b(\cdot, \mu(s)) \rangle \right) ds
\]

\[
= \int_{0}^{t} \langle \tilde{Y}^m(s), \phi'(\cdot) b(\cdot, \mu(s)) \rangle ds
\]

\[
+ \frac{a(m)}{\sqrt{m}} \sum_{i=1}^{m} \int_{0}^{t} \left( \phi'(\tilde{X}_i^m(s)) b(\tilde{X}_i^m(s), \tilde{\mu}^m(s)) - \phi'(\tilde{X}_i^m(s)) b(\tilde{X}_i^m(s), \mu(s)) \right) ds
\]

\[
= \int_{0}^{t} \langle \tilde{Y}^m(s), \phi'(\cdot) b(\cdot, \mu(s)) \rangle ds + \int_{0}^{t} \langle \tilde{Y}^m(s), \int_{\mathbb{R}} \phi'(x) \beta(x, \cdot) \mu_s(dx) \rangle ds + R_2^m(t),
\] (6.28)

where $R_2^m(t) = \int_{0}^{t} \tilde{R}_{21}^m(s) ds$ and

\[
\tilde{R}_{21}^m(s) = \frac{a(m)}{\sqrt{m}} \sum_{i=1}^{m} \phi'(\tilde{X}_i^m(s))[b(\tilde{X}_i^m(s), \tilde{\mu}^m(s)) - b(\tilde{X}_i^m(s), \mu(s))] \]

\[
- a(m) \sqrt{m} \int_{\mathbb{R}} \phi'(x)[b(x, \tilde{\mu}^m(s)) - b(x, \mu(s))] \mu_s(dx).
\]

101
In the following lemma we estimate the remainder term $\mathcal{R}_2^m$.

**Lemma 6.8.** There exists $\gamma_3 \in (0, \infty)$ such that $\mathbb{E}|\mathcal{R}_2^m|_{s,T} \leq \frac{\gamma_3||\phi||_2}{a(m)\sqrt{m}}$.

**Proof.** Define $\tilde{\mathcal{R}}_{21}^m(s)$ by replacing in the definition of $\tilde{\mathcal{R}}_{21}^m(s)$ the term $\tilde{\mu}^m(s)$ with $\tilde{\mu}^m(s)$, namely

\[
\tilde{\mathcal{R}}_{22}^m(s) \doteq \frac{a(m)\sqrt{m}}{m^2} \sum_{i=1}^{m} \sum_{j=1}^{m} \phi'(\tilde{X}_i^m(s))\left[b(\tilde{X}_i^m(s), \tilde{\mu}^m(s)) - b(\tilde{X}_i^m(s), \mu(s))\right]
\]

Using the representation of $b$ in terms of $\beta$ and suitably adding and subtracting terms we see that

\[
\tilde{\mathcal{R}}_{21}^m(s) - \tilde{\mathcal{R}}_{22}^m(s) = \frac{a(m)\sqrt{m}}{m^2} \sum_{i=1}^{m} \sum_{j=1}^{m} \left\{ \phi'(\tilde{X}_i^m(s)) - \phi'(\tilde{X}_i(s)) \right\} \left[\beta(\tilde{X}_i^m(s), \tilde{X}_j^m(s)) - \beta(\tilde{X}_i^m(s), \tilde{X}_j(s))\right]
\]

\[
- \phi'(\tilde{X}_i(s))\beta(\tilde{X}_i^m(s), \tilde{X}_j(s)) - \phi'(\tilde{X}_i(s))\beta(\tilde{X}_i^m(s), \tilde{X}_j(s))
\]

We will now compute the square of the first absolute moments of the various terms on the right side, which are denoted by $A_k^m(s)$, $k = 1, \ldots, 5$, for short. Using Cauchy–Schwarz inequality and Lipschitz estimates on $\beta$ and $\phi'$, we have

\[
A_1^m(s) \leq \kappa_1 a^2(m)m||\phi||_2^2 \left( \mathbb{E} \frac{1}{m} \sum_{i=1}^{m} |\tilde{X}_i^m(s) - \tilde{X}_i(s)|^2 \right) \left( \mathbb{E} \frac{1}{m} \sum_{j=1}^{m} |\tilde{X}_j^m(s) - \tilde{X}_j(s)|^2 \right).
\]

Using Taylor’s formula and the fact that $\beta \in C^2_b(\mathbb{R}^2)$, we have

\[
A_2^m(s) \leq \kappa_1 a^2(m)m||\phi||_2^2 \left( \mathbb{E} \frac{1}{m} \sum_{j=1}^{m} |\tilde{X}_j^m(s) - \tilde{X}_j(s)|^2 \right)^2.
\]

For $A_3^m(s)$, we again use Cauchy–Schwarz inequality yielding the same bound as in (6.31). $A_4^m(s)$ can be bounded using Taylor’s formula as for $A_2^m(s)$ by the same expression as in (6.32). Finally
using Cauchy–Schwarz inequality and the independence of the sequence \( \{\bar{X}_i\} \) one has

\[
A_5^m(s) \leq \kappa_1 a^2(m)m \left( \mathbf{E} \frac{1}{m} \sum_{j=1}^{m} |\bar{X}_j^m(s) - \bar{X}_j(s)|^2 \right) \frac{|\varphi|^2}{m}.
\]

Combing these estimates with Lemma 6.5 we now have

\[
(\mathbf{E} |\bar{\mathcal{R}}_{21}^m(s) - \bar{\mathcal{R}}_{22}^m(s)|)^2 \leq 5 \sum_{k=1}^{5} A_k^m(s) \leq \kappa_2 |\varphi|^2 \left\{ \frac{1}{a^2(m)m} + \frac{1}{m} \right\}.
\] (6.33)

The above estimate allows approximation of \( \bar{\mathcal{R}}_{21}^m(s) \) by \( \bar{\mathcal{R}}_{22}^m(s) \). Next we will approximate \( \bar{\mathcal{R}}_{22}^m(s) \) by the term \( \bar{\mathcal{R}}_{23}^m(s) \) that is obtained by replacing \( \bar{X}_i^m \) in (6.29) with \( \bar{X}_i \), namely

\[
\bar{\mathcal{R}}_{23}^m(s) = \frac{a(m)}{\sqrt{m}} \sum_{i=1}^{m} \phi'(\bar{X}_i(s))[b(\bar{X}_i(s), \bar{\mu}^m(s)) - b(\bar{X}_i(s), \mu(s))] - a(m) \sqrt{m} \int_{\mathbb{R}} \phi'(x)[b(x, \bar{\mu}^m(s)) - b(x, \mu(s))] \mu_s(dx).
\]

By similar addition and subtraction of terms as for (6.30),

\[
\bar{\mathcal{R}}_{22}^m(s) - \bar{\mathcal{R}}_{23}^m(s) = \frac{a(m)}{m^2} \sum_{i=1}^{m} m \left( \phi'(\bar{X}_i^m(s)) \right) \left\{ \beta(\bar{X}_i^m(s), \bar{X}_j(s)) - \beta(\bar{X}_i(s), \bar{X}_j(s)) \right\}
- (\bar{X}_i^m(s) - \bar{X}_i(s)) \beta_x(\bar{X}_i(s), \bar{X}_j(s))
- \int_{\mathbb{R}} [\beta(\bar{X}_i^m(s), y) - \beta(\bar{X}_i(s), y) - (\bar{X}_i^m(s) - \bar{X}_i(s)) \beta_x(\bar{X}_i(s), y)] \mu_s(dy)
+ [\bar{X}_i^m(s) - \bar{X}_i(s)][\beta_x(\bar{X}_i(s), \bar{X}_j(s)) - \int_{\mathbb{R}} \beta_x(\bar{X}_i(s), y) \mu_s(dy)]
+ [\phi'(\bar{X}_i^m(s)) - \phi'(\bar{X}_i(s))][\beta(\bar{X}_i(s), \bar{X}_j(s)) - \int_{\mathbb{R}} \beta(\bar{X}_i(s), y) \mu_s(dy)]
\]

As for the proof of (6.33), we have, once more using Cauchy–Schwarz inequality, Taylor series expansion, Lemma 6.5 and the independence of \( \{\bar{X}_i\} \),

\[
(\mathbf{E} |\bar{\mathcal{R}}_{22}^m(s) - \bar{\mathcal{R}}_{23}^m(s)|)^2 \leq \kappa_3 |\varphi|^2 \left( \frac{1}{a^2(m)m} + \frac{1}{m} \right).
\]
We omit the details. Finally noting that $b(x, \theta) \doteq \int_R \beta(x, y) \theta(dy)$ and $\bar{\mu}^m(t) = \frac{1}{m} \sum_{i=1}^m \delta_{\bar{X}_i(t)}$, we can write

\[ \hat{R}_{23}^m(s) = \frac{a(m) \sqrt{m}}{m^2} \sum_{i=1}^m \sum_{j=1}^m \left( \phi'(\bar{X}_i(s)) [\beta(\bar{X}_i(s), \bar{X}_j(s)) - b(\bar{X}_i(s), \mu(s))] - \int_R \phi'(x) [\beta(x, \bar{X}_j(s)) - b(x, \mu(s))] \mu_s(dx) \right), \]

Using independence of $\{\bar{X}_i\}$, we have

\[ E|\hat{R}_{23}^m(s)| \leq \kappa_4 a(m) |\phi|_1. \]

Combining the above estimates and using (6.1), (6.12) gives

\[ E|R_{2}^m|_{*, T} \leq E \int_0^T |\hat{R}_{21}^m(s)| ds \leq \frac{\kappa_5 |\phi|_2}{a(m) \sqrt{m}} \leq \frac{\kappa_6 \|\phi\|_3}{a(m) \sqrt{m}}. \]

This completes the proof of the lemma.

We will now estimate the term $T_3^m$. Define $\tilde{\nu}^m \in \mathcal{M}_T(\mathbb{R}^2 \times [0, T])$ as follows: For $A, B \in \mathcal{B}(\mathbb{R})$ and $t \in [0, T]$,

\[ \tilde{\nu}^m(A \times B \times [0, t]) \doteq \frac{1}{m} \sum_{i=1}^m \int_0^t \delta_{(\tilde{\nu}_i^m(s), \tilde{X}_i^m(s))}(A \times B) ds. \quad (6.34) \]

Then

\[ T_3^m(t) = \int_{\mathbb{R}^2 \times [0, t]} \phi'(x) \sigma(x, \mu(s)) y \tilde{\nu}^m(dy, dx, ds) + \mathcal{R}_3^m(t), \quad (6.35) \]

where

\[ \mathcal{R}_3^m(t) \doteq \frac{1}{m} \sum_{i=1}^m \int_0^t \phi'(\bar{X}_i^m(s)) [\sigma(\bar{X}_i^m(s), \bar{\mu}^m(s)) - \sigma(\bar{X}_i^m(s), \mu(s))] |\tilde{u}_i^m(s)| ds. \]

In the following lemma we estimate the remainder term $\mathcal{R}_3^m$.

**Lemma 6.9.** There exists $\gamma_4 \in (0, \infty)$ such that

\[ E|\mathcal{R}_3^m|_{*, T} \leq \gamma_4 \|\phi\|_2 (a^2(m)m)^{-\frac{1}{4}}. \]

104
Proof. Recall that we assume that (6.22) is satisfied. Using Cauchy–Schwarz inequality and boundedness of \(\alpha\), we have

\[
E|\mathcal{R}_{3i}^m|_{*T} \leq \left( E \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} \left[ \phi'(\bar{X}_i^m(s)) \right]^4 ds \right)^{\frac{1}{4}} \left( E \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} |\bar{\mu}_i^m(s)|^2 ds \right)^{\frac{1}{2}} \\
\cdot \left( E \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} \left[ \sigma(\bar{X}_i^m(s), \bar{\mu}_i^m(s)) - \sigma(\bar{X}_i^m(s), \mu(s)) \right]^4 ds \right)^{\frac{1}{4}} \\
\leq \kappa_1 |\phi|_1 \left( E \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} \left[ \sigma(\bar{X}_i^m(s), \bar{\mu}_i^m(s)) - \sigma(\bar{X}_i^m(s), \mu(s)) \right]^2 ds \right)^{\frac{1}{4}} \\
\leq \kappa_2 \|\phi\|_2 (\mathcal{R}_{31}^m + \mathcal{R}_{32}^m + \mathcal{R}_{33}^m)^{\frac{1}{4}},
\]

where

\[
\mathcal{R}_{31}^m \doteq E \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} \left[ \sigma(\bar{X}_i^m(s), \bar{\mu}_i^m(s)) - \sigma(\bar{X}_i(s), \bar{\mu}_i^m(s)) \right]^2 ds \\
\mathcal{R}_{32}^m \doteq E \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} \left[ \sigma(\bar{X}_i(s), \bar{\mu}_i^m(s)) - \sigma(\bar{X}_i(s), \mu(s)) \right]^2 ds \\
\mathcal{R}_{33}^m \doteq E \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} \left[ \sigma(\bar{X}_i(s), \mu(s)) - \sigma(\bar{X}_i^m(s), \mu(s)) \right]^2 ds.
\]

Using the Lipschitz property of \(\alpha\) and Lemma 6.5, for \(i = 1, 3\),

\[
\mathcal{R}_{3i}^m \leq \kappa_3 \left( E \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} |\bar{X}_i^m(s) - \bar{X}_i(s)|^2 ds \right) \leq \frac{\kappa_4}{a^2(m)m},
\]

Finally, by independence of \(\{X_i\}\), \(\mathcal{R}_{32}^m \leq \kappa_5 \frac{m}{m}\). Combining above estimates completes the proof. \(\square\)

Finally we consider \(\mathcal{T}_{4i}^m\).

\[
\mathcal{T}_{4i}^m(t) = \frac{1}{2} a(m) \sqrt{m} \sum_{i=1}^{m} \int_{0}^{t} \left( \phi''(\bar{X}_i^m(s)) \sigma^2(\bar{X}_i^m(s), \bar{\mu}_i^m(s)) - \langle \mu(s), \phi''(\cdot) \sigma^2(\cdot, \mu(s)) \rangle \right) ds \\
= \int_{0}^{t} \langle \bar{Y}_m(s), \frac{1}{2} \phi''(\cdot) \sigma^2(\cdot, \mu(s)) \rangle ds \\
+ \frac{1}{2} a(m) \sqrt{m} \sum_{i=1}^{m} \int_{0}^{t} \phi''(\bar{X}_i^m(s)) \left[ \sigma^2(\bar{X}_i^m(s), \bar{\mu}_i^m(s)) - \sigma^2(\bar{X}_i^m(s), \mu(s)) \right] ds.
\]
From this we can write
\[ T_{4}^{m}(t) = \int_{0}^{t} \left\langle \tilde{Y}^{m}(s), \frac{1}{2} \phi''(\cdot) \sigma^{2}(\cdot, \mu(s)) + \int_{\mathbb{R}} \phi''(x) \sigma(x, \mu(s)) \alpha(x, \cdot) \mu_{s}(dx) \right\rangle ds + \mathcal{R}_{4}^{m}(t), \tag{6.36} \]
where \( \mathcal{R}_{4}^{m}(t) = \int_{0}^{t} \tilde{R}_{41}^{m}(s) ds \) and
\[ \tilde{R}_{41}^{m}(s) = \frac{1}{2} a(m) \sum_{i=1}^{m} \phi''(\tilde{X}_{i}^{m}(s)) \left[ \sigma^{2}(\tilde{X}_{i}^{m}(s), \tilde{\mu}^{m}(s)) - \sigma^{2}(\tilde{X}_{i}^{m}(s), \mu(s)) \right] \]
\[ - a(m) \sqrt{m} \int_{\mathbb{R}} \phi''(x) \sigma(x, \mu(s)) [\sigma(x, \tilde{\mu}^{m}(s)) - \sigma(x, \mu(s))] \mu_{s}(dx). \]

The proof of the following lemma is similar to that of Lemma 6.8, the only difference being that one needs to estimate \( \phi'' \) rather than \( \phi' \). As a result the bound on the right side contains \( \|\phi\|_{4} \) instead of \( \|\phi\|_{3} \) as in Lemma 6.8. We omit the proof.

**Lemma 6.10.** There exists \( \gamma_{5} \in (0, \infty) \) such that
\[ E|\mathcal{R}_{4}^{m}|_{*,T} \leq \frac{\gamma_{5} \|\phi\|_{4}}{a(m) \sqrt{m}}. \]

We can now complete the proof of Proposition 6.6.

**Proof of Proposition 6.6:** Using (6.27), (6.28), (6.35) and (6.36), we have for \( t \in [0, T] \),
\[ \langle \tilde{Y}^{m}(t), \phi \rangle = \int_{0}^{t} \langle \tilde{Y}^{m}(s), L(s) \phi \rangle ds + \int_{\mathbb{R}^{2} \times [0, t]} \phi'(x) \sigma(x, \mu(s)) y \tilde{v}^{m}(dy dx ds) + \mathcal{R}^{m}(t), \]
where \( \mathcal{R}^{m}(t) = T_{1}^{m}(t) + \mathcal{R}_{2}^{m}(t) + \mathcal{R}_{3}^{m}(t) + \mathcal{R}_{4}^{m}(t) \).

The result now follows from Lemmas 6.7 – 6.10.

---

### 6.2.2 Tightness of \( \tilde{Y}^{m} \)

In this section we will argue the tightness of \( \tilde{Y}^{m} \) in \( \mathbb{C}([0, T] : \mathcal{S}_{v}) \) and identify the value of \( v \). Note that this tightness implies tightness in \( \mathbb{C}([0, T] : \mathcal{S}_{\tau}) \) for all \( \tau \geq v \).

**Theorem 6.11.** Suppose that Condition 6.1 holds and the control sequence \( \{\tilde{u}_{i}^{m}\} \) satisfies (6.22). Then the sequence \( \{(\tilde{Y}^{m}, \tilde{v}^{m})\} \) is tight in \( \mathbb{C}([0, T] : \mathcal{S}_{v}) \times \mathcal{M}_{T}(\mathbb{R}^{2} \times [0, T]) \) for some \( v > 4 \).
Proof. We first argue the tightness of $\tilde{Y}^m$. For this, we will make use of Proposition 6.6. Let for $t \in [0, T]$ and $\phi \in \mathcal{S}$,

$$A^m(t) = \int_0^t \left| \langle \tilde{Y}^m(s), L(s)\phi \rangle \right| ds$$

$$= \int_0^t a(m)\sqrt{m} \left| \langle \tilde{\mu}^m(s) - \bar{\mu}^m(s), L(s)\phi \rangle \right| ds, \quad B^m(t) = \int_{\mathbb{R}^2 \times [0, t]} \left| \phi'(x)\sigma(x, \mu(s))y \right| \tilde{v}^m(dy dx ds)$$

$$= \frac{1}{m} \sum_{i=1}^{m} \int_0^t \left| \phi'(\tilde{X}^m_i(s))\sigma(\tilde{X}^m_i(s), \mu(s))\tilde{u}^m_i(s) \right| ds.$$

Then for $t_1, t_2 \in [0, T]$,

$$|A^m(t_2) - A^m(t_1)|^2 \leq 2a^2(m) m |t_2 - t_1| \int_0^T \left( \langle \tilde{\mu}^m(s) - \bar{\mu}^m(s), L(s)\phi \rangle^2 + \langle \bar{\mu}^m(s) - \mu(s), L(s)\phi \rangle^2 \right) ds.$$

Note that

$$a^2(m)mE \int_0^T \left( \langle \tilde{\mu}^m(s) - \bar{\mu}^m(s), L(s)\phi \rangle^2 + \langle \bar{\mu}^m(s) - \mu(s), L(s)\phi \rangle^2 \right) ds$$

$$\leq \kappa_1 a^2(m) m \int_0^T \left( |L(s)\phi|^2 \mathbb{E} \frac{1}{m} \sum_{i=1}^{m} |\tilde{X}^m_i(s) - \bar{X}^m_i(s)|^2 + \frac{|L(s)\phi|^2}{m} \right) ds$$

$$\leq \kappa_2 |\phi|^2_3 (1 + a^2(m)),$$

where the last inequality uses Lemma 6.5 and the inequality $\sup_{0 \leq s \leq T} |L(s)\phi|_1 \leq \kappa_3 |\phi|_3$. This proves the tightness of $A^m$ in $\mathbb{C}([0, T] : \mathbb{R})$. Also for $t_1, t_2 \in [0, T]$,

$$|B^m(t_2) - B^m(t_1)|^2 \leq \kappa_4 |t_2 - t_1| |\phi|^2 \frac{1}{m} \sum_{i=1}^{m} \int_0^T |\tilde{u}^m_i(s)|^2 ds.$$

Combining this with (6.22) we now have the tightness of $B^m$ in $\mathbb{C}([0, T] : \mathbb{R})$. Also from Proposition 6.6 we have that $\mathcal{R}^m \Rightarrow 0$ in $\mathbb{C}([0, T] : \mathbb{R})$. The tightness of $t \mapsto \langle \tilde{Y}^m(t), \phi \rangle$ in $\mathbb{C}([0, T] : \mathbb{R})$, for each $\phi \in \mathcal{S}$, is now immediate. From the above estimates on $A^m$, $B^m$ and Proposition 6.6 we have that, for all $\phi \in \mathcal{S}$,

$$\sup_{m \in \mathbb{N}} \mathbb{E} \sup_{0 \leq t \leq T} \left| \langle \tilde{Y}^m(t), \phi \rangle \right| \leq \kappa_5 |\phi|_4.$$
This shows that for any $\epsilon_1 > 0$ and $\epsilon_2 > 0$, there is a $\delta > 0$ such that
\[
\mathbb{P} \left( \sup_{0 \leq t \leq T} \left| \langle \tilde{Y}^m(t), \phi \rangle \right| > \epsilon_1 \right) \leq \epsilon_2 \quad \text{if } \|\phi\|_4 \leq \delta, \quad m \in \mathbb{N}.
\]

Thus the induced measures $\mathbb{P} \circ (\tilde{Y}^m)^{-1}$ on $\mathcal{C}([0, T] : \mathcal{S}')$ are uniformly 4-continuous in the sense of [55]; see Remark (R.1) on page 997 there. It then follows from the same remark that the sequence $\{\tilde{Y}^m\}$ is tight in $\mathcal{C}([0, T] : \mathcal{S} - \tau)$ for some $v > 4$. Specifically, one can take any $v > 4$ for which there is an $r \in \mathbb{N}$, $4 < r < v$, such that $\sum_{j=1}^{\infty} \|\phi_j^v\|^2 < \infty$ and $\sum_{j=1}^{\infty} \|\phi_j^r\|_4^2 < \infty$.

Finally we argue tightness of $\tilde{\nu}^m$. Note that $g(\nu) = \int_{\mathbb{R}^2 \times [0, T]} (x^2 + y^2) \nu(dy dx ds)$, $\nu \in \mathcal{M}_T(\mathbb{R}^2 \times [0, T])$ (6.37)
is a tightness function on $\mathcal{M}_T(\mathbb{R}^2 \times [0, T])$, namely $g$ is bounded from below and has pre-compact level sets. Also, from (6.24) and the assumption that (6.22) holds,
\[
\sup_m \mathbb{E} |g(\tilde{\nu}^m)| = \sup_m \mathbb{E} \frac{1}{m} \sum_{i=1}^{m} \int_0^T \left( |\tilde{X}^m_i(s)|^2 + |\tilde{u}^m_i(s)|^2 \right) ds < \infty.
\]
This proves the tightness of $\{\tilde{\nu}^m\}$. \hfill \Box

### 6.2.3 Laplace upper bound

In this section we will establish under Conditions 6.1 and 6.2 the Laplace upper bound (6.17), where $I(\cdot)$ is as defined in (6.16), $F \in \mathcal{C}_b(\mathcal{C}([0, T] : \mathcal{S} - \tau))$, and $\tau \geq v$. Fix $\varepsilon \in (0, 1)$ and using (6.20) choose for each $m \in \mathbb{N}$ a sequence $\tilde{u}^m = \{\tilde{u}^m_i\}_{i=1}^{m}$ of controls such that
\[
-a^2(m) \log \mathbb{E} \exp \left\{ - \frac{1}{a^2(m)} F(Y^m) \right\} \geq \mathbb{E} \left\{ \frac{1}{2} m \sum_{i=1}^{m} \int_0^T |\tilde{u}^m_i(s)|^2 ds + F(\tilde{Y}^m) \right\} - \varepsilon, \quad (6.38)
\]
where $\tilde{Y}^m$ is as introduced in (6.11), with $\tilde{X}^m_i$ defined in (6.21) and above choice of $\tilde{u}^m$. Since the left side of (6.38) is bounded between $-\|F\|_\infty$ and $\|F\|_\infty$, we can assume $\tilde{u}^m$ are such that
\[
\mathbb{E} \frac{1}{m} \sum_{i=1}^{m} \int_0^T |\tilde{u}^m_i(s)|^2 ds \leq 4\|F\|_\infty + 2 \equiv C_F.
\]
Let $\tilde{\nu}^m$ be as introduced in (6.34). Then
\[
E \int_{\mathbb{R}^2 \times [0, T]} y^2 \tilde{\nu}^m(dy \, dx \, ds) = E \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} |\tilde{u}_i^m(s)|^2 ds \leq C_F. \tag{6.39}
\]

From Theorem 6.11 it follows that $\{(\tilde{Y}^m, \tilde{\nu}^m)\}$ is tight in $C([0, T] : \mathcal{S}_v) \times \mathcal{M}_T(\mathbb{R}^2 \times [0, T])$.

The following lemma will enable us to characterize its weak limit points. Recall that $w = v + 2$.

**Lemma 6.12.** Suppose Condition 6.2 holds. Then for each $n = 1, 2, \ldots, w$, there exists $c_n \in (0, \infty)$ such that for all $s \in [0, T]$ and $\phi \in \mathcal{S}_{n+2}$, $\|L(s)\phi\|_n \leq c_n \|\phi\|_{n+2}$. If Condition 6.3 holds then $w$ can be replaced by $\rho + 2$.

**Proof.** Note that from Condition 6.2(a), for $\phi \in \mathcal{S}$,
\[
\|\phi'(.)(\cdot, \mu(s))\|_n^2 = \sum_{0 \leq k \leq n} \int_{\mathbb{R}} (1 + x^2)^{2n} \left( \left[ \phi'(x) \sigma(x, \mu(s)) \right]^{(k)} \right)^2 dx \\
\leq \kappa_1 \sum_{0 \leq k \leq n} \int_{\mathbb{R}} (1 + x^2)^{2n+2} |\phi^{(k+1)}(x)|^2 dx \\
\leq \kappa_2 \|\phi\|^2_{n+1}.
\]

Similarly, $\|\frac{1}{2} \phi''(.)\sigma^2(\cdot, \mu(s))\|_n^2 \leq \kappa_3 \|\phi\|^2_{n+2}$. Also using Condition 6.2(b),
\[
\left\| \int_{\mathbb{R}} \phi'(y)\beta(y, \cdot) \mu_s(dy) \right\|_n^2 = \sum_{0 \leq k \leq n} \int_{\mathbb{R}} (1 + x^2)^{2n} \left( \left[ \int_{\mathbb{R}} \phi'(y)\beta(y, x) \mu_s(dy) \right]^{(k)} \right)^2 dx \\
\leq \int_{\mathbb{R}} \|\beta(y, \cdot)\|_n^2 \|\phi\|_{n+2}^2 \mu_s(dy) \\
\leq \kappa_4 \|\phi\|^2_2,
\]
and similarly $\|\int_{\mathbb{R}} \phi''(y)a(y, \mu(s))\alpha(y, \cdot) \mu_s(dy)\|_n^2 \leq \kappa_5 \|\phi\|^2_3$. The result follows on combining the above estimates. Proof of the second statement in the lemma is similar and hence omitted.

We can now establish the following characterization of the weak limit points of $\{(\tilde{Y}^m, \tilde{\nu}^m)\}$.

**Theorem 6.13.** Suppose that Condition 6.2 holds, the sequence of controls satisfies (6.22), and $\{(\tilde{Y}^m, \tilde{\nu}^m)\}$ converges weakly along a subsequence to $(\tilde{Y}, \tilde{\nu})$ in $C([0, T] : \mathcal{S}_v) \times \mathcal{M}_T(\mathbb{R}^2 \times [0, T])$.  

109
Then $\tilde{\nu} \in \mathcal{P}_\infty$ a.s. and $\tilde{Y}$ solves the following equation a.s.: For all $\phi \in \mathcal{S}$,

$$
\langle \tilde{Y}(t), \phi \rangle = \int_0^t \langle \tilde{Y}(s), L(s)\phi \rangle \, ds + \int_{\mathbb{R}^2 \times [0,t]} \phi'(x)\sigma(x, \mu(s))y\tilde{\nu}(dy \, dx \, ds).
$$

(6.40)

**Proof.** We assume without loss of generality that $(\tilde{Y}^m, \tilde{\nu}^m) \to (\tilde{Y}, \tilde{\nu})$ weakly along the full sequence.

We first verify $\tilde{\nu} \in \mathcal{P}_\infty$. Let $\tilde{\nu}^m \in \mathcal{M}_T(\mathbb{R} \times [0,T])$ be defined as

$$
\tilde{\nu}^m(B \times [0,t]) = \int_0^t \left( \frac{1}{m} \sum_{i=1}^m \delta_{\tilde{X}_i(s)}(B) \right) \, ds, \quad B \in \mathcal{B}(\mathbb{R}), \, t \in [0,T].
$$

It follows from (6.23) that

$$
\mathbb{E}d_{BL}(\tilde{\nu}^m_{(2,3)}, \tilde{\nu}^m) = \mathbb{E} \sup_{\|f\|_{BL} \leq 1} \left| \int_0^T \left( \frac{1}{m} \sum_{i=1}^m f(\tilde{X}_i(s), s) \right) \, ds - \int_0^T \left( \frac{1}{m} \sum_{i=1}^m f(\tilde{X}_i(s), s) \right) \, ds \right|^2
$$

$$
\leq \mathbb{E} \frac{T}{m} \sum_{i=1}^m \int_0^T |\tilde{X}_i^m(s) - \tilde{X}_i(s)|^2 \, ds
$$

$$
\leq \frac{\kappa_1}{a^2(m)m} \to 0.
$$

Also for each $f \in \mathcal{C}_b(\mathbb{R} \times [0,T])$, we have, with $\tilde{\nu}$ as in (6.13),

$$
\mathbb{E} \left| \langle \tilde{\nu}^m, f \rangle - \langle \tilde{\nu}, f \rangle \right|^2 = \mathbb{E} \left| \int_0^T \left( \frac{1}{m} \sum_{i=1}^m f(\tilde{X}_i(s), s) \right) \, ds - \int_0^T \langle f(\cdot, s), \mu(s) \rangle \, ds \right|^2
$$

$$
\leq \int_0^T \mathbb{E} \left( \frac{1}{m} \sum_{i=1}^m \left( f(\tilde{X}_i(s), s) - \langle f(\cdot, s), \mu(s) \rangle \right) \right)^2 \, ds \to 0.
$$

Combining the above two convergence properties with the fact that $\tilde{\nu}^m_{(2,3)} \to \tilde{\nu}_{(2,3)}$ weakly implies that $\tilde{\nu}_{(2,3)} = \tilde{\nu}$ a.s. Furthermore, it follows from Fatou’s lemma and (6.39) that

$$
\mathbb{E} \int_{\mathbb{R}^2 \times [0,T]} y^2 \tilde{\nu}(dy \, dx \, ds) \leq \liminf_{m \to \infty} \mathbb{E} \int_{\mathbb{R}^2 \times [0,T]} y^2 \tilde{\nu}^m(dy \, dx \, ds) \leq C_F.
$$

Thus we have shown that $\tilde{\nu} \in \mathcal{P}_\infty$ a.s.

Now we argue that $\tilde{Y}$ solves (6.40) a.s. Using Skorokhod’s representation theorem, we can assume that $(\tilde{Y}^m, \tilde{\nu}^m) \to (\tilde{Y}, \tilde{\nu})$ a.s. in $\mathcal{C}([0,T] : \mathcal{S}_\nu) \times \mathcal{M}_T(\mathbb{R}^2 \times [0,T])$. Then for each $\phi \in \mathcal{S}_\nu$,

$$
\langle \tilde{Y}^m(t), \phi \rangle \to \langle \tilde{Y}(t), \phi \rangle, \quad \forall t \in [0,T].
$$

(6.41)
It follows from Lemma 6.12 that for each \( \phi \in \mathcal{S}_w \),

\[
\sup_{m \in \mathbb{N}} \sup_{s \in [0, T]} |\langle \tilde{Y}^m(s), L(s)\phi \rangle| \leq \sup_{m \in \mathbb{N}} \sup_{s \in [0, T]} \|\tilde{Y}^m(s)\|_{-v} \|L(s)\phi\|_v \\
\leq \kappa_2 \sup_{m \in \mathbb{N}} \sup_{s \in [0, T]} \|\tilde{Y}^m(s)\|_{-v} \|\phi\|_v < \infty
\]

and hence by bounded convergence theorem, for every \( t \in [0, T] \),

\[
\int_0^t \langle \tilde{Y}^m(s), L(s)\phi \rangle \, ds \to \int_0^t \langle \tilde{Y}(s), L(s)\phi \rangle \, ds.
\]

(6.42)

In view of (6.41), (6.42) and Proposition 6.6, to finish the proof, it suffices to show that for each \( \phi \in \mathcal{S} \) and \( t \in [0, T] \),

\[
\int \int_{\mathbb{R}^2 \times [0, t]} \phi'(x)\sigma(x, \mu(s)) y \tilde{\nu}^m(dy \, dx \, ds) \to \int \int_{\mathbb{R}^2 \times [0, t]} \phi'(x)\sigma(x, \mu(s)) y \tilde{\nu}(dy \, dx \, ds)
\]

(6.43)

in probability. For this, first note that by convergence of \( \tilde{\nu}^m \) to \( \tilde{\nu} \), for each \( K \in (0, \infty) \),

\[
\int \int_{\mathbb{R}^2 \times [0, t]} \phi'(x)\sigma(x, \mu(s)) h_K(y) \tilde{\nu}^m(dy \, dx \, ds) \to \int \int_{\mathbb{R}^2 \times [0, t]} \phi'(x)\sigma(x, \mu(s)) h_K(y) \tilde{\nu}(dy \, dx \, ds)
\]

a.s. as \( m \to \infty \), where \( h_K(y) = y 1_{\{|y| \leq K\}} + K 1_{\{y > K\}} - K 1_{\{y < -K\}} \) for \( y \in \mathbb{R} \). Also it follows from (6.39) that

\[
\sup_m \mathbb{E} \left| \int \int_{\mathbb{R}^2 \times [0, t]} \phi'(x)\sigma(x, \mu(s))(y - h_K(y)) \tilde{\nu}^m(dy \, dx \, ds) \right| \\
\leq \|\phi\|_1 \|\alpha\|_\infty \sup_m \mathbb{E} \int \int_{\mathbb{R}^2 \times [0, T]} \frac{y^2}{K} \tilde{\nu}^m(dy \, dx \, ds) \\
\leq \|\phi\|_1 \|\alpha\|_\infty C_F \to 0
\]

as \( K \to \infty \). Similarly, using Fatou’s lemma,

\[
\mathbb{E} \left| \int \int_{\mathbb{R}^2 \times [0, t]} \phi'(x)\sigma(x, \mu(s))(y - h_K(y)) \tilde{\nu}(dy \, dx \, ds) \right| \to 0 \quad \text{as} \ K \to \infty.
\]

Combining the above convergence properties we have (6.43), which completes the proof. \( \square \)
We can now complete the proof of the Laplace upper bound under Conditions 6.1 and 6.2.

**Proof of the Laplace upper bound:** Recall that \( \{(\tilde{Y}_m, \tilde{\nu}_m)\} \) is tight in \( C([0, T] : S_{-v}) \times \mathcal{M}_T(\mathbb{R}^2 \times [0, T]) \). By a standard subsequential argument we can assume without loss of generality that \((\tilde{Y}_m, \tilde{\nu}_m)\) converges in distribution, along the full sequence, to a limit \((\tilde{Y}, \tilde{\nu})\) in \( C([0, T] : S_{-v}) \times \mathcal{M}_T(\mathbb{R}^2 \times [0, T]) \). It follows from Theorem 6.13 that \( \tilde{\nu} \in \mathcal{T}(\tilde{Y}) \) a.s. Also, from (6.38) we have that

\[
\liminf_{m \to \infty} -a^2(m) \log \mathbb{E} \exp \left\{ -\frac{1}{a^2(m)} F(Y^m) \right\} \geq \lim_{m \to \infty} \mathbb{E} \left[ \frac{1}{2} \int_{\mathbb{R}^2 \times [0, T]} y^2 \tilde{\nu}_m(dy \, dx \, ds) + F(\tilde{Y}^m) \right] - \varepsilon
\]

\[
\geq \mathbb{E} \left[ \frac{1}{2} \int_{\mathbb{R}^2 \times [0, T]} y^2 \tilde{\nu}(dy \, dx \, ds) + F(\tilde{Y}) \right] - \varepsilon
\]

\[
\geq \inf_{\zeta \in C([0, T] : S_{-v})} \left[ \inf_{\nu \in \mathcal{T}(\zeta)} \left\{ \frac{1}{2} \int_{\mathbb{R}^2 \times [0, T]} y^2 \nu(dy \, dx \, ds) \right\} + F(\zeta) \right] - \varepsilon
\]

\[
= \inf_{\zeta \in C([0, T] : S_{-v})} \{ I(\zeta) + F(\zeta) \} - \varepsilon,
\]

where the second inequality uses Fatou’s lemma and weak convergence of \((\tilde{Y}_m, \tilde{\nu}_m)\) to \((\tilde{Y}, \tilde{\nu})\) in \( C([0, T] : S_{-v}) \times \mathcal{M}_T(\mathbb{R}^2 \times [0, T]) \). Since \( \varepsilon > 0 \) is arbitrary, the desired Laplace upper bound follows.

**6.2.4 Laplace lower bound**

In this section we prove the inequality (6.18) under Conditions 6.1 and 6.2 for every \( \tau \geq v \) and \( F \in C_b(C([0, T] : S_{-\tau})) \). Fix \( \varepsilon \in (0, 1) \). Let \( \zeta^* \in C([0, T] : S_{-v}) \) be such that

\[
I(\zeta^*) + F(\zeta^*) \leq \inf_{\zeta \in C([0, T] : S_{-v})} \{ I(\zeta) + F(\zeta) \} + \varepsilon.
\]

Recalling the definition of \( I \) in (6.16), choose \( \nu^* \in \mathcal{T}(\zeta^*) \) such that

\[
\frac{1}{2} \int_{\mathbb{R}^2 \times [0, T]} y^2 \nu^*(dy \, dx \, ds) \leq I(\zeta^*) + \varepsilon.
\]

Recalling that \( \nu^*_{(2,3)} = \tilde{\nu} \), we can disintegrate \( \nu^* \) as

\[
\nu^*(A \times B \times [0, t]) = \int_{B \times [0, t]} \vartheta(s, x, A) \mu_s(dx) \, ds, \quad A, B \in \mathcal{B}(\mathbb{R}), \quad t \in [0, T],
\]
for some \( \vartheta : [0, T] \times \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0, 1] \) such that for each \( A \in \mathcal{B}(\mathbb{R}) \), \( \vartheta(\cdot, \cdot, A) \) is a measurable map and for each \( (s, x) \in [0, T] \times \mathbb{R} \), \( \vartheta(s, x, \cdot) \in \mathcal{P}(\mathbb{R}) \). Define

\[
\phi(s, x) = \int_{\mathbb{R}} y \vartheta(s, x, dy)
\]

for \( (s, x) \in [0, T] \times \mathbb{R} \). Note that this is finite \( (\mu_s(dx)ds \text{ a.e.}) \) since

\[
\int_{\mathbb{R} \times [0, T]} \left( \int_{\mathbb{R}} |y| \vartheta(s, x, dy) \right)^2 \mu_s(dx) ds \leq \int_{\mathbb{R}^2 \times [0, T]} y^2 \nu^*(dy \, dx \, ds) < \infty. \tag{6.44}
\]

Recall the sequence \( \{\bar{X}_i\} \) defined through (6.3) in terms of an i.i.d. sequence of real Brownian motions \( \{W_i\} \) on the filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})\). Using the same sequence of Brownian motions, let \( \{\tilde{X}_i^m\} \) be the solution of the system of SDE in (6.21), where

\[
\tilde{\mu}^m(t) = \frac{1}{m} \sum_{i=1}^m \delta_{\tilde{X}_i^m(t)}, \quad \tilde{u}_i^m(s) = u(s, \tilde{X}_i(s)).
\]

It follows from (6.44) that

\[
\mathbf{E} \frac{1}{m} \sum_{i=1}^m \int_0^T |\tilde{u}_i^m(s)|^2 \, ds = \int_{\mathbb{R} \times [0, T]} |u(s, x)|^2 \mu_s(dx) ds \leq \int_{\mathbb{R}^2 \times [0, T]} y^2 \nu^*(dy \, dx \, ds) < \infty. \tag{6.45}
\]

We note that the controls \( \tilde{u}^m \) are defined using the given processes \( \{\bar{X}_i\} \) and hence clearly \( \{\mathcal{F}_t\} \)-progressively measurable. In particular, \( \{\tilde{u}_i^m\}_{i=1}^m \) is a control sequence of the form on which infimum is taken in (6.20). Consequently,

\[
-a^2(m) \log \mathbf{E} \exp \left\{ -\frac{1}{a^2(m)} F(Y^m) \right\} \leq \mathbf{E} \left\{ \frac{1}{2} \frac{1}{m} \sum_{i=1}^m \int_0^T |\tilde{u}_i^m(s)|^2 \, ds + F(\bar{Y}^m) \right\}, \tag{6.46}
\]

where \( \bar{Y}^m \) is defined as in (6.11) with \( \tilde{X}_i^m \) given through (6.21).

We now claim that as \( m \to \infty \), for each \( \phi \in \mathcal{S} \) and \( t \in [0, T] \)

\[
\int_{\mathbb{R}^2 \times [0, t]} \phi'(x) \sigma(x, \mu(s)) y \tilde{\nu}^m(dy \, dx \, ds) \to \int_{\mathbb{R}^2 \times [0, t]} \phi'(x) \sigma(x, \mu(s)) y \nu^*(dy \, dx \, ds) \tag{6.47}
\]
in probability. To verify this convergence, let

\[ \tilde{A}_m(t) = \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{t} \phi'(\tilde{X}_i^m(s)) \sigma(\tilde{X}_i^m(s), \mu(s)) \tilde{u}_i^m(s) \, ds \]

\[ \tilde{B}_m(t) = \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{t} \phi'(\bar{X}_i(s)) \sigma(\bar{X}_i(s), \mu(s)) \bar{u}_i^m(s) \, ds \]

\[ \tilde{C}(t) = \int_{\mathbb{R} \times [0,t]} \phi'(x) \sigma(x, \mu(s)) u(s, x) \mu_s(dx) \, ds. \]

Using Cauchy–Schwarz inequality and the boundedness and Lipschitz property of \( \phi' \) and \( \sigma \), we have

\[ \mathbb{E}[|\tilde{A}_m - \tilde{B}_m|_{s,T}^2] \leq \kappa_1 \left( \mathbb{E} \frac{1}{m} \sum_{i=1}^{m} |\tilde{X}_i^m - \bar{X}_i|_{s,T}^2 \right) \left( \mathbb{E} \frac{1}{m} \sum_{i=1}^{m} \int_{0}^{T} |\bar{u}_i^m(s)|^2 \, ds \right) \]

\[ \leq \frac{\kappa_2}{a^2(m)m} \rightarrow 0, \]

where the second inequality follows from (6.45) and Lemma 6.5. Also, since \( \{\tilde{X}_i\} \) are i.i.d.,

\[ \mathbb{E}[|\tilde{B}_m - \tilde{C}|_{s,T}^2] \leq \frac{\kappa_3}{m} \int_{\mathbb{R} \times [0,T]} |u(s, x)|^2 \mu_s(dx) \, ds \rightarrow 0. \]

Thus \( \tilde{A}_m(t) \rightarrow \tilde{C}(t) \) in probability for each \( t \in [0, T] \), which proves the claim (6.47).

Next, from Theorem 6.11 and (6.45) we have that \( \{\tilde{Y}_m\} \) is tight in \( \mathbb{C}([0, T] : \mathcal{S}_v) \). Finally using Proposition 6.6, the convergence in (6.47), and an analogous application of bounded convergence theorem used below (6.41), we see that any limit point \( \tilde{Y} \) of \( \tilde{Y}_m \) solves (6.15) with \( \eta \) replaced by \( \tilde{Y} \) and \( \nu \) by \( \nu^* \). In particular, since \( \{\tilde{Y}_m\} \) is tight, (6.15) (with \( \nu \) replaced by \( \nu^* \)) admits at least one solution. The following lemma shows that the equation admits only one solution, in particular, since \( \nu^* \in \mathcal{T}(\zeta^*) \), any limit point \( \tilde{Y} \) satisfies \( \tilde{Y} = \zeta^* \) a.s.

**Lemma 6.14.** Suppose Conditions 6.1 and 6.2 hold. Then for each \( \nu \in \mathcal{P}_\infty \), there exists a unique solution of (6.15) in \( \mathbb{C}([0, T] : \mathcal{S}_v) \). If in addition Condition 6.3 is satisfied, uniqueness holds in \( \mathbb{C}([0, T] : \mathcal{S}_\rho) \).

**Proof.** We only prove the first statement in the lemma; the second statement is proved in a similar manner. Existence of solutions was argued above; we now argue uniqueness. Suppose \( \eta \) and \( \tilde{\eta} \) are
two solutions of (6.15) in $C([0,T]: \mathcal{S}_v)$. Let $\xi := \eta - \tilde{\eta}$. Then $\xi$ satisfies

$$\langle \xi(t), \phi \rangle = \int_{0}^{t} \langle \xi(s), L(s)\phi \rangle ds.$$  \hfill (6.48)

It suffices to show $\xi = 0$. We adapt arguments of Kurtz and Xiong (see Lemma 4.2 and Appendix in [49]). By an analogous argument to Lemma A.6 in [49], using Condition 6.2, we have for all $f \in \mathcal{S}_v$,

$$\sup_{0 \leq s \leq T} \langle f, L^*(s)f \rangle_{-w} \leq \kappa_1\|f\|^2_{-w},$$  \hfill (6.49)

where $L^*(s): \mathcal{S}_v \to \mathcal{S}_w$ is the adjoint of $L(s): \mathcal{S}_w \to \mathcal{S}_v$. Recall that $\{\phi^w_j\}$ is an orthonormal basis for $\mathcal{S}_w$. We can choose this basis such that for each $j \in \mathbb{N}$, $\phi^w_j \in \mathcal{S}_v$. From (6.48) we have

$$\langle \xi(t), \phi^w_j \rangle^2 = 2 \int_{0}^{t} \langle \xi(s), \phi^w_j \rangle d\langle \xi(s), \phi^w_j \rangle = 2 \int_{0}^{t} \langle \xi(s), \phi^w_j \rangle \langle \xi(s), L(s)\phi^w_j \rangle ds.$$  

Therefore,

$$\|\xi(t)\|^2_{-w} = 2 \int_{0}^{t} \langle \xi(s), L^*(s)\xi(s) \rangle_{-w} ds \leq \kappa_2 \int_{0}^{t} \|\xi(s)\|^2_{-w} ds,$$  \hfill (6.50)

where the last inequality follows from (6.49). By Gronwall’s lemma $\xi(t) = 0 \forall t \in [0,T]$ and uniqueness follows.

We can now complete the proof of the Laplace lower bound.

**Proof of the Laplace lower bound:** The above lemma shows that $\tilde{Y}^m \Rightarrow \zeta^*$ in $C([0,T]: \mathcal{S}_v)$. Combining this with (6.46) and (6.45) gives us

$$\limsup_{m \to \infty} -a^2(m) \log E \exp \left\{ -\frac{1}{a^2(m)} F(Y^m) \right\} \leq \limsup_{m \to \infty} E \left\{ \frac{1}{2} \sum_{i=1}^{m} \int_{0}^{T} |\tilde{u}_i^m(s)|^2 ds + F(\tilde{Y}^m) \right\} \leq \frac{1}{2} \int_{\mathbb{R}^2 \times [0,T]} y^2 \nu^*(dy \, dx \, ds) + F(\zeta^*),$$

which can be further bounded by

$$I(\zeta^*) + F(\zeta^*) + \varepsilon \leq \inf_{\xi \in C([0,T]: \mathcal{S}_v)} \{I(\xi) + F(\xi)\} + 2\varepsilon.$$  

Since $\varepsilon > 0$ is arbitrary, we have the desired lower bound. \qed
6.2.5 \( I \) is a rate function

In this section we prove that under Conditions 6.1 and 6.3, \( I \) defined in (6.16) regarded as a map from \( C([0, T] : \mathcal{S}_{-\rho}) \) to \([0, \infty] \) has compact level sets and is therefore a rate function on \( C([0, T] : \mathcal{S}_{-\rho}) \).

Fix \( K \in (0, \infty) \) and let \( \Theta_K = \{ \eta \in C([0, T] : \mathcal{S}_{-\rho}) : I(\eta) \leq K \} \). Let \( \{ \eta^m \}_{m \in \mathbb{N}} \subset \Theta_K \). Then for each \( m \in \mathbb{N} \) there exists \( \nu^m \in T(\eta^m) \) such that

\[
\frac{1}{2} \int_{\mathbb{R}^2 \times [0, T]} y^2 \nu^m(dy \, dx \, ds) \leq K + \frac{1}{m}. \tag{6.51}
\]

It follows from (6.14) and (6.25) that

\[
\sup_{m \in \mathbb{N}} \int_{\mathbb{R}^2 \times [0, T]} x^2 \nu^m(dy \, dx \, ds) = \int_{\mathbb{R} \times [0, T]} x^2 \mu_s(dx) \, ds < \infty.
\]

So we have \( \sup_{m \in \mathbb{N}} |g(\nu^m)| < \infty \), where \( g \) is the tightness function on \( M_T(\mathbb{R}^2 \times [0, T]) \) defined in (6.37). Hence \( \{ \nu^m \} \) is pre-compact. Let \( \nu^m \) converge along a subsequence (labeled once more as \( \{ m \} \)) to \( \hat{\nu} \). It follows from Fatou’s lemma and (6.51) that \( \hat{\nu} \in \mathcal{P}_\infty \). Now let \( \hat{\eta} \) be defined as in (6.15) with \( \nu \) replaced by \( \hat{\nu} \). Note that from Lemma 6.14 there is a unique such \( \hat{\eta} \in C([0, T] : \mathcal{S}_{-\rho}) \).

We claim that \( \eta^m \to \hat{\eta} \) in \( C([0, T] : \mathcal{S}_{-\rho}) \). Once the claim is verified, it will follow from (6.51) and Fatou’s lemma that \( I(\hat{\eta}) \leq K \), which will prove the desired compact level set property. Note that both \( \eta^m \) and \( \hat{\eta} \) are in \( C([0, T] : \mathcal{S}_{-\rho}) \) and if one could show that \( \eta^m \to \hat{\eta} \) in \( C([0, T] : \mathcal{S}_{-\rho}) \), we would have that \( I \) is a rate function on \( C([0, T] : \mathcal{S}_{-\rho}) \). However, that convergence is not immediately obvious and we only argue the weaker statement made in the claim.

Disintegrate \( \nu^m \) as

\[\nu^m(A \times B \times [0, t]) = \int_0^t \nu^m_s(A \times B) \, ds, \quad A, B \in \mathcal{B}(\mathbb{R}), t \in [0, T].\]

Since \( \nu^m \in \mathcal{P}_\infty \), we have \( \nu^m_s \in \mathcal{P}(\mathbb{R}^2) \) for a.e. \( s \in [0, T] \). Define for \( s \in [0, T] \), the function \( J^m(s) : \mathcal{S}_w \to \mathbb{R} \) as follows:

\[
\langle J^m(s), \phi \rangle := \int_{\mathbb{R}^2} \phi'(x) \sigma(x, \mu(s)) y \nu^m_s(dy \, dx), \quad \phi \in \mathcal{S}_w.
\]
It is easy to see that $J^m(s) \in \mathcal{S}_{-w}$ for a.e. $s \in [0, T]$, in fact it follows from (6.12) and (6.51) that
\[
\sup_{m \in \mathbb{N}} \int_0^T \|J^m(s)\|_{-w}^2 \, ds = \sup_{m \in \mathbb{N}} \int_0^T \sup_{\phi \in \mathbb{R}^2, L(s), \phi} \left( \int_{\mathbb{R}^2} \phi(x) \sigma(x, \mu(s)) \nu^m_x(dy \, dx) \right)^2 \, ds \\
\leq \kappa_1 \sup_{m \in \mathbb{N}} \int_{\mathbb{R}^2 \times [0, T]} \gamma^2 \nu^m(dy \, dx \, ds) < \infty.
\] (6.52)

Since $\nu^m \in \mathcal{T}(\eta^m)$, it follows that
\[
\langle \eta^m(t), \phi \rangle = \int_0^t \langle \eta^m(s), L(s)\phi \rangle \, ds + \int_0^t \langle J^m(s), \phi \rangle \, ds, \quad \forall \phi \in \mathcal{S}.
\]

Analogous to the proof of Lemma 6.14 we have
\[
\langle \eta^m(t), \phi \rangle^2 = 2 \int_0^t \langle \eta^m(s), \phi \rangle \langle \eta^m(s), L(s)\phi \rangle \, ds + 2 \int_0^t \langle \eta^m(s), \phi \rangle \langle J^m(s), \phi \rangle \, ds.
\]

Also note that since $I(\eta^m) \leq K < \infty$, we must have that $\eta^m \in \mathbb{C}([0, T] : \mathcal{S}_{-w}) \subset \mathbb{C}([0, T] : \mathcal{S}_{-w})$.

Thus, as for the proof of (6.50),
\[
\|\eta^m(t)\|_{-w}^2 = 2 \int_0^t \langle \eta^m(s), L^*(s)\eta^m(s) \rangle_{-w} \, ds + 2 \int_0^t \langle \eta^m(s), J^m(s) \rangle_{-w} \, ds \\
\leq \kappa_2 \int_0^t \|\eta^m(s)\|_{-w}^2 \, ds + \kappa_2 \int_0^t \|\eta^m(s)\|_{-w} \|J^m(s)\|_{-w} \, ds \\
\leq \kappa_3 \int_0^t \|\eta^m(s)\|_{-w}^2 \, ds + \kappa_3 \int_0^t \|J^m(s)\|_{-w}^2 \, ds.
\]

Applying Gronwall’s lemma and using (6.52), we have
\[
\sup_{m \in \mathbb{N}} \sup_{t \in [0, T]} \|\eta^m(t)\|_{-w}^2 < \infty.
\] (6.53)

Next note that for $t_1, t_2 \in [0, T]$ and $\phi \in \mathcal{S}$, by Cauchy–Schwarz inequality
\[
|\langle \eta^m(t_2), \phi \rangle - \langle \eta^m(t_1), \phi \rangle|^2 \leq 2|t_2 - t_1| \left( \int_0^T \|\eta^m(s), L(s)\phi \|^2 \, ds + \int_0^T \|J^m(s), \phi \|^2 \, ds \right) \\
\leq 2|t_2 - t_1| \left( \int_0^T \|\eta^m(s)\|_{-w}^2 \|L(s)\phi\|^2 \, ds + \int_0^T \|J^m(s)\|_{-w}^2 \|\phi\|^2 \, ds \right) \\
\leq \kappa_4 |t_2 - t_1| \|\phi\|^2_{w+2},
\]
where the last inequality follows from (6.53), Lemma 6.12 and (6.52). This together with (6.53) implies that \( \{\eta^m\} \) is pre-compact in \( \mathbb{C}([0,T] : \mathcal{S}_\rho) \), where \( \rho \) is as introduced below (6.16) (see e.g. Theorem 2.5.2 in [44]). Suppose now that \( \eta^m \) converges in \( \mathbb{C}([0,T] : \mathcal{S}_\rho) \) along a subsequence (labeled once more as \( \{m\} \)) to \( \tilde{\eta} \). Under Condition 6.3, for every \( \phi \in \mathcal{S} \) and \( s \in [0,T] \), \( L(s)\phi \in \mathcal{S}_\rho \).

Thus \( \langle \eta^m(s), L(s)\phi \rangle \to \langle \tilde{\eta}(s), L(s)\phi \rangle \ \forall \ s \in [0,T] \). Finally, using (6.51), (6.53), the convergence of \( \nu^m \) to \( \hat{\nu} \), and an estimate analogous to the one below (6.41) (with \( v \) replaced by \( \rho \) and \( w \) by \( \rho + 2 \)), we see that any limit point \( \tilde{\eta} \) of \( \{\eta^m\} \) solves (6.15) with \( \nu \) replaced by \( \hat{\nu} \). From the second statement in Lemma 6.14 this equation has a unique solution in \( \mathbb{C}([0,T] : \mathcal{S}_\rho) \) and so we must have that \( \tilde{\eta} = \hat{\eta} \). This proves the desired compactness of \( \Theta_K \).
CHAPTER 7
MODERATE DEVIATION PRINCIPLES FOR WEAKLY INTERACTING
PURE JUMP MARKOV PROCESSES

In this chapter we consider a family of models of weakly interacting Markov processes with a countable state space. Consider for \( m \in \mathbb{N} \), a pure jump Markov process \( \{(X^m(t_1), \ldots, X^m(t_m)) : t \in [0, T]\} \) taking values in \( \mathbb{N}^m \) with \( X^m(0) = x^m \). The evolution of the process is described through the jump intensities that are given as follows:

Given \( (X^m_1(t-), \ldots, X^m_m(t-)) = (x_1, \ldots, x_m) \in \mathbb{N}^m \), for \( i \in \{1, \ldots, m\} \) and \( y \in \mathbb{N}, y \neq x_i \),

\[
(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_m) \mapsto (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_m) \quad (7.1)
\]
at rate \( \Gamma_{x_i,y}(\mu^m(t-)) \), where \( \mu^m(t-) = \frac{1}{m} \sum_{i=1}^{m} \delta_{x_i} = \frac{1}{m} \sum_{i=1}^{m} \delta_{X^m_i(t-)} \in \mathcal{P}(\mathbb{N}) \). All other forms of jump have rate 0. Here \( \Gamma(q) = (\Gamma_{ij}(q)) \in \mathcal{P}(\mathbb{N}) \) is a rate matrix for each \( q \in \mathcal{P}(\mathbb{N}) \), namely \( \Gamma_{ij}(q) \geq 0 \) for \( i \neq j \) and \( \Gamma_{ii}(q) = -\sum_{j \neq i}^{\infty} \Gamma_{ij}(q) > -\infty \). We identify \( \mathcal{P}(\mathbb{N}) \) with the simplex

\[
\hat{S} = \left\{ q = (q_1, q_2, \ldots) \in l_2 \mid \sum_{j=1}^{\infty} q_j = 1, q_j \geq 0 \ \forall j \in \mathbb{N} \right\}
\]
in \( l_2 \) (the Hilbert space of square-summable sequences, equipped with the usual inner product).

With suitable assumptions on the intensity function \( \Gamma \) and the initial configuration of the particles, it can be proved (see Theorem 7.2) that for each \( T > 0 \), \( \mu^m \) converges to \( p \) in \( \mathcal{D}([0, T] : l_2) \) for a continuous function \( p \) characterized as the unique solution of an \( l_2 \)-valued ODE (see (7.9)). We are interested in the asymptotics of the centered and scaled quantity \( Z^m = a(m) \sqrt{m} (\mu^m - p) \) regarded as a random variable with values in \( \mathcal{D}([0, T] : l_2) \), where \( a(m) \) is as defined in (6.1). In Theorem 7.4 we will establish a moderate deviation principle for \( \mu^m \) (which is formulated in terms of a LDP for \( Z^m \)) with the associated rate function \( \bar{I} \) introduced in (7.10). We also give an alternative expression for the rate function in (7.12) which is somewhat easier to interpret in terms of the model parameters.
The chapter is organized as follows. In Section 7.1.1 we give a convenient representation for
the associated empirical measure process in terms of a Poisson random measure on a suitable point
space. This section also presents some basic well-posedness results and a law of large numbers result
under a natural condition (Condition 7.1). Proof of such result follows from standard arguments,
however for completeness we provide a sketch in Section 7.3. The MDP for the empirical measure
process in this setting is given in Section 7.1.2. The main result is Theorem 7.4 which establishes a
MDP for \( \mu^m \) under Condition 7.3. Theorem 7.5 gives an alternative expression for the rate function.
Proofs of Theorems 7.4 and 7.5 are given in Section 7.2.

7.1 Model and main results

7.1.1 Model and law of large numbers

Recall the pure jump Markov process \( \{(X_1^m(t), \ldots, X_m^m(t)), t \in [0, T]\} \) governed by intensity
function \( \Gamma \) that was introduced at the beginning of this chapter. It will be convenient to describe
the evolution of the associated empirical measure process \( \{\mu^m(t)\} \) through an SDE driven by a
Poisson random measure. We now introduce some notation that will be needed to formulate this
evolution equation.

For a locally compact Polish space \( S \), let \( \mathcal{M}_{FC}(S) \) be the space of all measures \( \nu \) on \((S, \mathcal{B}(S))\)
such that \( \nu(K) < \infty \) for every compact \( K \subset S \). We equip \( \mathcal{M}_{FC}(S) \) with the usual vague topology.
This topology can be metrized such that \( \mathcal{M}_{FC}(S) \) is a Polish space (see for example [13]). A Poisson
random measure (PRM) \( n \) on \( S \) with mean measure (or intensity measure) \( \nu \) is a \( \mathcal{M}_{FC}(S) \)-valued
random variable such that for each \( B \in \mathcal{B}(S) \) with \( \nu(B) < \infty \), \( n(B) \) is Poisson distributed with
mean \( \nu(B) \) and for disjoint \( B_1, \ldots, B_k \in \mathcal{B}(S) \), \( n(B_1), \ldots, n(B_k) \) are mutually independent random
variables (cf. [41]).

Let \( l_2 \) be the Hilbert space of square-summable sequences, equipped with the usual inner
product and norm denoted by \( \langle \cdot, \cdot \rangle \) and \( \| \cdot \| \), respectively. For each \( i \in \mathbb{N} \) let \( e_i \) be the unit vector
in \( l_2 \) with 1 for the \( i \)-th coordinate and 0 otherwise. For Banach spaces \( B_1 \) and \( B_2 \), \( L(B_1, B_2) \) will
denote the space of bounded linear operators from \( B_1 \) to \( B_2 \).

We are interested in characterizing the limit of the empirical measure process \( \{\mu^m(t)\} \) in the
space \( \mathbb{D}([0, T]; l_2) \), and to establish a moderate deviation principle for \( \{\mu^m(t)\} \). We begin by giving
an equivalent in law representation of this empirical measure process using a PRM on a suitable
point space. We will follow the notation in [13].
Let $X \equiv \mathbb{R}_+^2$, $Y \equiv X \times \mathbb{R}_+ = \mathbb{R}_+^3$, $X_T \equiv [0, T] \times \mathbb{R}$ and $Y_T \equiv [0, T] \times \mathbb{R}$. Let $\lambda_T$, $\lambda_X$ and $\lambda_\infty$ be the Lebesgue measures on $[0, T]$, $X$ and $\mathbb{R}_+$, respectively. Let $N$ be a PRM on $Y_T$ with intensity $\lambda_{Y_T} = \lambda_T \otimes \lambda_X \otimes \lambda_\infty$, defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ with a $\mathbb{P}$-complete right-continuous filtration. We assume that $N([0, a] \times A)$ is $\mathcal{F}_a$-measurable and $N((a, b] \times A)$ is independent of $\mathcal{F}_a$ for all $0 \leq a < b \leq T$ and $A \in \mathcal{B}(\mathbb{R})$. Given $m \geq 1$, let $N^m$ be a counting process on $X_T$ defined as

$$N^m([0, t] \times A) \equiv \int_{[0, t] \times A} 1_{[0, m]}(r) N(ds \, dy \, dr), \quad t \in [0, T], A \in \mathcal{B}(X).$$

We will make the following assumption on $\Gamma$ (this assumption will be restated in Condition 7.1)

$$\|\Gamma\|_\infty \equiv \sup_{q \in \hat{S}} \sup_{i \in \mathbb{N}} |\Gamma_{ii}(q)| < \infty. \quad (7.2)$$

Given $i, j \in \mathbb{N}$ with $i \neq j$ and $q \in \hat{S}$, let

$$A_{ij}(q) \equiv \{y \in X : i - 1 < y_1 \leq i, (j - 1)\|\Gamma\|_\infty < y_2 \leq (j - 1)\|\Gamma\|_\infty + q, \Gamma_{ij}(q)\}. \quad (7.3)$$

Note that for every $q \in \hat{S}$,

$$A_{ij}(q) \cap A_{i'j'}(q) = \emptyset, \text{ if } (i, j) \neq (i', j'). \quad (7.3)$$

For $q \in \hat{S}$ and $y \in X$, let

$$G(q, y) \equiv \sum_{i=1}^\infty \sum_{j=1, j \neq i}^\infty (e_j - e_i)1_{A_{ij}(q)}(y) = \sum_{i=1}^\infty G_i(q, y)e_i, \quad (7.4)$$

$$G_i(q, y) \equiv \sum_{j=1, j \neq i}^\infty (1_{A_{ji}(q)}(y) - 1_{A_{ij}(q)}(y)), \quad i \in \mathbb{N}. \quad (7.5)$$

Note that $\|G(q, y)\| \leq \sqrt{2}$ for all $q \in \hat{S}$ and $y \in X$, in particular $G$ is a well defined map from $l_2 \times X$ to $l_2$. Define the stochastic process $\{\mu^m(t), t \in [0, T]\}$ as

$$\mu^m(t) = \mu^m(0) + \frac{1}{m} \int_{X_t} G(\mu^m(s-), y) N^m(ds \, dy), \quad (7.6)$$

121
where \( \mu^m(0) = \frac{1}{m} \sum_{i=1}^{m} \delta_{x_i^m} \) and \( X_t = [0,t] \times \mathbb{X} \) for each \( t \in [0,T] \). This describes a pure jump Markov process for which jump at time instant \( t \) is \( \frac{1}{m}(e_j - e_i) \) at rate \( m\lambda_x(A_{ij}(q)) = mq_i \Gamma_{ij}(q) \) with \( q = \mu^m(t-) \) for \( i, j \in \mathbb{N} \) with \( i \neq j \). Thus \( \{\mu^m(t)\} \) defined by (7.6) has the same law as the empirical process \( \{\frac{1}{m} \sum_{i=1}^{m} \delta_X(t_i)\} \) introduced at the beginning of this chapter. Throughout this chapter we will use the representation for \( \{\mu^m(t)\} \) given by (7.6). With this representation \( \{\mu^m(t)\} \) can be viewed as an Hilbert space \((l^2)\)-valued small noise stochastic dynamical system driven by a PRM. Moderate deviation principles for such small noise processes have been studied in [12]. However one key difference between the models in [12] from that considered here is that unlike in [12] the map \( x \mapsto G(x,y) \) is not Lipschitz (in fact not even continuous). This lack of regularity is one of the challenges in the large deviation analysis.

Let \( \tilde{N}^m(ds \, dy) \equiv N^m(ds \, dy) - m \lambda(ds \, dy) \), where \( \lambda = \lambda_T \otimes \lambda_X \) is the Lebesgue measure on \( \mathbb{X}_T \). Then (7.6) can be written as:

\[
\mu^m(t) = \mu^m(0) + \int_0^t b(\mu^m(s)) \, ds + \frac{1}{m} \int_{\mathbb{X}^m} G(\mu^m(s-), y) \, \tilde{N}^m(ds \, dy),
\]

(7.7)

where \( b: \hat{S} \rightarrow l_2 \) is defined as

\[
b(q) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} q_j \Gamma_{ji}(q) e_i, \quad q = (q_1, q_2, \ldots) \in \hat{S}.
\]

(7.8)

To see that \( b(q) \) defined by (7.8) is in \( l_2 \), note that for \( q = (q_1, q_2, \ldots) \in \hat{S} \),

\[
\|b(q)\|^2 = \sum_{i=1}^{\infty} \left| \sum_{j=1}^{\infty} q_j \Gamma_{ji}(q) \right|^2 \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} q_j^2 \Gamma^2_{ji}(q) \leq 2\|\Gamma\|^2 \infty.
\]

Note also that \( b(q) = \int_{\mathbb{X}} G(q, y) \lambda_X(dy) \).

We now introduce an assumption under which a law of large numbers result holds. Note that part (a) was previously stated in (7.2).

**Condition 7.1.** (a) \( \|\Gamma\|_\infty < \infty \).

(b) The map \( b: \hat{S} \rightarrow l_2 \) defined in (7.8) is Lipschitz, namely there exists \( L_b \in (0, \infty) \) such that for all \( q, \tilde{q} \in \hat{S} \), \( \|b(q) - b(\tilde{q})\| \leq L_b \|q - \tilde{q}\| \).

(c) \( \|\mu^m(0) - p(0)\| \rightarrow 0 \) as \( m \rightarrow \infty \) for some probability measure \( p(0) \in \mathcal{P}(\mathbb{N}) \).
Remark 7.1. Condition 7.1(c) is trivially satisfied if \( p(0) = \delta_x \) and \( x^n_i = x \) for all \( m, i \in \mathbb{N} \), for some \( x \in \mathbb{N} \). An elementary application of Scheffé’s lemma and the strong law of large numbers shows that it is also satisfied for a.e. \( \omega \) if \( x^n_i = \xi_i(\omega) \) where \( \xi_i \) are i.i.d. with common law \( p(0) \).

Let \( \mathbb{M} = \mathcal{M}_{FC}(\mathbb{X}_T) \), namely the space of all measures \( \nu \) on \((\mathbb{X}_T, \mathcal{B}(\mathbb{X}_T))\) such that \( \nu(K) < \infty \) for every compact \( K \subset \mathbb{X}_T \). The proof of the following result giving unique solvability and law of large numbers is standard. We provide a sketch in Section 7.3 for completeness.

Theorem 7.2. Under Condition 7.1, the following conclusions hold.

(a) For each \( m \in \mathbb{N} \) there is a measurable map \( \tilde{G}^m: \mathbb{M} \to \mathbb{D}([0,T]: l_2) \) such that for any probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})\) on which is given a Poisson random measure \( n_m \) on \( \mathbb{X}_T \) with intensity measure \( m\lambda \), \( \tilde{\mu}^m \approx \tilde{G}^m \left( \frac{1}{m} n_m \right) \) is an \( \tilde{\mathcal{F}}_t \approx \sigma \left\{ n_m([0,s] \times A), s \leq t, A \in \mathcal{B}(\mathbb{X}) \right\} \)-adapted RCLL process that is the unique adapted solution of the stochastic integral equation

\[
\tilde{\mu}^m(t) = \tilde{\mu}^m(0) + \frac{1}{m} \int_{\mathbb{X}_t} G(\tilde{\mu}^m(s-), y) \, n_m(ds \, dy), \quad t \in [0,T].
\]

In particular \( \mu^m \approx \tilde{G}^m \left( \frac{1}{m} N^m \right) \) is the unique \( \{\mathcal{F}_t\} \)-adapted solution of (7.6).

(b) The process \( \mu^m \) converges in probability to \( p \) in \( \mathbb{D}([0,T]: l_2) \), where \( p \) is given as the unique solution of the following integral equation in \( l_2 \):

\[
p(t) = p(0) + \int_0^t b(p(s)) \, ds, \quad t \in [0,T]. \tag{7.9}
\]

7.1.2 Moderate deviation principle

Let \( a(m) \) be as in (6.1). We will now study a large deviation principle for \( Z^m = a(m) \sqrt{m}(\mu^m - p) \). We make the following stronger assumption in place of Condition 7.1.

Condition 7.3. (a) \( \|\Gamma\|_\infty < \infty \).

(b) There exists \( c_{\Gamma} \in (0, \infty) \) such that for all \( q \in \tilde{S} \) and \( j \in \mathbb{N} \), \( \sum_{i=1}^\infty |\Gamma_{ij}(q)| \leq c_{\Gamma} \). There exists \( L_{\Gamma} \in (0, \infty) \) such that for all \( \tilde{q}, q \in \tilde{S} \) and \( i \in \mathbb{N} \), \( \sum_{j=1,j \neq i}^\infty |\Gamma_{ij}(\tilde{q}) - \Gamma_{ij}(q)| \leq L_{\Gamma} \|\tilde{q} - q\| \).

(c) For the map \( b: \tilde{S} \to l_2 \) defined in (7.8), there exist \( c_b \in (0, \infty) \); a map \( Db: \tilde{S} \to L(l_2, l_2) \); and \( \theta_b: \tilde{S} \times \tilde{S} \to l_2 \) such that for all \( q, \tilde{q} \in \tilde{S} \),

\[
b(\tilde{q}) - b(q) = Db(q)[\tilde{q} - q] + \theta_b(q, \tilde{q})
\]
and \( \| \theta_b(q, \tilde{q}) \| \leq c_b \| \tilde{q} - q \|^2 \), \( \| Db(q) \|_{L^2(\mathbb{L}_1)} \leq c_b \).

(d) \( a(m) \sqrt{m} \| \mu^m(0) - p(0) \| \to 0 \) as \( m \to \infty \) for some probability measure \( p(0) \in \mathcal{P}(\mathbb{N}) \).

**Remark 7.2.** (i) Condition 7.1(b) is implied by Condition 7.3(c) while Condition 7.1(c) is implied by Condition 7.3(d).

(ii) Condition 7.3(b) is satisfied in particular for finite range jump models of the following form: There exists some \( K \in (0, \infty) \) such that for all \( q \in \hat{S} \), \( \Gamma_{ij}(q) = 0 \) for \( |i - j| > K \) and \( q \mapsto \Gamma_{ij}(q) \) is Lipschitz continuous with \( \| \Gamma_{ij} \|_L \leq K \) for \( |i - j| \leq K \).

(iii) Condition 7.3(d) is trivially satisfied if \( p(0) = \delta_x \) and \( x^m_i = x \) for all \( m, i \in \mathbb{N} \), for some \( x \in \mathbb{N} \). It is also satisfied for a.e. \( \omega \) if \( x^m_i = \xi_i(\omega) \) where \( \xi_i \) are i.i.d. with common distribution \( p(0) \) and \( \sum_{m=1}^{\infty} [a(m)]^{2n} < \infty \) for some \( n \in \mathbb{N} \) (for a proof see Section 7.3). This summability property is satisfied quite generally, e.g. when \( a(m) = O(m^{-\theta}) \) for some \( \theta \in (0, 1/2) \) or \( a(m) = O((\log m)^k m^{-\theta}) \) for some \( \theta \in (0, 1/2] \) and \( k > 0 \).

The following theorem is the main result of this section. The proof will be given in Section 7.2.1.

**Theorem 7.4.** Under Condition 7.3, \( \{Z^m\} \) satisfies a large deviation principle in \( \mathbb{D}([0, T] : l_2) \) with speed \( a^2(m) \) and the rate function given by

\[
\bar{I}(\eta) = \inf_{\psi} \left\{ \frac{1}{2} \| \psi \|^2_{L^2(\lambda)} \right\}, \quad \eta \in \mathbb{D}([0, T] : l_2),
\]

where the infimum is taken over all \( \psi \in L^2(\lambda) \) such that

\[
\eta(t) = \int_0^t Db(p(s)) [\eta(s)] \, ds + \int_{\mathbb{X}_t} G(p(s), y) \psi(s, y) \lambda(ds \, dy), \quad t \in [0, T].
\]

Along the lines of the proof of Theorem 7.2(b), it is easy to check that under Condition 7.3(c), (7.11) has a unique solution in \( \mathbb{C}([0, T] : l_2) \) for each \( \psi \in L^2(\lambda) \). In particular, \( \bar{I}(\eta) = \infty \) for all \( \eta \in \mathbb{D}([0, T] : l_2) \setminus \mathbb{C}([0, T] : l_2) \).

The rate function \( \bar{I} \) introduced in (7.10) is somewhat indirect in that its definition involves the extraneous function \( G \) that was introduced for the convenience of representation of \( \mu^m \) as a small noise stochastic dynamical system. The following result gives an alternative representation that is
more intrinsic. For $\eta \in \mathbb{D}([0, T] : l_2)$ let

$$I(\eta) = \inf_u \left\{ \frac{1}{2} \int_0^T \sum_{i=1}^\infty \sum_{j=1, j \neq i}^\infty u_{ij}^2(s) \, ds \right\}, \quad (7.12)$$

where the infimum is taken over all $u = \{u_{ij}\}_{i,j=1}^\infty$ with each $u_{ij} \in L^2([0, T] : \mathbb{R})$ such that

$$\eta(t) = \int_0^t Db(p(s))[\eta(s)] \, ds + \int_0^t \sum_{i=1}^\infty \sum_{j=1, j \neq i}^\infty (e_j - e_i) \sqrt{p_i(s)} \Gamma_{ij}(p(s)) u_{ij}(s) \, ds. \quad (7.13)$$

The proof of the following result will be given in Section 7.2.2.

**Theorem 7.5.** *Under the conditions of Theorem 7.4, $I = \bar{I}$.***

### 7.2 Proofs for the pure jump case

In this section we will prove Theorems 7.4 and 7.5. Throughout the section we assume that Condition 7.3 holds.

#### 7.2.1 Proof of Theorem 7.4

The basic idea is to make use of a sufficient condition for MDP presented in [12]. We begin with some notation.

Recall the PRM $N$ introduced in Section 7.1.1. Let $\bar{P}$ be the $\{\mathcal{F}_t\}$-predictable $\sigma$-field on $\Omega \times [0, T]$. We denote by $\bar{A}_+$ [resp. $\bar{A}$] the class of all $(\bar{P} \otimes \mathcal{B}(\mathbb{X}))/\mathcal{B}(\mathbb{R}_+)$ [resp. $(\bar{P} \otimes \mathcal{B}(\mathbb{X}))/\mathcal{B}(\mathbb{R})]$-measurable maps from $\Omega \times \mathbb{X}_T$ to $\mathbb{R}_+$ [resp. $\mathbb{R}$]. For $\varphi \in \bar{A}_+$, define a counting process $N^\varphi$ on $\mathbb{X}_T$ by

$$N^\varphi([0, t] \times U) = \int_{[0, t] \times U \times [0, \infty]} 1_{[0, \varphi(s,y)]}(r) \lambda(ds \, dy \, dr), \quad t \in [0, T], U \in \mathcal{B}(\mathbb{X}). \quad (7.14)$$

We think of $N^\varphi$ as a controlled random measure, with $\varphi$ the control process that produces a thinning of the point process $N$ in a random but non-anticipative manner.

Define $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$\ell(r) = r \log r - r + 1, \quad r \in \mathbb{R}_+.$$

For any $\varphi \in \bar{A}_+$ and $t \in [0, T]$ the quantity $L_t(\varphi) \triangleq \int_{\mathbb{X}_T} \ell(\varphi(s, y)) \lambda(ds \, dy)$ is well defined as a $[0, \infty]$-valued random variable. This quantity will appear as a cost term in the representation presented below. It will be convenient to restrict to the following smaller collection of controls. For each
\[ n \in \mathbb{N} \text{ let } \bar{A}_{b,n} = \{ \varphi \in \bar{A}_+ : \text{for all } (\omega, t) \in \Omega \times [0, T], \frac{1}{n} \leq \varphi(\omega, t, y) \leq n \text{ if } y \in [0, n]^2 \] and \( \varphi(\omega, t, y) = 1 \text{ if } y \notin [0, n]^2 \}\]

and let \( \bar{A}_b = \bigcup_{n=1}^{\infty} \bar{A}_{b,n} \).

For \( m \in \mathbb{N} \) and \( M \in (0, \infty) \), consider the spaces

\[ S^M_{+,m} = \left\{ g : \mathbb{X}_T \to \mathbb{R}_+ \mid L_T(g) \leq \frac{M}{a^2(m)n} \right\}, \tag{7.15} \]

\[ S^M_{m} = \left\{ f : \mathbb{X}_T \to \mathbb{R} \mid 1 + \frac{1}{a(m)\sqrt{n}}f \doteq g \in S^M_{+,m} \right\}, \tag{7.16} \]

\[ U^M_{+,m} = \{ \varphi \in \bar{A}_b \mid \varphi(\omega, \cdot, \cdot) \in S^M_{+,m}, \mathbb{P} \text{ a.s.} \}. \tag{7.17} \]

Given \( M \in (0, \infty) \), denote by \( B_2(M) \) the ball of radius \( M \) in \( L^2(\lambda) \). A set \( \{ \psi^m \} \subset \bar{A} \) with the property that \( \sup_{m \in \mathbb{N}} \| \psi^m \|_{L^2(\lambda)} \leq M \) a.s. for some \( M < \infty \) will be regarded as a collection of \( B_2(M) \)-valued random variables, where \( B_2(M) \) is equipped with the weak topology on the Hilbert space \( L^2(\lambda) \). Since \( B_2(M) \) is weakly compact, such a collection of random variables is automatically tight. Throughout this section \( B_2(M) \) will be regarded as the compact metric space obtained by equipping it with the weak topology on \( L^2(\lambda) \).

It follows from the paper [12] (see Lemma 7.8 below) that if \( g \in S^M_{+,m} \) then, with \( f \doteq a(m)\sqrt{m}(g - 1), \ f1\{|f| \leq a(m)\sqrt{m}\} \in B_2(C_M) \), where \( C_M \doteq \sqrt{M\tilde{\gamma}_2(1)} \) and \( \tilde{\gamma}_2(1) \in (0, \infty) \) is as in Lemma 7.7 below. Let \( \mathbb{S} \) be a Polish space. The following condition on a sequence \( \{G^m\} \) of measurable maps from \( \mathbb{M} \) to \( \mathbb{S} \) and a measurable map \( G_0 : L^2(\lambda) \to \mathbb{S} \) was introduced in [12] (see Condition 2.2 therein).

**Condition 7.6.** (a) Given \( M \in (0, \infty) \), suppose that \( g^m, g \in B_2(M) \) and \( g^m \to g \) as \( m \to \infty \). Then \( G_0(g^m) \to G_0(g) \).

(b) Given \( M \in (0, \infty) \), let \( \{\varphi^m\}_{m \in \mathbb{N}} \) be such that for every \( m \in \mathbb{N} \), \( \varphi^m \in U^M_{+,m} \) and for some \( \beta \in (0, 1] \), \( \psi^m1\{|\psi^m| \leq \beta a(m)\sqrt{m}\} \to \psi \text{ in } B_2(C_M) \text{ as } m \to \infty \), where \( \psi^m \doteq a(m)\sqrt{m}(\varphi^m - 1) \). Then

\[ G^m \left( \frac{1}{m} N^m \varphi^m \right) \Rightarrow G_0(\psi). \]
Theorem 2.3 of [12] says that under Condition 7.6, \( \{G^m(\frac{1}{m}N^m\varphi^m)\}_{m \in \mathbb{N}} \) satisfies a LDP on \( S \) with speed \( a^2(m) \) and rate function \( I \) given by

\[
I(\eta) = \inf_{\psi \in L^2(\lambda); \eta = G_0(\psi)} \left\{ \frac{1}{2} \| \psi \|^2_{L^2(\lambda)} \right\}, \quad \eta \in S. \tag{7.18}
\]

We will now use this theorem to establish a MDP for \( \mu^m \).

From Theorem 7.2 we have that there exists a measurable map \( \bar{G}^m: \mathbb{M} \rightarrow D([0,T]: l_2) \) such that \( \mu^m = \bar{G}^m(\frac{1}{m}N^m) \), and hence there is a map \( G^m: \mathbb{M} \rightarrow D([0,T]: l_2) \) such that with \( Z^m \) defined as above Condition 7.3, \( Z^m = G^m(\frac{1}{m}N^m) \). Define \( G_0: L^2(\lambda) \rightarrow D([0,T]: l_2) \) by

\[
G_0(\psi) \doteq \eta \text{ if for } \psi \in L^2(\lambda), \eta \text{ solves (7.11)} \tag{7.19}
\]

Note that the map is well defined since for each \( \psi \in L^2(\lambda) \) there is a unique \( \eta \in C([0,T]: l_2) \) solving (7.11). It is easy to check that with the above choice of \( G_0 \), \( I \) defined in (7.18) (with \( S = D([0,T]: l_2) \)) is same as the function \( \bar{I} \) introduced in (7.10). Thus in order to prove Theorem 7.4 it suffices to check that Condition 7.6 holds with \( S = D([0,T]: l_2) \) and the above choice of \( \{G^m\} \) and \( G_0 \). Rest of the section is devoted to the verification of this condition.

All statements except the last one in Lemma 7.7(a) below have been established in [12] (see Lemma 3.1 therein). The last statement in Lemma 7.7(a) is crucially used in our proofs and is a key ingredient in overcoming the lack of regularity of \( G \) (see proof of Proposition 7.17).

**Lemma 7.7.** (a) For each \( \beta > 0 \), there exist \( \bar{\gamma}_1(\beta), \bar{\gamma}'_1(\beta) \in (0, \infty) \) such that

\[
|x - 1| \leq \bar{\gamma}_1(\beta)\ell(x) \text{ for } |x - 1| \geq \beta, x \geq 0, \text{ and } x \leq \bar{\gamma}'_1(\beta)\ell(x) \text{ for } x \geq \beta > 1.
\]

Furthermore, \( \bar{\gamma}_1 \) can be selected to be such that for \( \beta \in (0, \frac{1}{2}) \), \( \bar{\gamma}_1(\beta) \leq \frac{4}{\beta} \).

(b) For each \( \beta > 0 \), there exists \( \bar{\gamma}_2(\beta) \in (0, \infty) \) such that \( |x - 1|^2 \leq \bar{\gamma}_2(\beta)\ell(x) \) for \( |x - 1| \leq \beta, x \geq 0 \).

**Proof.** We only need to prove the last statement in part (a). Note that we can set

\[
\bar{\gamma}_1(\beta) = \sup_{|x - 1| \geq \beta, x \geq 0} \frac{|x - 1|}{\ell(x)}.
\]
For $\beta \in (0, \frac{1}{2})$, $x \geq 0$ and $|x-1| \geq \beta$, consider function $f(x) \triangleq \frac{x-1}{\ell(x)}$. Since $\log(1+u) \leq u$ for $u \geq -1$, we have $f'(x) = \frac{\log x-(x-1)}{\ell^2(x)} < 0$ for $x \geq 0, |x-1| \geq \beta$. Since $f(0) = -1$, $\lim_{x \to 1^-} f(x) = -\infty$, $\lim_{x \to 1^+} f(x) = \infty$ and $\lim_{x \to \infty} f(x) = 0$, we have

$$\tilde{\gamma}_1(\beta) = \sup_{|x-1| \geq \beta, x \geq 0} \frac{|x-1|}{\ell(x)} = \sup_{|x-1| \geq \beta, x \geq 0} |f(x)| = \max\{|f(1+\beta)|, |f(1-\beta)|\}.$$ 

For $\beta \in (0, \frac{1}{2})$, let $h_1(\beta) = \beta |f(1+\beta)|$ and $h_2(\beta) = \beta |f(1-\beta)|$. Using elementary calculus, one can easily show that $h_i(\beta) \leq 4$ for $i = 1, 2$. The result follows.

The following lemma is taken from [12] (see Lemma 3.2 therein).

Lemma 7.8. Suppose $g \in S^M_+, \text{ for some } M \in (0, \infty)$. Let $f = a(m)\sqrt{m}(g-1) \in S^M_+$. Then

\[(a) \int_{\mathbb{X}} |f| \mathbf{1}_{\{|f| \leq \beta a(m)\sqrt{m}\}} \, d\lambda \leq \frac{M\tilde{\gamma}_1(\beta)}{a(m)\sqrt{m}} \text{ for } \beta > 0,\]
\[(b) \int_{\mathbb{X}} g \mathbf{1}_{\{g \geq \beta\}} \, d\lambda \leq \frac{M\tilde{\gamma}'_1(\beta)}{a^2(m)m} \text{ for } \beta > 1,\]
\[(c) \int_{\mathbb{X}} |f|^2 \mathbf{1}_{\{|f| \leq \beta a(m)\sqrt{m}\}} \, d\lambda \leq M\tilde{\gamma}_2(\beta) \text{ for } \beta > 0,\]

where $\tilde{\gamma}_1, \tilde{\gamma}'_1$ and $\tilde{\gamma}_2$ are as in Lemma 7.7.

We will now proceed to the verification of Condition 7.6. We begin with verifying part (a) of the condition. The following moment bounds on $G$ will be useful.

Lemma 7.9. For all $k \in \mathbb{N}$ and $q \in \hat{S}$, we have $\int_{\mathbb{X}} \|G(q, y)\|^k \lambda_{\mathbb{X}}(dy) \leq 2^{k/2}\|\Gamma\|_{\infty}$.

Proof. Recalling the definition of $G$ in (7.4), we have

$$\int_{\mathbb{X}} \|G(q, y)\|^k \lambda_{\mathbb{X}}(dy) = \sum_{i=1}^{\infty} \sum_{j \neq i} \int_{A_{ij}(q)} \|G(q, y)\|^k \lambda_{\mathbb{X}}(dy) = \sum_{i=1}^{\infty} \sum_{j \neq i} 2^{k/2} \lambda_{\mathbb{X}}(A_{ij}(q))$$

$$= \sum_{i=1}^{\infty} \sum_{j \neq i} 2^{k/2} q_i \Gamma_{ij}(q) \leq \sum_{i=1}^{\infty} 2^{k/2}\|\Gamma\|_{\infty} q_i = 2^{k/2}\|\Gamma\|_{\infty},$$

where the first two equalities use the property (7.3) of the sets $\{A_{ij}(q)\}$. The result follows. \qed

The following lemma provides a key convergence property.
Lemma 7.10. Fix $M \in (0, \infty)$. Suppose that $g^m, g \in B_2(M)$ and $g^m \to g$. Then

$$\int_{[0,\cdot] \times \mathbb{X}} g^m(s,y)G(p(s), y) \lambda(ds\,dy) \to \int_{[0,\cdot] \times \mathbb{X}} g(s,y)G(p(s), y) \lambda(ds\,dy) \text{ in } \mathbb{C}([0,T]:L_2).$$

Proof. It follows from Lemma 7.9 that $(s,y) \mapsto G_i(p(s),y)$ is in $L^2(\lambda)$ for each $i \in \mathbb{N}$. Thus, since $g^m \to g$ in $B_2(M)$, we have for every $t \in [0,T]$ and $i \in \mathbb{N}$,

$$\int_{\mathbb{X}_t} g^m(s,y)G_i(p(s), y) \lambda(ds\,dy) \to \int_{\mathbb{X}_t} g(s,y)G_i(p(s), y) \lambda(ds\,dy).$$

Note that by Cauchy–Schwarz inequality, for each $i \in \mathbb{N}$,

$$\left| \int_{\mathbb{X}_t} g^m(s,y)G_i(p(s), y) \lambda(ds\,dy) \right| \leq M \left( \int_{\mathbb{X}_t} G_i^2(p(s), y) \lambda(ds\,dy) \right)^{1/2} = \alpha_i.$$

From Lemma 7.9 we see that $\sum_{i=1}^{\infty} \alpha_i^2 < \infty$ and so by dominated convergence theorem,

$$\int_{\mathbb{X}_t} g^m(s,y)G(p(s), y) \lambda(ds\,dy) \to \int_{\mathbb{X}_t} g(s,y)G(p(s), y) \lambda(ds\,dy), \forall t \in [0,T]. \quad (7.20)$$

Also note that by Cauchy–Schwarz inequality and Lemma 7.9 once again, for $0 \leq s \leq t \leq T$

$$\left\| \int_{[s,t] \times \mathbb{X}} g^m(s,y)G(p(s), y) \lambda(ds\,dy) \right\|^2 \leq M^2 \int_{[s,t] \times \mathbb{X}} \|G(p(s), y)\|^2 \lambda(ds\,dy) \leq 2\|\Gamma\|_\infty M^2 |t-s|.$$

This implies equicontinuity, which shows that the convergence in (7.20) is in fact uniform in $t$. \qed

Now we are able to verify part (a) of Condition 7.6.

Proposition 7.11. Fix $M \in (0, \infty)$. Suppose that $g^m, g \in B_2(M)$ and $g^m \to g$. Let $G_0$ be as defined in (7.19). Then $G_0(g^m) \to G_0(g)$.

Proof. Let $\eta^m \doteq G_0(g^m)$ and $\eta \doteq G_0(g)$. Then

$$\eta^m(t) - \eta(t) = \int_0^t Db(p(s))[\eta^m(s) - \eta(s)] ds + \int_{\mathbb{X}_t} (g^m(s,y) - g(s,y))G(p(s), y) \lambda(ds\,dy).$$

The result now follows from Gronwall’s lemma together with Condition 7.3(c) and Lemma 7.10. \qed
In order to verify part (b) of Condition 7.6, we first prove some estimates. Recall spaces $\mathcal{S}^M_{+m}$ and $\mathcal{S}^M_m$ introduced in (7.15) and (7.16).

**Lemma 7.12.** Let $M \in (0, \infty)$. Then there exists $\tilde{\gamma}_3 \in (0, \infty)$ such that for all measurable maps $q : [0, T] \to \hat{S}$,

$$
\sup_{m \in \mathbb{N}} \sup_{g \in \mathcal{S}^M_{+m}} \int_{\mathcal{X} \times T} \|G(q(s), y)\|^2 g(s, y) \lambda(ds \, dy) \leq \tilde{\gamma}_3.
$$

**Proof.** Fix $g \in \mathcal{S}^M_{+m}$. By Lemmas 7.8(b), 7.9 and recalling that $\|G(q, y)\| \leq \sqrt{2}$, we have

$$
\int_{\mathcal{X} \times T} \|G(q(s), y)\|^2 g(s, y) \lambda(ds \, dy) \leq 2 \int_{\{g \geq 2\}} g(s, y) \lambda(ds \, dy) + 2 \int_{\mathcal{X} \times T} \|G(q(s), y)\|^2 \lambda(ds \, dy)
$$

$$
\leq \frac{2M\tilde{\gamma}'(2)}{a^2(m)m} + 4\|\Gamma\|_{\infty}T.
$$

The result follows since $a^2(m)m \to \infty$ as $m \to \infty$. \qed

The following lemma will be needed in the proof of the estimate (7.24) in Lemma 7.14 and (7.28) in Proposition 7.17.

**Lemma 7.13.** Let $M \in (0, \infty)$. Then there exists a map $\tilde{\gamma}_4 : (0, \infty) \to (0, \infty)$ such that for all $m \in \mathbb{N}$, $\beta \in (0, \infty)$, measurable $I \subset [0, T]$ and measurable maps $q : [0, T] \to \hat{S}$,

$$
\sup_{f \in \mathcal{S}^M_m} \int_{I \times \mathcal{X}} \|G(q(s), y)\| |f(s, y)| \mathbf{1}_{\{|f| \geq \beta a(m) \sqrt{m}\}} \lambda(ds \, dy) \leq \frac{\tilde{\gamma}_4(\beta)}{a(m) \sqrt{m}}
$$

and

$$
\sup_{f \in \mathcal{S}^M_m} \left\| \int_{I \times \mathcal{X}} G(q(s), y) f(s, y) \lambda(ds \, dy) \right\| \leq \tilde{\gamma}_4(\beta) \left( \frac{1}{a(m) \sqrt{m}} + \sqrt{|I|} \right).
$$

**Proof.** Fix $f \in \mathcal{S}^M_m$. Note that

$$
\left\| \int_{I \times \mathcal{X}} G(q(s), y) f(s, y) \lambda(ds \, dy) \right\|
$$

$$
\leq \int_{I \times \mathcal{X}} \|G(q(s), y)\| |f(s, y)| \mathbf{1}_{\{|f| \geq \beta a(m) \sqrt{m}\}} \lambda(ds \, dy)
$$

$$
+ \int_{I \times \mathcal{X}} G(q(s), y) \|f(s, y)| \mathbf{1}_{\{|f| < \beta a(m) \sqrt{m}\}} \lambda(ds \, dy).
$$

130
It follows from Lemma 7.8(a) that

\[
\int_{I \times X} \|G(q(s), y)\| |f(s, y)| \mathbf{1}_{\{|f(s, y)| \geq \beta a(m) \sqrt{m}\}} \lambda(ds \, dy) \\
\leq \sqrt{2} \int_{I \times X} |f(s, y)| \mathbf{1}_{\{|f(s, y)| \geq \beta a(m) \sqrt{m}\}} \lambda(ds \, dy) \leq \frac{\sqrt{2} M \tilde{G}_1(\beta)}{a(m) \sqrt{m}}.
\]

This proves the first inequality in the lemma. From Cauchy–Schwarz inequality it follows that

\[
\int_{I \times X} \|G(q(s), y)\| |f(s, y)| \mathbf{1}_{\{|f(s, y)| < \beta a(m) \sqrt{m}\}} \lambda(ds \, dy) \\
\leq \left( \int_{I \times X} \|G(q(s), y)\|^2 \lambda(ds \, dy) \right)^{1/2} \left( \int_{I \times X} |f(s, y)|^2 \mathbf{1}_{\{|f(s, y)| < \beta a(m) \sqrt{m}\}} \lambda(ds \, dy) \right)^{1/2} \\
\leq \sqrt{2} \|\Gamma\|_{\infty} M \tilde{G}_2(\beta)|I|,
\]

where the last inequality follows from Lemmas 7.8(c) and 7.9. The second inequality in the lemma now follows by combining the above two displays. \[\Box\]

Recall the space \(U_{+}^M\) and the map \(G^m\) introduced in (7.17) and above (7.19), respectively. Let for \(\varphi \in U_{+}^M\), \(\bar{Z}^{m, \varphi} \doteq G^m(\frac{1}{m} N^{m, \varphi})\), where \(N^{m, \varphi}\) is as defined in (7.14). Then it follows from an application of Girsanov’s theorem that (see for example the arguments above Lemma 4.4 in [12])

\[
\bar{Z}^{m, \varphi} = a(m) \sqrt{m} (\bar{\mu}^{m, \varphi} - p),
\]

where \(\bar{\mu}^{m, \varphi}\) is the unique pathwise solution of

\[
\bar{\mu}^{m, \varphi}(t) = \mu^m(0) + \frac{1}{m} \int_{X_t} G(\bar{\mu}^{m, \varphi}(s-), y) N^{m, \varphi}(ds \, dy).
\]

The following moment bounds on \(\bar{Z}^{m, \varphi}\) will be useful for our analysis.

**Lemma 7.14.** For every \(M \in (0, \infty)\),

\[
\sup_{m \in \mathbb{N}} \sup_{\varphi \in U_{+}^M} \mathbf{E} \left\| \bar{Z}^{m, \varphi} \right\|_{s, T}^2 < \infty.
\]

131
Proof. Given \( \varphi \in U_{+}^{M} \), let

\[
\tilde{N}^{m\varphi}(ds\,dy) = N^{m\varphi}(ds\,dy) - m\varphi(s,y)\lambda(ds\,dy).
\]

Then recalling (7.9) and that \( b(q) = \int_{X} G(q,y) \lambda_{X}(dy) \),

\[
\begin{align*}
\bar{\mu}^{m,\varphi}(t) - p(t) &= \mu^{m}(0) - p(0) + \frac{1}{m} \int_{X_{t}} G(\bar{\mu}^{m,\varphi}(s-),y) \tilde{N}^{m\varphi}(ds\,dy) \\
&\quad + \int_{X_{t}} \left( G(\bar{\mu}^{m,\varphi}(s),y) - G(p(s),y) \right) \lambda(ds\,dy) \\
&\quad + \int_{X_{t}} G(\bar{\mu}^{m,\varphi}(s),y)(\varphi(s,y) - 1) \lambda(ds\,dy).
\end{align*}
\]

Let \( \psi \doteq a(m)\sqrt{m}(\varphi - 1) \). Using above display and (7.21), we can write

\[
Z^{m,\varphi} = A^{m} + M^{m,\varphi} + B^{m,\varphi} + C^{m,\varphi},
\]

(7.22)

where

\[
\begin{align*}
A^{m} &\doteq a(m)\sqrt{m}(\mu^{m}(0) - p(0)), \\
M^{m,\varphi}(t) &\doteq \frac{a(m)}{\sqrt{m}} \int_{X_{t}} G(\bar{\mu}^{m,\varphi}(s-),y) \tilde{N}^{m\varphi}(ds\,dy), \\
B^{m,\varphi}(t) &\doteq a(m)\sqrt{m} \int_{X_{t}} \left( G(\bar{\mu}^{m,\varphi}(s),y) - G(p(s),y) \right) \lambda(ds\,dy) \\
&\quad = a(m)\sqrt{m} \int_{0}^{t} \left( b(\bar{\mu}^{m,\varphi}(s)) - b(p(s)) \right) ds \\
C^{m,\varphi}(t) &\doteq \int_{X_{t}} G(\bar{\mu}^{m,\varphi}(s),y)\psi(s,y) \lambda(ds\,dy).
\end{align*}
\]

Noting that \( M^{m,\varphi} \) is a martingale, Doob’s inequality gives us

\[
\mathbf{E}\|M^{m,\varphi}\|_{*,T}^{2} \leq \frac{4a^{2}(m)}{m}\mathbf{E}\int_{X_{T}} \|G(\bar{\mu}^{m,\varphi}(s),y)\|^{2} m\varphi(s,y) \lambda(ds\,dy).
\]

It then follows from Lemma 7.12 that

\[
\sup_{\varphi \in U_{+}^{M}} \mathbf{E}\|M^{m,\varphi}\|_{*,T}^{2} \leq \kappa_{1}a^{2}(m).
\]

(7.23)
Using Cauchy–Schwarz inequality and Condition 7.1(b) we have for all $\varphi \in U^M_{m,m}$,

$$\|B_{m,\varphi}\|_{s,t}^2 \leq a^2(m)t \int_0^t \|b(\bar{\mu}^{m,\varphi}(s)) - b(p(s))\|^2 \, ds \leq TL_b^2 \int_0^t \|\bar{Z}^{m,\varphi}\|_{s,s}^2 \, ds.$$ 

Since $\psi \in S^M_m$ a.s., it follows from Lemma 7.13 that

$$\sup_{\varphi \in U^M_{m,m}} \|C_{m,\varphi}\|_{s,T}^2 \leq \kappa^2 \left( \frac{1}{a^2(m)\sqrt{m}} + \sqrt{T} \right)^2 \leq 2\kappa^2 \left( \frac{1}{a^2(m)\sqrt{m}} + T \right). \quad (7.24)$$

Collecting these estimates we have for some $\kappa^3 \in (0, \infty)$ and all $\varphi \in U^M_{m,m}$, $t \in [0, T]$,

$$E \|\bar{Z}^{m,\varphi}\|_{s,t}^2 \leq \kappa^3 \left( \|A^m\|^2 + a^2(m) + \frac{1}{a^2(m)\sqrt{m}} + 1 + \int_0^t E \|\bar{Z}^{m,\varphi}\|_{s,s}^2 \, ds \right).$$

The result now follows from Gronwall’s inequality, (6.1) and Condition 7.3(d). $\square$

Although $G(q, y)$ is not a continuous map, using the specific form of $G$ and properties of $\Gamma$, we can establish the following Lipschitz property.

**Lemma 7.15.** There exists $\tilde{\gamma}_5 \in (0, \infty)$ such that for all $g \in M_0(\mathcal{X})$ and all $q, \tilde{q} \in \tilde{S}$,

$$\left\| \int_{\mathcal{X}} \left( G(\tilde{q}, y) - G(q, y) \right) g(y) \lambda_X(dy) \right\| \leq \tilde{\gamma}_5 \|g\|_{\infty} \|\tilde{q} - q\|.$$

**Proof.** Observing that

$$\lambda_X(A_{ij}(\tilde{q}) \Delta A_{ij}(q)) = |\tilde{q}_i \Gamma_{ij}(\tilde{q}) - q_i \Gamma_{ij}(q)|$$

for $i \neq j$, where “$\Delta$” denotes the symmetric difference, we see

$$\left\| \int_{\mathcal{X}} \left( G(\tilde{q}, y) - G(q, y) \right) g(y) \lambda_X(dy) \right\|^2 \leq \|g\|_{\infty}^2 \sum_{i=1}^\infty \left( \int_{\mathcal{X}} \left| G_i(\tilde{q}, y) - G_i(q, y) \right| \lambda_X(dy) \right)^2 \leq \|g\|_{\infty}^2 \sum_{i=1}^\infty \sum_{j \neq i} \left( |\tilde{q}_i \Gamma_{ij}(\tilde{q}) - q_i \Gamma_{ij}(q)| + |\tilde{q}_j \Gamma_{ji}(\tilde{q}) - q_j \Gamma_{ji}(q)| \right)^2.$$
By adding and subtracting terms, we can bound the above display by

\[
4\|g\|_\infty^2 \left( \sum_{i=1}^{\infty} \left( \sum_{j \neq i}^{\infty} |\tilde{q}_j \Gamma_{ij}(\tilde{q}) - q_j \Gamma_{ij}(q)| \right)^2 + \sum_{i=1}^{\infty} \left( \sum_{j \neq i}^{\infty} |q_i \Gamma_{ij}(\tilde{q}) - q_i \Gamma_{ij}(q)| \right)^2 \right) + \sum_{i=1}^{\infty} \left( \sum_{j \neq i}^{\infty} |\tilde{q}_j \Gamma_{ji}(\tilde{q}) - q_j \Gamma_{ji}(q)| \right)^2 + \sum_{i=1}^{\infty} \left( \sum_{j \neq i}^{\infty} |q_j \Gamma_{ji}(\tilde{q}) - q_j \Gamma_{ji}(q)| \right)^2 \right] = 4\|g\|_\infty^2 \sum_{k=1}^{4} T_k.
\]

The terms \( T_k \) for \( k = 1, 2, 3, 4 \), can be estimated as follows.

\[
T_1 = \sum_{i=1}^{\infty} \left( \tilde{q}_i - q_i \right)^2 \left( \sum_{j \neq i}^{\infty} \Gamma_{ij}(\tilde{q}) \right)^2 \leq \|\Gamma\|_\infty^2 \|\tilde{q} - q\|^2.
\]

Also, from Condition 7.3(b),

\[
T_2 \leq \sum_{i=1}^{\infty} q_i^2 L_i^2 \|\tilde{q} - q\|^2 \leq L_i^2 \|\tilde{q} - q\|^2,
\]

\[
T_3 = \sum_{i=1}^{\infty} \left( \sum_{j \neq i}^{\infty} |\tilde{q}_j - q_j \Gamma_{ji}(\tilde{q})| \right)^2 \leq \sum_{i=1}^{\infty} \sum_{j \neq i}^{\infty} \left( |\tilde{q}_j - q_j|^2 \Gamma_{ji}(\tilde{q}) \right) \sum_{j \neq i}^{\infty} \Gamma_{ji}(\tilde{q})
\]

\[
\leq c_T \sum_{i=1}^{\infty} \sum_{j \neq i}^{\infty} |\tilde{q}_j - q_j|^2 \Gamma_{ji}(\tilde{q}) \leq c_T \|\Gamma\|_\infty \|\tilde{q} - q\|^2,
\]

\[
T_4 \leq \left( \sum_{i=1}^{\infty} \sum_{j \neq i}^{\infty} q_j |\Gamma_{ji}(\tilde{q}) - \Gamma_{ji}(q)| \right)^2 \leq \left( \sum_{j=1}^{\infty} q_j L \|\tilde{q} - q\| \right)^2 \leq L_i^2 \|\tilde{q} - q\|^2.
\]

The result follows by combining above estimates. \( \square \)

The following lemma will allow us to apply the continuous mapping theorem to deduce the key weak convergence property in the proof of Proposition 7.17.

**Lemma 7.16.** Let \( M \in (0, \infty) \). Given \( \varepsilon \in \mathbb{D}([0, T] : l_2) \) and \( f \in B_2(M) \), there exists a unique \( z \in \mathbb{D}([0, T] : l_2) \) solving the following equation:

\[
z(t) = \varepsilon(t) + \int_{0}^{t} Db(p(s))[z(s)] ds + \int_{X} G(p(s), y)f(s, y) \lambda(ds dy), \quad t \in [0, T], \quad (7.25)
\]

namely there exists a measurable map \( h : \mathbb{D}([0, T] : l_2) \times B_2(M) \to \mathbb{D}([0, T] : l_2) \) such that the solution to (7.25) can be written as \( z = h(\varepsilon, f) \). Moreover, \( h \) is continuous at \( (0, f) \) for every \( f \in B_2(M) \).
Proof. The existence and uniqueness of solutions of (7.25) and the measurability of the solution map are easy to check using Condition 7.3(c) in a manner similar to the proof of Theorem 7.2(b). To see the continuity at \((0, f)\) for \(f \in B_2(M)\), first note that (7.25) can be written as

\[
z(t) - \varepsilon(t) = \int_0^t \! \! Db(p(s)) [z(s) - \varepsilon(s)] ds + \int_0^t \! \! Db(p(s)) [\varepsilon(s)] ds + \int_{\mathbb{R}^t} G(p(s), y) f(s, y) \lambda(ds \, dy).
\]

Suppose \((\varepsilon^m, f^m) \to (0, f)\) in \(\mathbb{D}([0, T] : l_2) \times B_2(M)\) as \(m \to \infty\). Let \(z^m = h(\varepsilon^m, f^m)\) and \(z = h(0, f)\). Using the convergence that \(\varepsilon^m \to 0\) in \(\mathbb{D}([0, T] : l_2)\) and Condition 7.3(c), we see that \(\int_0^t \! \! Db(p(s)) [\varepsilon^m(s)] ds \to 0\) in \(\mathbb{C}([0, T] : l_2)\). It follows from Lemma 7.10 that

\[
\int_{[0, \cdot] \times \mathbb{X}} G(p(s), y) f^m(s, y) \lambda(ds \, dy) \to \int_{[0, \cdot] \times \mathbb{X}} G(p(s), y) f(s, y) \lambda(ds \, dy) \text{ in } \mathbb{C}([0, T] : l_2).
\]

Combining above results and applying Gronwall’s lemma gives us \(z^m - \varepsilon^m \to z - 0\) in \(\mathbb{C}([0, T] : l_2)\). Since \(\varepsilon^m \to 0\), we have that \(z^m \to z\) in \(\mathbb{D}([0, T] : l_2)\) and the result follows. 

We can now verify part (b) of Condition 7.6. Recall that for \(M \in (0, \infty)\), \(C_M = \sqrt{M\gamma_2(1)}\).

Proposition 7.17. Fix \(M \in (0, \infty)\). Let \(\{\varphi^m\}_{m \in \mathbb{N}}\) be such that for every \(m \in \mathbb{N}\), \(\varphi^m \in U_{a, m}^M\). Let \(\psi^m \mathbf{1}_{\{\psi^m| \leq \beta a(m)\sqrt{m}\}} \Rightarrow \psi\) in \(B_2(C_M)\) for some \(\beta \in (0, 1]\), where \(\psi^m = a(m)\sqrt{m}(\varphi^m - 1)\). Let \(G_0\) and \(G^m\) be as defined in and above (7.19), respectively. Then \(G^m(\frac{1}{m} N^m \varphi^m) \Rightarrow G_0(\psi)\).

Proof. We will use the notation from the proof of Lemma 7.14. From (6.1), (7.23) and Condition 7.3(d) we have that \(\mathbb{E}\|M^m \varphi^m\|_{s, T}^2 \to 0\) and \(\|A^m\|^2 \to 0\) as \(m \to \infty\). It follows from Condition 7.3(c) that

\[
b(\bar{m}^m \varphi^m(s)) - b(p(s)) = \frac{1}{a(m)\sqrt{m}} \int_0^t \! \! Db(p(s)) [\bar{Z}^m \varphi^m(s)] + \theta_b(p(s), \bar{m}^m \varphi^m(s)).
\]

Hence \(B^m \varphi^m = \tilde{B}^m \varphi^m + e_1^m \varphi^m\), where

\[
\tilde{B}^m \varphi^m(t) = \int_0^t \! \! Db(p(s)) [\bar{Z}^m \varphi^m(s)] ds,
\]

\[
e_1^m \varphi^m(t) = a(m)\sqrt{m} \int_0^t \! \! \theta_b(p(s), \bar{m}^m \varphi^m(s)) ds.
\]

135
From Condition 7.3(c) and Lemma 7.14 we see that

$$
\mathbb{E} \left\| \mathcal{E}^{m, \varphi_m} \right\|_{*, T} \leq a(m) \sqrt{m} \mathbb{E} \int_0^T \left\| \theta_b(p(s), \tilde{\mu}^{m, \varphi_m}(s)) \right\| ds \leq \frac{c_T}{a(m) \sqrt{m}} \mathbb{E} \left\| \tilde{Z}^{m, \varphi_m} \right\|_{*, T}^2 \to 0
$$
as \( m \to \infty \). Write \( C^{m, \varphi_m} = \tilde{C}^{m, \varphi_m} + \mathcal{E}^{m, \varphi_m}_2 + \mathcal{E}^{m, \varphi_m}_3 + \mathcal{E}^{m, \varphi_m}_4 \), where

\[
\begin{align*}
\tilde{C}^{m, \varphi_m}(t) & \doteq \int_{X_t} G(p(s), y) \psi^m(s, y) 1_{\{ |\psi^m(s, y)| \leq \beta a(m) \sqrt{m} \}} \lambda(ds, dy), \\
\mathcal{E}^{m, \varphi_m}_2(t) & \doteq \int_{X_t} G(p(s), y) \psi^m(s, y) 1_{\{ |\psi^m(s, y)| > \beta a(m) \sqrt{m} \}} \lambda(ds, dy), \\
\mathcal{E}^{m, \varphi_m}_3(t) & \doteq \int_{X_t} \left( G(\tilde{\mu}^{m, \varphi_m}(s), y) - G(p(s), y) \right) \psi^m(s, y) 1_{\{ |\psi^m(s, y)| > \delta_m \}} \lambda(ds, dy), \\
\mathcal{E}^{m, \varphi_m}_4(t) & \doteq \int_{X_t} \left( G(\tilde{\mu}^{m, \varphi_m}(s), y) - G(p(s), y) \right) \psi^m(s, y) 1_{\{ |\psi^m(s, y)| \leq \delta_m \}} \lambda(ds, dy),
\end{align*}
\]

and \( \delta_m \doteq (a(m) \sqrt{m})^{1/2} \to \infty \) as \( m \to \infty \). Then using Lemma 7.13 we see that

\[
\left\| \mathcal{E}^{m, \varphi_m}_2 \right\|_{*, T} \leq \frac{\bar{\gamma}_4(\beta)}{a(m) \sqrt{m}} \to 0
\]
as \( m \to \infty \). Also applying Lemma 7.8(a) with \( \beta = \frac{\delta_m}{a(m) \sqrt{m}} \), we see that as \( m \to \infty \),

\[
\left\| \mathcal{E}^{m, \varphi_m}_3 \right\|_{*, T} \leq \int_{X_T} \left\| G(\tilde{\mu}^{m, \varphi_m}(s), y) - G(p(s), y) \right\| |\psi^m(s, y)| 1_{\{ |\psi^m(s, y)| > \delta_m \}} \lambda(ds, dy)
\]
\[
\leq 2 \sqrt{2} \int_{X_T} |\psi^m(s, y)| 1_{\{ |\psi^m(s, y)| > \delta_m \}} \lambda(ds, dy)
\]
\[
\leq \frac{2 \sqrt{2} M \bar{\gamma}_1(\frac{\delta_m}{a(m) \sqrt{m}})}{a(m) \sqrt{m}}
\]
\[
\leq \frac{8 \sqrt{2} M}{\delta_m} \to 0;
\]

where the last inequality is a consequence of the last statement in Lemma 7.7(a). Next, it follows from Lemma 7.15 that

\[
\left\| \mathcal{E}^{m, \varphi_m}_4 \right\|_{*, T} \leq \int_0^T \left\| \int_{X} \left( G(\tilde{\mu}^{m, \varphi_m}(s), y) - G(p(s), y) \right) \psi^m(s, y) 1_{\{ |\psi^m(s, y)| \leq \delta_m \}} \lambda(dy) \right\| ds
\]
\[
\leq \bar{\gamma}_5 \delta_m \int_0^T \left\| \tilde{\mu}^{m, \varphi_m}(s) - p(s) \right\| ds \leq \bar{\gamma}_5 T \frac{\delta_m}{a(m) \sqrt{m}} \left\| \tilde{Z}^{m, \varphi_m} \right\|_{*, T}.
\]
Since
\[ \frac{\delta_m}{a(m)\sqrt{m}} = (a(m)\sqrt{m})^{-\frac{1}{4}} \to 0, \]
it follows from Lemma 7.14 that \( E\|E_4^{m,\varphi^m}\|_{s,T} \to 0 \) as \( m \to \infty \). Putting above estimates together we have from (7.22)
\[ \tilde{Z}^{m,\varphi^m}(t) = E^{m,\varphi^m}(t) + B^{m,\varphi^m}(t) + C^{m,\varphi^m}(t) \]
\[ = E^{m,\varphi^m}(t) + \int_0^t Db(p(s))\left[ \tilde{Z}^{m,\varphi^m}(s) \right] ds \]
\[ + \int_{\mathcal{X}_t} G(p(s), y)\psi^m(s, y) 1_{\{\psi^m \leq \beta a(m)\sqrt{m}\}} \lambda(ds dy), \]
where \( E^{m,\varphi^m} = M^{m,\varphi^m} + A^m + \sum_{k=1}^4 E_k^{m,\varphi^m} \Rightarrow 0 \) in \( \mathbb{D}([0, T] : l_2) \). Thus we have
\[ G^m\left( \frac{1}{m} \Sigma^{m,\varphi^m} \right) = \tilde{Z}^{m,\varphi^m} = h(E^{m,\varphi^m}, \psi^m 1_{\{\psi^m \leq \beta a(m)\sqrt{m}\}}), \]
where \( h \) is as introduced in Lemma 7.16. It follows from Lemma 7.8(c) that \( \psi^m 1_{\{\psi^m \leq \beta a(m)\sqrt{m}\}} \) takes values in \( B_2(C_M) \) for all \( m \in \mathbb{N} \). Finally note that \( G_0(\psi) = h(0, \psi) \) and
\[ (E^{m,\varphi^m}, \psi^m 1_{\{\psi^m \leq \beta a(m)\sqrt{m}\}}) \Rightarrow (0, \psi). \]
The result now follows by combining above observations and applying continuous mapping theorem together with Lemma 7.16.

Now we can complete the proof of Theorem 7.4.

**Proof of Theorem 7.4:** As noted earlier, it suffices to show that Condition 7.6 holds with \( G^n \) and \( G_0 \) above and in (7.19), respectively. Part (a) of the condition was verified in Proposition 7.11, while part (b) was verified in Proposition 7.17.

**7.2.2 Proof of Theorem 7.5**

Fix \( \eta \in \mathbb{D}([0, T] : l_2) \). We first argue that \( \tilde{I}(\eta) \leq I(\eta) \). Let \( \delta > 0 \) be arbitrary. Let \( u = \{u_{ij}\}_{i,j=1}^\infty \) be such that \( u_{ij} \in L^2([0, T] : \mathbb{R}) \),
\[ \frac{1}{2} \int_0^T \sum_{i=1}^\infty \sum_{j=1, j \neq i}^\infty u_{ij}^2(s) ds \leq I(\eta) + \delta, \]
and \((\eta, u)\) satisfies (7.13). Define \(\psi : X_T \rightarrow \mathbb{R}\) by

\[
\psi(s, y) = \sum_{i=1}^{\infty} \sum_{j=1, j \neq i}^{\infty} 1_{A_{ij}(p(s))}(y) \frac{u_{ij}(s)}{\sqrt{p_i(s)\Gamma_{ij}(p(s))}} 1_{\{p_i(s)\Gamma_{ij}(p(s)) \neq 0\}}, \quad (s, y) \in [0, T] \times X.
\]

Then we have

\[
\int_{X_T} \psi^2(s, y) \lambda(dy) = \int_0^T \sum_{i=1}^{\infty} \sum_{j=1, j \neq i}^{\infty} u_{ij}^2(s) 1_{\{p_i(s)\Gamma_{ij}(p(s)) \neq 0\}} ds < \infty
\]

and hence \(\psi \in L^2(\lambda)\). Also note that

\[
\int_X G(q, y) 1_{A_{ij}(q)}(y) \lambda(dy) = (e_j - e_i) q_i \Gamma_{ij}(q)
\]

for \(q \in \hat{S}\) and \(i \neq j\). From this it follows that \((\eta, \psi)\) satisfies (7.11). Thus

\[
\bar{I}(\eta) \leq \frac{1}{2} \int_{X_T} \psi^2(s, y) \lambda(dy)
\]

\[
= \frac{1}{2} \int_0^T \sum_{i=1}^{\infty} \sum_{j=1, j \neq i}^{\infty} u_{ij}^2(s) 1_{\{p_i(s)\Gamma_{ij}(p(s)) \neq 0\}} ds
\]

\[
\leq I(\eta) + \delta.
\]

Since \(\delta > 0\) is arbitrary, we have that \(\bar{I}(\eta) \leq I(\eta)\).

Conversely, suppose \(\psi \in L^2(\lambda)\) is such that

\[
\frac{1}{2} \int_{X_T} \psi^2(s, y) \lambda(dy) \leq \bar{I}(\eta) + \delta
\]

and (7.11) holds. For \(i, j \in \mathbb{N}\) with \(i \neq j\) and \(s \in [0, T]\), define \(u_{ij} : [0, T] \rightarrow \mathbb{R}\) by

\[
u_{ij}(s) \approx \int_X 1_{A_{ij}(p(s))}(y) \psi(s, y) \lambda_X(dy) \sqrt{p_i(s)\Gamma_{ij}(p(s))} 1_{\{p_i(s)\Gamma_{ij}(p(s)) \neq 0\}}.
\]

An application of Cauchy–Schwarz inequality shows that

\[
u_{ij}^2(s) \leq \int_X 1_{A_{ij}(p(s))}(y) \psi^2(s, y) \lambda_X(dy),
\]

and hence \(u_{ij}^2 \in L^2(\lambda)\). Also note that

\[
\int_X G(q, y) 1_{A_{ij}(q)}(y) \lambda(dy) = (e_j - e_i) q_i \Gamma_{ij}(q)
\]

for \(q \in \hat{S}\) and \(i \neq j\). From this it follows that \((\eta, \psi)\) satisfies (7.11). Thus

\[
\bar{I}(\eta) \leq \frac{1}{2} \int_{X_T} \psi^2(s, y) \lambda(dy)
\]

\[
= \frac{1}{2} \int_0^T \sum_{i=1}^{\infty} \sum_{j=1, j \neq i}^{\infty} u_{ij}^2(s) 1_{\{p_i(s)\Gamma_{ij}(p(s)) \neq 0\}} ds
\]

\[
\leq I(\eta) + \delta.
\]

Since \(\delta > 0\) is arbitrary, we have that \(\bar{I}(\eta) \leq I(\eta)\).
and hence \( u_{ij} \in L^2([0,T] : \mathbb{R}) \) for all \( i \neq j \). We set \( u_{ii} = 0 \) for \( i \in \mathbb{N} \) and let \( u = \{u_{ij}\}_{i,j=1}^{\infty} \). It is easy to check that \((\eta,u)\) satisfies (7.13), and hence

\[
I(\eta) \leq \frac{1}{2} \int_0^T \sum_{i=1}^{\infty} \sum_{j=1,j \neq i}^{\infty} u_{ij}^2(s) \, ds
\]

\[
\leq \frac{1}{2} \int_{\mathcal{X}_T} \sum_{i=1}^{\infty} \sum_{j=1,j \neq i}^{\infty} 1_{A_{ij}(p(s))}(y) \psi^2(s,y) \lambda(\,ds \, dy) 
\]

\[
\leq \frac{1}{2} \int_{\mathcal{X}_T} \psi^2(s,y) \lambda(\,ds \, dy)
\]

\[
\leq \bar{I}(\eta) + \delta.
\]

Since \( \delta > 0 \) is arbitrary we have \( I(\eta) \leq \bar{I}(\eta) \). The result follows.

### 7.3 Proofs of Theorem 7.2 and Remark 7.2(iii)

#### Proof of Theorem 7.2

Part (a) can be established using a recursive construction of the solution from one jump to the next. Note that although \( \hat{E}n_m(\mathcal{X}_t) = \infty \) for all \( t > 0 \), the property that

\[
\lambda(\bigcup_{i=1}^{\infty} \bigcup_{j=1,j \neq i}^{\infty} A_{ij}(q)) \leq \sum_{i=1}^{\infty} \sum_{j=1,j \neq i}^{\infty} q_i \Gamma_{ij}(q) \leq \|\Gamma\|_{\infty} < \infty, \quad \forall q \in \hat{S},
\]

allows one to enumerate the jump instants \( t \) at which the state of \( \hat{\mu}^m(t) \) changes. At any such jump instant we define \( \hat{\mu}^m(t) \equiv \hat{\mu}^m(t-) + \frac{1}{m} G(\hat{\mu}^m(t-),y) \) if the jump corresponds to the point \((t,y)\) of the point process \( n_m \). We omit the details.

For part (b), uniqueness of solution of (7.9) follows from an application of Gronwall’s lemma along with Condition 7.1(b). For existence of solution we follow a standard iteration scheme. Define \( p^0(t) \equiv p(0) \) and \( p^{n+1}(t) \equiv p(0) + \int_0^t b(p^n(s)) \, ds \) for \( n \in \mathbb{N} \). From Condition 7.1(b),

\[
\|p^{n+1} - p^n\|_{s,t} = \left\| \int_0^t (b(p^n(s)) - b(p^{n-1}(s))) \, ds \right\|_{s,t} \leq L_b \int_0^t \|p^n - p^{n-1}\|_{s,s} \, ds,
\]

which implies that \( \{p^n\}_{n=0}^{\infty} \) is a Cauchy sequence in \( C([0,T] : L_2) \). Hence there exists some \( \tilde{p} \in C([0,T] : L_2) \) such that \( p^n \to \tilde{p} \) and it is easy to see that \( \tilde{p} \) is a solution to (7.9).
We now argue that \( \mu^m \Rightarrow p \) as \( m \to \infty \). For \( t \in [0,T] \),

\[
\mathbb{E} \sup_{s \leq t} \| \mu^m(s) - p(s) \|^2 \leq 3 \| \mu^m(0) - p(0) \|^2 + 3 \mathbb{E} \sup_{s \leq t} \left\| \int_0^s (b(\mu^m(u)) - b(p(u))) \, du \right\|^2 \\
+ 3 \mathbb{E} \sup_{s \leq t} \left\| \frac{1}{m} \int_{X_s} G(\mu^m(u), y) \tilde{N}^m(du \, dy) \right\|^2 \\
\leq 3 \| \mu^m(0) - p(0) \|^2 + 3T \int_0^t \mathbb{E} \| b(\mu^m(s)) - b(p(s)) \|^2 \, ds \\
+ \frac{12}{m} \int_{X_t} \mathbb{E} \| G(\mu^m(s), y) \|^2 \lambda(ds \, dy) \\
\leq \kappa \| \mu^m(0) - p(0) \|^2 + \frac{\kappa}{m} + \kappa \int_0^t \mathbb{E} \sup_{u \leq s} \| \mu^m(u) - p(u) \|^2 \, ds,
\]

where the second inequality follows from Doob’s inequality and the third inequality follows from Lemma 7.9. The result now follows from Gronwall’s inequality and Condition 7.1(c).

**Proof of Remark 7.2(iii)**

Suppose that for some \( n \in \mathbb{N} \), \( \sum_{m=1}^{\infty} [a(m)]^{2n} < \infty \). We need to show that \( a(m) \sqrt{m} \| \mu^m(0) - p(0) \| \to 0 \) almost surely. To simplify the notation, we will abbreviate \( \mu^m(0), p(0), \mu^m_i(0), p_i(0) \) as \( \mu, p, \mu_i, p_i \). It follows from Markov’s inequality that for \( \varepsilon > 0 \),

\[
\mathbb{P}(a(m) \sqrt{m} \| \mu^m - p \| > \varepsilon) \leq \left( \frac{a(m) \sqrt{m}}{\varepsilon} \right)^{2n} \mathbb{E} \| \mu^m - p \|^{2n} = \left( \frac{[a(m)]^{2n}}{\varepsilon^{2n} m^n} \mathbb{E} \sum_{i=1}^{\infty} (\mu_i^m - p_i)^2 \right)^n.
\]

Since \( \sum_{m=1}^{\infty} [a(m)]^{2n} < \infty \), by Borel-Cantelli lemma, it suffices to show that for every \( n \in \mathbb{N} \) there exists some \( \bar{\gamma}_n \in (0, \infty) \) such that

\[
\mathbb{E} \left[ \sum_{i=1}^{\infty} (\mu_i^m - p_i)^2 \right]^n \leq \frac{\bar{\gamma}_n}{m^n}, \quad (7.30)
\]

We will prove (7.30) when \( n = 2 \) in detail and then sketch the argument for \( n > 2 \). First write

\[
\mu_i^m - p_i = \frac{1}{m} \sum_{j=1}^{m} 1_{\{\xi_j = i\}} - p_i = \frac{1}{m} \sum_{j=1}^{m} Y_{ij},
\]

where \( Y_{ij} \equiv 1_{\{\xi_j = i\}} - p_i \). Note that for all \( \alpha, \beta \in \mathbb{N} \),

\[
|Y_{ij}| \leq 1, \quad \mathbb{E} Y_{ij} = 0, \quad \mathbb{E} Y_{ij}^{2\alpha} \leq \mathbb{E} Y_{ij}^2 \leq p_i, \quad \mathbb{E} Y_{ij}^{2\alpha} Y_{kl}^{2\beta} \leq \mathbb{E} Y_{ij}^2 Y_{kl}^2 \leq p_i p_k \text{ for } i \neq k.
\]
So we have

\[
E \left[ \sum_{i=1}^{\infty} (\mu_i^m - p_i)^2 \right]^2 = \frac{1}{m^4} E \left[ \sum_{i=1}^{\infty} \sum_{j,j' = 1}^{m} Y_{ij} Y_{ij'} \right]^2 = \frac{1}{m^4} \sum_{i,k=1}^{m} \sum_{j,j',l,l' = 1}. Y_{ij} Y_{ij'} Y_{kl} Y_{kl'}.
\]  

(7.31)

From independence of \{\xi_j\} it follows that \(Y_{ij}\) and \(Y_{kl}\) are independent for \(j \neq l\). Hence we have \(EY_{ij}Y_{ij'}Y_{kl}Y_{kl'} \neq 0\) only if \(j, j', l, l'\) are matched in pairs (e.g. \(j = j'\) and \(l = l'\)). Using this observation, (7.31) can be written as

\[
\frac{1}{m^4} \sum_{r=1}^{5} \sum_{(i,k,j,j',l,l') \in \mathcal{H}_r} Y_{ij} Y_{ij'} Y_{kl} Y_{kl'} = \sum_{r=1}^{5} T_r^m,
\]

where \(\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4\) and \(\mathcal{H}_5\) are collections of \((i, k) \in \mathbb{N}^2\) and \((j, j', l, l') \in \{1, \ldots, m\}^4\) such that \(\{j = j' \neq l = l', \{j = l \neq j' = l', \{j = j' = l = l', i = k\}\) and \(\{j = j' = l = l', i \neq k\}\), respectively. For \(T_1^m\), it follows from independence of \(\xi_j\) that

\[
T_1^m = \frac{1}{m^4} E \sum_{i,k=1}^{m} \sum_{j,j'=1}^{m} Y_{ij}^2 Y_{kl} = \frac{1}{m^4} \sum_{i,k=1}^{m} \sum_{j,j'=1}^{m} \frac{EY_{ij}^2 EY_{kl}^2}{m^2} \leq \frac{1}{m^2} \sum_{i,k=1}^{m} p_i p_k = \frac{1}{m^2}.
\]

For \(T_2^m\), using independence of \{\xi_j\} and Cauchy–Schwarz inequality we have

\[
T_2^m = \frac{1}{m^4} E \sum_{i,k=1}^{m} \sum_{j,j'=1}^{m} Y_{ij} Y_{ij'} Y_{kj} Y_{kl'} = \frac{1}{m^4} \sum_{i,k=1}^{m} \sum_{j,j'=1}^{m} E(Y_{ij} Y_{kj}) E(Y_{ij'} Y_{kl'}) \leq \frac{1}{m^2} \sum_{i,k=1}^{m} p_i p_k = \frac{1}{m^2}.
\]

Similarly, \(T_3^m \leq \frac{1}{m^2}\). For \(T_4^m\) and \(T_5^m\), we have

\[
T_4^m = \frac{1}{m^4} E \sum_{i=1}^{\infty} \sum_{j=1}^{m} Y_{ij}^4 \leq \frac{1}{m^3} \sum_{i=1}^{\infty} p_i = \frac{1}{m^3},
\]

\[
T_5^m = \frac{1}{m^4} E \sum_{i,k=1}^{m} \sum_{j=1}^{m} \sum_{i \neq k} Y_{ij}^2 Y_{kj}^2 \leq \frac{1}{m^3} \sum_{i,k=1}^{\infty} \sum_{i \neq k} p_i p_k \leq \frac{1}{m^3}.
\]

Combining above estimates we can bound (7.31) by \(\frac{3}{m^2} + \frac{2}{m^2}\). This proves (7.30) when \(n = 2\).
For the case $n > 2$, write

$$
E \left[ \sum_{i=1}^{\infty} (\mu_i^m - p_i)^2 \right]^n = E \left[ \sum_{i=1}^{\infty} \left( \frac{1}{m} \sum_{j=1}^{m} Y_{ij} \right)^2 \right]^n = \frac{1}{m^{2n}} E \left[ \sum_{i=1}^{\infty} \sum_{j=1}^{m} \sum_{k=1}^{m} Y_{ij} Y_{ik} \right]^n
$$

$$
= \frac{1}{m^{2n}} \sum_{i_1, \ldots, i_n = 1}^{\infty} \sum_{j_1, k_1, \ldots, j_n, k_n = 1}^{m} Y_{i_1j_1} Y_{i_1k_1} \cdots Y_{i_nj_n} Y_{i nk_n}, \quad (7.32)
$$

Once again $E (Y_{i_1j_1} Y_{i_1k_1} \cdots Y_{i_nj_n} Y_{i nk_n}) \neq 0$ only if $j_1, k_1, \ldots, j_n, k_n$ are matched in pairs. Hence the $2n$-fold summation over $j_1, k_1, \ldots, j_n, k_n$ in (7.32) can be reduced to no more than an $n$-fold sum. We can break up the outer sum into $n$ terms where the $M$-th term, $M = 1, \ldots, n$, corresponds to indices $(i_1, \ldots, i_n)$ of which exactly $M$ indices are distinct. Similarly the inner sum can be split into $n$ terms where the $N$-th term, $N = 1, \ldots, n$, corresponds to indices $(j_1, k_1, \ldots, j_n, k_n)$ matched in pairs with exactly $N$ distinct pairs. Furthermore each such $(M, N)$-term can be split into a finite number of terms, each of which corresponds to a collection $\{c_{\alpha\beta}, \alpha = 1, \ldots, M, \beta = 1, \ldots, N\}$ of non-negative integers representing how $\{Y_{ij}\}$ is paired up, with $\sum_{\alpha=1}^{M} c_{\alpha\beta} \geq 1$, $\sum_{\beta=1}^{N} c_{\alpha\beta} \geq 1$ and $\sum_{\alpha=1}^{M} \sum_{\beta=1}^{N} c_{\alpha\beta} = 2n$. By independence of $\{\xi_j\}$, the contribution of each such $(M, N, \{c_{\alpha\beta}\})$-term to (7.32) is at most

$$
\frac{\kappa_n}{m^{2n}} \sum_{i_1, \ldots, i_M = 1}^{\infty} \sum_{j_1, \ldots, j_N = 1}^{m} E \left( Y_{i_1j_1}^{c_{11}} Y_{i_2j_1}^{c_{21}} \cdots Y_{i_Mj_1}^{c_{M1}} \right) \cdots E \left( Y_{i_1j_N}^{c_{1N}} Y_{i_2j_N}^{c_{2N}} \cdots Y_{i_Mj_N}^{c_{MN}} \right), \quad (7.33)
$$

where $\kappa_n \in (0, \infty)$ only depends on $n$. A simple calculation gives that for all $\beta = 1, \ldots, N$,

$$
\left| E \left( Y_{i_1j_\beta}^{c_{1\beta}} Y_{i_2j_\beta}^{c_{2\beta}} \cdots Y_{i_Mj_\beta}^{c_{M\beta}} \right) \right| \leq \tilde{\kappa}_n P_{i_1}^{c_{1\beta} \wedge 1} \cdots P_{i_M}^{c_{M\beta} \wedge 1},
$$

where $\tilde{\kappa}_n \in (0, \infty)$ only depends on $n$. Hence (7.33) is bounded by $\frac{\kappa_n \tilde{\kappa}_n^N}{m^{2n-N}} \leq \frac{\kappa_n \tilde{\gamma}_n^N}{m^N}$. So (7.32) is bounded by $\frac{\tilde{\gamma}_n}{m^\gamma}$ for some $\tilde{\gamma}_n \in (0, \infty)$, which gives (7.30) and completes the proof. \qed
CHAPTER 8
LARGE DEVIATIONS FOR PARTICLE APPROXIMATION OF A NONLINEAR HEAT EQUATION BY BROWNIAN MOTIONS WITH KILLING

8.1 Introduction

The goal of this chapter is to study a large deviation principle (LDP) for an interacting particle system associated with the non-linear (and non-local) heat equation

\[
\frac{\partial u}{\partial t} = \frac{1}{2} \Delta u - \langle \zeta, u \rangle u, \tag{8.1}
\]

where \(\Delta\) is the \(d\)-dimensional Laplacian operator and \(\zeta : \mathbb{R}^d \to \mathbb{R}_+\) is a continuous and bounded function that is bounded away from 0. Roughly speaking the system is described by \(n\) independent Brownian particles where each particle is killed independently at a rate determined by the empirical measure of current particle states. Let \(\{X_i\}_{i \geq 1}\) be a sequence of i.i.d. exponential random variables with rate 1 and let \(\{B_i(t), t \geq 0\}_{i \geq 1}\) be independent \(d\)-dimensional standard Brownian motions independent of \(\{X_i\}_{i \geq 1}\). Define for \(t \geq 0\) the random sub-probability measure \(\mu^n(t)\) as the solution to the following equation

\[
\mu^n(t) = \frac{1}{n} \sum_{i=1}^{n} \delta_{B_i(t)} \mathbf{1}_{\{X_i > \int_0^t \langle \zeta, \mu^n(s) \rangle \, ds\}}. \tag{8.2}
\]

Since a.s., we can enumerate \(\{X_i\}_{i=1}^{n}\) in a strictly increasing order, the unique solution of (8.2) can be written explicitly in a recursive manner. We are interested in an LDP for the stochastic process \(\mu^n_t \doteq \{\mu^n(t)\}_{t \in [0, T]}\) in \(\mathcal{D}([0, T] : \mathcal{M}(\mathbb{R}^d))\) where \(\mathcal{M}(\mathbb{R}^d)\) is the space of sub probability measures on \(\mathbb{R}^d\) with the usual weak convergence topology.

The chapter is organized as follows. Section 8.2 introduces the basic model and notation. In this section we also state a law of large numbers result for \(\mu^n\), a proof of which is sketched in Section 8.9. Section 8.3 presents the main large deviations result of this chapter, namely Theorem 8.3. Rest of this chapter gives the proof of Theorem 8.3. We begin in Section 8.4 by establishing a variational representation for functionals of independent Brownian motions and an i.i.d. sequence of random
variables. Tightness of controls and related processes in this representation is argued in Section 8.5. Section 8.6 proves the Laplace principle upper bound for the LDP in Theorem 8.3 while Section 8.7 gives the proof of the lower bound. Compactness of sub-level sets of the candidate rate function is established in Section 8.8. The large deviation principle for \(\mu^n_T\) is immediate on combining results of Sections 8.6, 8.7 and 8.8.

### 8.2 Model and notation

The following notation will be used. For a Polish space \(S\), let \(\mathcal{M}(S)\) denote the collection of sub-probability measures on \(S\) with the usual weak convergence topology. For measurable \(f: S \to \mathbb{R}^d\) and \(\gamma \in \mathcal{M}(S)\), let \(\langle f, \gamma \rangle = \int_S f \, d\gamma\) whenever the integral is well defined. For \(\gamma_1, \gamma_2 \in \mathcal{P}(S)\), we write \(\gamma_1 \ll \gamma_2\) if \(\gamma_1\) is absolutely continuous with respect to \(\gamma_2\). We will denote \(C = C([0, \infty) : \mathbb{R}^d)\), \(S = C \times \mathbb{R}_+\) and \(D = D([0, \infty) : \mathcal{M}(\mathbb{R}^d))\). Spaces \(C_T, S_T, D_T\) are defined similarly by replacing \([0, \infty)\) with \([0, T]\).

Let \(\{(B_t, X_t)\}_{t \in \mathbb{N}}\) be a sequence of independent \(d\)-dimensional standard Brownian motions and exponential random variables defined on \((\Omega, \mathcal{F}, P)\). Let \(\mathcal{F}_t\) be the \(P\)-completion of the \(\sigma\)-field generated by \(\{B_t(s), X_i, s \leq t, i \in \mathbb{N}\}\). Fix \(\zeta \in \mathcal{C}_b(\mathbb{R}^d)\) with \(\inf_{x \in \mathbb{R}^d} \zeta(x) = \zeta > 0\). Let \(\theta \overset{\text{d}}{=} \mathcal{L}(X_1)\) be the exponential distribution (with rate 1). The main result of this chapter is a large deviation principle (LDP) for \(\mu^n_T = \{\mu^n(t)\}_{t \in [0, T]}\) in \(D_T\) for fixed \(T < \infty\). We begin by first establishing a law of large numbers result for \(\mu^n = \{\mu^n(t)\}_{t \geq 0}\).

Let \(\mu_0 = \mathcal{L}(B_1)\) be the Wiener measure on \(\mathcal{C}\), \(\mu_{0, t} = \mathcal{L}(B_1(t))\) be the marginal distribution of \(\mu_0\) at time instant \(t\), and \(b(t) = \langle \zeta, \mu_{0, t} \rangle\). Let \(\mu_1 \in \mathcal{P}([0, \infty])\) be such that \(\mu_1([t, \infty)) = a(t), t \geq 0\), where \(a(t)\) is the unique solution of the ordinary differential equation (ODE)

\[
\dot{a}(t) = -a^2(t)b(t), \quad a(0) = 1. \tag{8.3}
\]

Note that since \(\zeta > 0\), one has \(\mu_1(\{\infty\}) = 0\). Indeed \(a(t) \leq \frac{1}{1 + \overline{\zeta}}\) for all \(t \geq 0\). Define \(\mu: [0, \infty) \to \mathcal{M}(\mathbb{R}^d)\) as \(\mu(t) = a(t)\mu_{0, t}\). Clearly, \(\mu \in \mathcal{C}([0, \infty) : \mathcal{M}(\mathbb{R}^d))\) and \(\mu(t)\) is absolutely continuous with respect to Lebesgue measure. Writing \(\mu_t(dx) = u(t, x)dx\), it is easily checked that \(u(t, x)\) is the solution to (8.1). Further we have the following law of large numbers. A sketch of proof is provided in Section 8.9.

**Theorem 8.1.** As \(n \to \infty\), \(\mu^n\) converges to \(\mu\) in probability in \(D\).
8.3 Large deviation principle

Fix $T > 0$. The goal of this chapter is to establish a LDP for $\{\mu^n_T\}$ in $\mathcal{D}_T$. For notational simplicity, we omit $T$ from the notation $C_T, \mathcal{D}_T, S_T, \mu^n_T$, etc.

8.3.1 Canonical space and processes

Let $\mathcal{R}^W$ be the space of finite measures $r$ on $\mathcal{B}(\mathbb{R}^d \times [0, T])$ such that $r(\mathbb{R}^d \times [0, t]) = t$ for all $t \in [0, T]$ and

$$\int_{\mathbb{R}^d \times [0, T]} \|y\| r(dydt) < \infty.$$ 

Such a measure can be disintegrated as $r(dydt) = r_t(dy) dt$, where $t \mapsto r_t(\cdot)$ is a measurable map from $[0, T]$ to $\mathcal{P}(\mathbb{R}^d)$. We equip $\mathcal{R}^W$ with the topology of weak convergence plus convergence of first moments. This topology can be metrized with the Wasserstein-1 distance and the space with this metric is Polish (cf. [60], Section 6.3).

We now introduce the canonical space $\Xi = \mathcal{C} \times \mathcal{C} \times \mathcal{R}^W \times \mathbb{R}_+$ and canonical variables on $(\Xi, \mathcal{B}(\Xi))$: For $\xi = (\tilde{b}, b, \rho, \tilde{\sigma}) \in \Xi$,

$$\tilde{b}(\xi) \doteq \tilde{b}, \quad b(\xi) \doteq b, \quad \rho(\xi) \doteq \rho, \quad \tilde{\sigma}(\xi) \doteq \tilde{\sigma}.$$ 

Let $\{\mathcal{G}_t\}_{0 \leq t \leq T}$ be the canonical filtration on $(\Xi, \mathcal{B}(\Xi))$, namely

$$\mathcal{G}_t \doteq \sigma(\tilde{b}_s, b_s, \rho(A \times [0, s]), s \leq t, A \in \mathcal{B}(\mathbb{R}^d), \tilde{\sigma}), \quad t \in [0, T].$$ 

For $\Theta \in \mathcal{P}(\Xi)$, denote by $\Theta_{(i)}$ the marginal of $\Theta$ on the $i$–th coordinate, $i = 1, 2, 3, 4$.

8.3.2 Rate function and statement of the LDP

Let $\mathcal{P}_\infty$ denote the collection of all $\Theta \in \mathcal{P}(\Xi)$ such that, under $\Theta$ the following hold:

1. $\mathbf{b}$ is a $d$-dimensional standard $\mathcal{G}_t$-Brownian motion.
2. $\int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dydt) < \infty$, a.s.
3. $\tilde{b}_t = b_t + \int_{\mathbb{R}^d \times [0, t]} y \rho(dyds)$, for all $t \in [0, T]$, a.s.
4. $\Theta_{(4)} \ll \theta$.

Remark 8.1. Note that property (4) above implies that for $\Theta \in \mathcal{P}_\infty$, $\Theta_{(4)}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}_+$. This observation will be used in the proof of Lemmas 8.2, 8.10 and 8.11.
For $\Theta \in \mathcal{P}(\Xi)$ let
\[
J(\Theta) = \mathbf{E}^\Theta \left[ \frac{1}{2} \int_{\mathbb{R}^d \times [0,T]} \| y \|^2 \mathbf{1}_{\tilde{\sigma} > f_0(\zeta, \omega_\Theta(u))} \, dy \, dt \right] + R(\Theta(4), \| \theta \|),
\]
where $R(\gamma \| \theta)$ for $\gamma \in \mathcal{P}(\mathbb{R}^d)$ is the relative entropy defined as
\[
R(\gamma \| \theta) = \begin{cases} 
\int_{\mathbb{R}^d} \left( \log \frac{d\gamma}{d\theta} \right) \, d\gamma, & \gamma \ll \theta, \\
\infty, & \text{otherwise}.
\end{cases}
\]

For $\Theta \in \mathcal{P}_\infty$ with $J(\Theta) < \infty$, let $\omega_\Theta \in \mathcal{D}$ be the unique solution (which is guaranteed by Lemma 8.2 below) to the following equation:
\[
\langle f, \omega_\Theta(t) \rangle = \mathbf{E}^\Theta \left[ f(\tilde{b}_t) \mathbf{1}_{\tilde{\sigma} > f_0(\zeta, \omega_\Theta(s))} \right], \quad \forall f \in \mathcal{C}_b(\mathbb{R}^d), t \in [0,T]. \quad (8.4)
\]

Define function $I : \mathcal{D} \to [0, \infty]$ as
\[
I(\pi) = \inf_{\Theta \in \mathcal{P}_\infty : \omega_\Theta = \pi} \{ J(\Theta) \}, \quad \pi \in \mathcal{D}. \quad (8.5)
\]

**Lemma 8.2.** Let $\Theta \in \mathcal{P}_\infty$ satisfy $J(\Theta) < \infty$. Then there exists a unique solution $\omega_\Theta$ of (8.4). Moreover $\omega_\Theta \in \mathcal{C}([0,T] : \mathcal{M}(\mathbb{R}^d))$.

**Proof.** The existence of $\omega_\Theta$ follows from the argument in Section 8.7 (see arguments below (8.47)). For any solution $\omega_\Theta$ of (8.4), it follows from Property (4) of $\mathcal{P}_\infty$ and Remark 8.1 that for each $f \in \mathcal{C}_b(\mathbb{R}^d)$ and $t, s \in [0,T]$, as $t \to s$,
\[
f(\tilde{b}_t) \mathbf{1}_{\tilde{\sigma} > f_0(\zeta, \omega_\Theta(u))} \to f(\tilde{b}_s) \mathbf{1}_{\tilde{\sigma} > f_0(\zeta, \omega_\Theta(u))}, \quad \text{a.s. } \Theta,
\]
and hence $\langle f, \omega_\Theta(t) \rangle \to \langle f, \omega_\Theta(s) \rangle$. This implies that $\omega_\Theta \in \mathcal{C}([0,T] : \mathcal{M}(\mathbb{R}^d))$. Now it remains to argue the uniqueness of $\omega_\Theta$. Let $\pi_1$ and $\pi_2$ be two solutions of (8.4) in $\mathcal{C}([0,T] : \mathcal{M}(\mathbb{R}^d))$. Consider functions $h(t) \doteq \langle \zeta, \pi_1(t) \rangle - \langle \zeta, \pi_2(t) \rangle$ and $H(t) \doteq \int_0^t h(s) \, ds$ for $t \in [0,T]$. We will argue via contradiction that $H(t) \equiv 0$. This will show $\pi_1 = \pi_2$, proving the desired uniqueness. Suppose $M \doteq \sup_{0 \leq t \leq T} |H(t)| > 0$. Without loss of generality we assume that $M = \sup_{0 \leq t \leq T} H(t)$. Since
$H$ is a continuous function, there exists some $t \in [0, T]$ such that $H(t) = M$. Define $t^* = \inf\{t : H(t) = M\}$. Then $H(t^*) = M$. Since $H(0) = 0$, we have $t^* > 0$. Again by continuity of $H$, there exists $t_0 \in (0, t^*)$ such that for all $t \in [t_0, t^*], H(t) \geq M/2 > 0$, and hence

$$\int_0^t \langle \zeta, \pi_1(s) \rangle \, ds > \int_0^t \langle \zeta, \pi_2(s) \rangle \, ds, \forall t \in [t_0, t^*].$$

Noting that $\zeta \geq 0$, we have for all $t \in [t_0, t^*],$

$$h(t) = \langle \zeta, \pi_1(t) \rangle - \langle \zeta, \pi_2(t) \rangle = \mathbb{E}^\Theta \left[ \zeta(\tilde{b}_t) \left( 1_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \pi_1(s) \rangle \, ds \}} - 1_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \pi_2(s) \rangle \, ds \}} \right) \right] \leq 0.$$ 

This implies that

$$M = H(t^*) = H(t_0) + \int_{t_0}^{t^*} h(s) \, ds \leq H(t_0),$$

which contradicts with the definition of $t^*$. Hence we must have $H(t) \equiv 0$ and this completes the proof.

The following is the main result of this chapter.

**Theorem 8.3.** The sequence $\{\mu^n\}_{n \in \mathbb{N}}$ satisfies a LDP in $\mathcal{D}$ with rate function $I$.

In order to prove Theorem 8.3 it suffices to show that:

1. $I$ defined in (8.5) is a rate function.
2. For every $F \in \mathcal{C}_b(\mathcal{D}),$

$$\lim_{n \to \infty} -\frac{1}{n} \log \mathbb{E} \left[ \exp \left( -nF(\mu^n) \right) \right] = \inf_{\pi \in \mathcal{D}} \{ F(\pi) + I(\pi) \}. \quad (8.6)$$

Proof of item (1) is given in Section 8.8 while the proof of item (2) is carried out in two steps. First in Section 8.6 we will prove the **Laplace upper bound**:

$$\liminf_{n \to \infty} -\frac{1}{n} \log \mathbb{E} \left[ \exp \left( -nF(\mu^n) \right) \right] \geq \inf_{\pi \in \mathcal{D}} \{ F(\pi) + I(\pi) \}. \quad (8.7)$$

The proof of (8.6) is then completed in Section 8.7 by proving the complementary **Laplace lower bound**:

$$\limsup_{n \to \infty} -\frac{1}{n} \log \mathbb{E} \left[ \exp \left( -nF(\mu^n) \right) \right] \leq \inf_{\pi \in \mathcal{D}} \{ F(\pi) + I(\pi) \}. \quad (8.8)$$

147
8.4 A general variational representation formula

In order to prove Theorem 8.3 we need to study the asymptotics of

\[- \frac{1}{n} \log \mathbb{E} \left[ \exp \left( - nF(\mu^n) \right) \right], \quad (8.9)\]

where \( F \in \mathbb{C}_b(D) \). For this we will use certain variational representations. Note that \( F(\mu^n) \) can be written as \( \Psi(B^n, X^n) \), where \( B^n \doteq \{B_i\}_{i=1}^n \) is an \( nd \)-dimensional Brownian motion, \( X^n \doteq \{X_i\}_{i=1}^n \) is a \((\mathbb{R}_+)^n\)-valued random variable and \( \Psi \) is a suitable map. When \( \Psi \) is just a function of the Brownian motion \( B^n \), a variational representation for quantities as in (8.9) was obtained in [8] (see also [10] where a more convenient form that allows for an arbitrary filtration is given). In this section, we will establish an extension of this result that gives a variational representation for positive functionals of both \( B^n \) and \( X^n \).

Throughout this section let \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})\) be a \( \mathbb{P} \)-complete filtered probability space on which are given a \( d \)-dimensional standard \( \mathcal{F}_t \)-Brownian motion \( B \) and an \( \mathcal{F}_0 \)-measurable random variable \( X \), which takes values in a Polish space \( S \) with law \( \rho \). Note that since \( X \) is \( \mathcal{F}_0 \)-measurable, \( B \) and \( X \) are independent. We will now establish a convenient variational representation for

\[- \log \mathbb{E} \left[ \exp \left( - f(B, X) \right) \right], \quad f \in \mathbb{M}_b(C \times S) \text{ and } C \doteq C([0, T] : \mathbb{R}^d). \]

Consider the probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})\), where \( \tilde{\Omega} = \mathcal{C} \), \( \tilde{\mathcal{F}} = \mathcal{B}(\mathcal{C}) \) and \( \tilde{\mathbb{P}} \) is the \( d \)-dimensional Wiener measure. Under \( \tilde{\mathbb{P}} \) the coordinate mapping process \( \tilde{W} \doteq \{\tilde{W}_t(\tilde{\omega}) = \tilde{\omega}(t), 0 \leq t \leq T \} \) is a standard \( d \)-dimensional Brownian motion with respect to the filtration \( \{\tilde{\mathcal{F}}^W_t\} \doteq \{\sigma(\tilde{W}(s) : s \leq t)\} \). Let \( \{\tilde{\mathcal{F}}_t\} \) be the augmented filtration, namely \( \tilde{\mathcal{F}}_t \doteq \sigma(\tilde{\mathcal{F}}^W_t \cup \tilde{\mathcal{N}}) \) and \( \tilde{\mathcal{N}} \) is the collection of all \( \tilde{\mathbb{P}} \)-null sets.

For \( f \in \mathbb{M}_b(C \times S) \), define

\[ \tilde{f}(x) \doteq - \log \mathbb{E} \left[ \exp \left( - f(\tilde{W}, x) \right) \right], \quad x \in S. \quad (8.10) \]

It follows from the independence between \( B \) and \( X \) that

\[- \log \mathbb{E} \left[ \exp \left( - f(B, X) \right) \right] = - \log \mathbb{E} \left[ \exp \left( - \tilde{f}(X) \right) \right]. \]
From classical results of Donsker–Varadhan (cf. [27], Proposition 1.4.2) we have the following representation formula from the above equality

$$\log \mathbb{E} \left[ \exp \left( -f(B, X) \right) \right] = \inf_{\Pi \in \mathcal{P}(\mathcal{S})} \left[ R(\Pi \| \rho) + \int_{\mathcal{S}} \tilde{f}(x) \Pi(dx) \right].$$  \hspace{1cm} (8.11)

Consider the collection of processes

$$\tilde{A} \doteq \{ \tilde{\psi} : \text{the process } \tilde{\psi}(s, \tilde{\omega}) \text{ is } \tilde{\mathcal{F}}_t\text{-progressively measurable and } \mathbb{E} \int_0^T \|\tilde{\psi}(s)\|^2 \, ds < \infty \}. $$

From Theorem 3.1 in [8] we now have the following variational formula

$$\tilde{f}(x) = \inf_{\tilde{\psi} \in \tilde{A}} \mathbb{E} \left[ \frac{1}{2} \int_0^T \|\tilde{\psi}(s)\|^2 \, ds + f \left( \tilde{W} + \int_0^T \tilde{\psi}(s) \, ds, x \right) \right],$$  \hspace{1cm} (8.12)

which together with (8.11) gives

$$- \log \mathbb{E} \left[ \exp \left( -f(B, X) \right) \right] = \inf_{\Pi \in \mathcal{P}(\mathcal{S})} \left\{ R(\Pi \| \rho) + \int_{\mathcal{S}} \mathbb{E} \left[ \frac{1}{2} \int_0^T \|\tilde{\psi}(s)\|^2 \, ds + f \left( \tilde{W} + \int_0^T \tilde{\psi}(s) \, ds, x \right) \right] \Pi(dx) \right\}. $$ \hspace{1cm} (8.13)

We will now give an equivalent variational representation that is simpler than (8.13) and more convenient to use. Let \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})\) be a probability space on which are given a \(d\)-dimensional standard Brownian motion \(\tilde{W}\) and an \(\mathcal{S}\)-valued random variable \(\tilde{X}\), which is independent of \(\tilde{W}\), with law \(\Pi\).

Let \(\{\tilde{\mathcal{F}}_t \equiv \sigma(\tilde{\mathcal{F}}^{W, X}_t \cup \mathcal{N})\}\) be the augmented filtration, where \(\mathcal{N}\) is the collection of all \(\mathbb{P}\)-null sets and \(\mathcal{N}^{W, X}_t \equiv \sigma(\tilde{X}, \tilde{W}(s) : s \leq t)\). Consider the following collection of processes

$$\mathcal{A} \doteq \{ \tilde{\psi} : \text{the process } \tilde{\psi}(s, \tilde{\omega}) \text{ is } \mathcal{F}_t\text{-progressively measurable and } \mathbb{E} \int_0^T \|\tilde{\psi}(s)\|^2 \, ds < \infty \}. $$

For each \(N < \infty\), let

$$\mathcal{A}_N \doteq \left\{ \tilde{\psi} \in \mathcal{A} : \int_0^T \|\tilde{\psi}(s)\|^2 \, ds \leq N, \mathbb{P}\text{-a.s.} \right\}. $$ \hspace{1cm} (8.14)

For \(\Pi \in \mathcal{P}(\mathcal{S})\), \(\mathcal{A} \subset \mathcal{A}\) and \(g \in \mathcal{M}_b(\mathbb{C} \times \mathcal{S})\), define

$$\Lambda_{\Pi}(\mathcal{A}, g) \doteq \inf_{\tilde{\psi} \in \mathcal{A}} \mathbb{E} \left[ \frac{1}{2} \int_0^T \|\tilde{\psi}(s)\|^2 \, ds + g \left( \tilde{W} + \int_0^T \tilde{\psi}(s) \, ds, \tilde{X} \right) \right]. $$ \hspace{1cm} (8.15)
The following lemma will be used in establishing a simpler representation formula. The proof is very similar to that of Lemma 3.4 in [10] (see also proof of Theorem 3.1 in [8]).

**Lemma 8.4.** Let \( \{f_n\} \) be a uniformly bounded sequence of real-valued measurable functions on \( C \times S \) converging to \( f \) a.s. \( \tilde{P} \times \Pi \). Then for every \( N < \infty \), \( \Lambda_\Pi(\tilde{A}^N, f_n) \to \Lambda_\Pi(\tilde{A}^N, f) \) as \( n \to \infty \).

**Proof.** Fix \( \varepsilon > 0 \). For each \( n \in \mathbb{N} \) pick an \( \varepsilon \)-optimal element \( \tilde{\psi}_{n,\varepsilon} \in \tilde{A}^N \) such that

\[
\Lambda_\Pi(\tilde{A}^N, f_n) \geq \mathbb{E} \left[ \frac{1}{2} \int_0^T \|\tilde{\psi}_{n,\varepsilon}(s)\|^2 \, ds + f_n \left( \tilde{W} + \int_0^\cdot \tilde{\psi}_{n,\varepsilon}(s) \, ds, \tilde{X} \right) \right] - \varepsilon.
\]

By definition, for each \( n \in \mathbb{N} \)

\[
\mathbb{E} \left[ \frac{1}{2} \int_0^T \|\tilde{\psi}_{n,\varepsilon}(s)\|^2 \, ds + f \left( \tilde{W} + \int_0^\cdot \tilde{\psi}_{n,\varepsilon}(s) \, ds, \tilde{X} \right) \right] \geq \Lambda_\Pi(\tilde{A}^N, f).
\]

We claim that as \( n \to \infty \),

\[
\mathbb{E} f_n \left( \tilde{W} + \int_0^\cdot \tilde{\psi}_{n,\varepsilon}(s) \, ds, \tilde{X} \right) - \mathbb{E} f \left( \tilde{W} + \int_0^\cdot \tilde{\psi}_{n,\varepsilon}(s) \, ds, \tilde{X} \right) \to 0.
\]

Once the claim is verified, combining above three displays gives us \( \limsup_{n \to \infty} \Lambda_\Pi(\tilde{A}^N, f_n) \geq \Lambda_\Pi(\tilde{A}^N, f) - \varepsilon \). Since \( \varepsilon > 0 \) is arbitrary, this shows \( \limsup_{n \to \infty} \Lambda_\Pi(\tilde{A}^N, f_n) \geq \Lambda_\Pi(\tilde{A}^N, f) \). In order to prove the claim, from Lemma 2.8(b) in [8], it suffices to show that the relative entropies

\[
R \left( \tilde{P} \circ \left( \tilde{W} + \int_0^\cdot \tilde{\psi}_{n,\varepsilon}(s) \, ds, \tilde{X} \right)^{-1} \, \left\| \tilde{P} \times \Pi \right\| \right)
\]

are uniformly bounded in \( n \). For this, first consider the probability measure \( \tilde{P}^{n,\varepsilon} \) defined by

\[
\frac{d\tilde{P}^{n,\varepsilon}}{d\tilde{P}} = \exp \left\{ - \int_0^T \tilde{\psi}_{n,\varepsilon}(s) \, d\tilde{W}(s) - \frac{1}{2} \int_0^T \|\tilde{\psi}_{n,\varepsilon}(s)\|^2 \, ds \right\}.
\]

By Girsanov’s theorem, on the probability space \( (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}^{n,\varepsilon}) \), \( \tilde{W} + \int_0^\cdot \tilde{\psi}_{n,\varepsilon}(s) \, ds \) is an \( \{\tilde{\mathcal{F}}_t\} \)-Brownian motion independent of the \( \tilde{\mathcal{F}}_0 \)-measurable random variable \( \tilde{X} \), and \( \tilde{P}^{n,\varepsilon} \circ (\tilde{X})^{-1} = \tilde{P} \times \Pi \). So we have

\[
\tilde{P}^{n,\varepsilon} \circ \left( \tilde{W} + \int_0^\cdot \tilde{\psi}_{n,\varepsilon}(s) \, ds, \tilde{X} \right)^{-1} = \tilde{P} \times \Pi
\]
and hence
\[
R\left( \bar{P} \circ \left( \bar{W} + \int_{0}^{T} \bar{\psi}_{n,\varepsilon}(s) \, ds, \bar{X} \right) \right)^{-1} \leq R(\bar{P} \| \bar{P}^{n,\varepsilon}) = \bar{E} \left[ \frac{1}{2} \int_{0}^{T} \| \bar{\psi}_{n,\varepsilon}(s) \|^2 \, ds \right] \leq \frac{N}{2} < \infty.
\]

Thus the claim holds and we have \( \liminf_{n \to \infty} \Lambda_{\Pi}(\bar{A}_{N}, f_n) \geq \Lambda_{\Pi}(\bar{A}_{N}, f) \).

For the reverse inequality, pick an \( \varepsilon \)-optimal element \( \bar{\psi}_{\varepsilon} \in \bar{A}_{N} \) such that
\[
\bar{E} \left[ \frac{1}{2} \int_{0}^{T} \| \bar{\psi}_{\varepsilon}(s) \|^2 \, ds + f \left( \bar{W} + \int_{0}^{T} \bar{\psi}_{\varepsilon}(s) \, ds, \bar{X} \right) \right] \leq \Lambda_{\Pi}(\bar{A}_{N}, f) + \varepsilon.
\]
Since \( \bar{P} \circ (\bar{W} + \int_{0}^{T} \bar{\psi}_{\varepsilon}(s) \, ds, \bar{X})^{-1} \ll \bar{P} \times \Pi \), we have that as \( n \to \infty \),
\[
\bar{E} f_n \left( \bar{W} + \int_{0}^{T} \bar{\psi}_{\varepsilon}(s) \, ds, \bar{X} \right) \to \bar{E} f \left( \bar{W} + \int_{0}^{T} \bar{\psi}_{\varepsilon}(s) \, ds, \bar{X} \right).
\]

Also, by definition, for each \( n \in \mathbb{N} \)
\[
\Lambda_{\Pi}(\bar{A}_{N}, f_n) \leq \bar{E} \left[ \frac{1}{2} \int_{0}^{T} \| \bar{\psi}_{\varepsilon}(s) \|^2 \, ds + f_n \left( \bar{W} + \int_{0}^{T} \bar{\psi}_{\varepsilon}(s) \, ds, \bar{X} \right) \right].
\]

Combining the last three displays gives \( \limsup_{n \to \infty} \Lambda_{\Pi}(\bar{A}_{N}, f_n) \leq \Lambda_{\Pi}(\bar{A}_{N}, f) + \varepsilon \). The result follows since \( \varepsilon > 0 \) is arbitrary.

The following lemma provides a simpler representation than (8.13).

**Lemma 8.5.** Let \( f \in M_{b}(C \times S) \). Then
\[
- \log \bar{E} [\exp (-f(B, X))] = \inf_{\Pi \in \mathcal{P}(S), \bar{\psi} \in \mathcal{A}} \left\{ R(\Pi \| \rho) + \bar{E} \left[ \frac{1}{2} \int_{0}^{T} \| \bar{\psi}(s) \|^2 \, ds + f \left( \bar{W} + \int_{0}^{T} \bar{\psi}(s) \, ds, \bar{X} \right) \right] \right\}. \tag{8.16}
\]

**Proof.** In view of (8.11), it suffices to prove that for every \( \Pi \in \mathcal{P}(S) \) such that \( R(\Pi \| \rho) < \infty \),
\[
\int_{S} \hat{f}(x) \Pi(\, dx) = \inf_{\bar{\psi} \in \bar{A}} \bar{E} \left[ \frac{1}{2} \int_{0}^{T} \| \bar{\psi}(s) \|^2 \, ds + f \left( \bar{W} + \int_{0}^{T} \bar{\psi}(s) \, ds, \bar{X} \right) \right]. \tag{8.17}
\]
where $\tilde{f}$ is as defined in (8.10). We claim that it suffices to prove (8.17) for $f \in C_b(C \times \mathcal{S})$. To see this, let $\{f_n\}$ be a sequence of elements in $C_b(C \times \mathcal{S})$ such that $\|f_n\|_\infty \leq \|f\|_\infty$ and $f_n$ converges to $f$ a.s. $\tilde{P} \times \Pi$. It then follows from dominated convergence theorem that as $n \to \infty$, $\tilde{f}_n \to \tilde{f}$ a.s. $\Pi$ and hence

$$\int_{\mathcal{S}} \tilde{f}_n(x)\Pi(dx) \to \int_{\mathcal{S}} \tilde{f}(x)\Pi(dx).$$

To prove the claim, it then remains to show that $\Lambda_{\Pi}(\tilde{A}, f_n) \to \Lambda_{\Pi}(\tilde{A}, f)$ as $n \to \infty$, where $\Lambda_{\Pi}$ is as in (8.15). Let

$$\tilde{A}_f = \left\{ \tilde{\psi} \in \tilde{A} : E \int_0^T \|\tilde{\psi}(s)\|^2 ds \leq 4\|f\|_\infty \right\}.$$  

Then clearly

$$\Lambda_{\Pi}(\tilde{A}, f_n) = \Lambda_{\Pi}(\tilde{A}_f, f_n) \text{ and } \Lambda_{\Pi}(\tilde{A}, f) = \Lambda_{\Pi}(\tilde{A}_f, f).$$

Fix $\varepsilon \in (0, 1)$ and choose $N > 0$ such that $4\|f\|_\infty^2 / N \leq \varepsilon/2$. Fix $\tilde{\psi} \in \tilde{A}_f$ and define the stopping time

$$\tau_N(\tilde{\omega}) = \inf \left\{ t \in [0, T] : \int_0^t \|\tilde{\psi}(s, \tilde{\omega})\|^2 ds \geq N \right\} \wedge T, \ \tilde{\omega} \in \tilde{\Omega}.$$

Let $\tilde{\psi}_N(s) = \tilde{\psi}(s)1_{[0, \tau_N]}(s)$. Then $\tilde{\psi}_N \in \tilde{A}^N$, where $\tilde{A}^N$ is as defined in (8.14), and

$$P(\tilde{\psi}_N \neq \tilde{\psi}) \leq P(\tau_N < T) \leq P \left( \int_0^T \|\tilde{\psi}(s)\|^2 ds \geq N \right) \leq 4\|f\|_\infty / N.$$

Hence we have

$$\Lambda_{\Pi}(\tilde{A}^N, f_n) \leq E \left[ \frac{1}{2} \int_0^T \|\tilde{\psi}_N(s)\|^2 ds + f_n \left( \tilde{W} + \int_0^\cdot \tilde{\psi}_N(s) ds, \tilde{X} \right) \right]$$

$$\leq E \left[ \frac{1}{2} \int_0^T \|\tilde{\psi}(s)\|^2 ds + f_n \left( \tilde{W} + \int_0^\cdot \tilde{\psi}(s) ds, \tilde{X} \right) \right] + \varepsilon. $$

Taking the infimum over all $\tilde{\psi} \in \tilde{A}_f$ in the above inequality we have

$$\Lambda_{\Pi}(\tilde{A}_f, f_n) \leq \Lambda_{\Pi}(\tilde{A}^N, f_n) \leq \Lambda_{\Pi}(\tilde{A}_f, f_n) + \varepsilon.$$

Exactly the same argument with $f_n$ replaced by $f$ gives

$$\Lambda_{\Pi}(\tilde{A}_f, f) \leq \Lambda_{\Pi}(\tilde{A}^N, f) \leq \Lambda_{\Pi}(\tilde{A}_f, f) + \varepsilon.$$
From Lemma 8.4 we have that as $n \to \infty$,

$$\Lambda_\Pi(\bar{A}^N, f_n) \to \Lambda_\Pi(\bar{A}^N, f).$$

This proves the claim since $\varepsilon > 0$ is arbitrary.

Henceforth we will assume that $f \in C_b(C \times \mathbb{S})$. We first argue that LHS ≤ RHS in (8.17). Fix $\bar{\psi} \in \bar{A}$. Since $\bar{\psi}$ is $\bar{F}_t$-progressively measurable, there exists (cf. [65], Exercise 1.5.6) a $\mathcal{B}([0, T]) \times \mathcal{B}(\mathbb{S}) \times \mathcal{B}(\mathcal{C}) / \mathcal{B}(\mathbb{R}^d)$-measurable map $F: [0, T] \times \mathbb{S} \times \mathcal{C} \to \mathbb{R}^d$ such that

$$\bar{\psi}(s, \bar{\omega}) = F(s, \bar{X}(\bar{\omega}), \bar{W}_{s\wedge} (\bar{\omega})), \quad \mathcal{P}$$. a.s. $(s, \bar{\omega}) \in [0, T] \times \bar{\Omega}.$

Define a collection of processes $\{\bar{\psi}_x\}_{x \in \mathbb{S}}$ on $\bar{\Omega}$ as

$$\bar{\psi}_x(s, \bar{\omega}) = F(s, x, \bar{W}_{s\wedge} (\bar{\omega})), \quad (x, s, \bar{\omega}) \in \mathbb{S} \times [0, T] \times \bar{\Omega}.$$

Then $\bar{\psi}_x \in \bar{A}$ for $\Pi$-a.e. $x \in \mathbb{S}$. By independence between $\bar{X}$ and $\bar{W}$,

$$\bar{\mathbb{E}} \left[ \frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + f \left( \bar{W} + \int_0^s \bar{\psi}(s) ds, \bar{X} \right) \right]$$

$$= \int_{\mathbb{S}} \bar{\mathbb{E}} \left[ \frac{1}{2} \int_0^T \|F(s, x, \bar{W}_{s\wedge})\|^2 ds + f \left( \bar{W} + \int_0^s F(s, x, \bar{W}_{s\wedge}) ds, x \right) \right] \Pi(dx)$$

$$= \int_{\mathbb{S}} \bar{\mathbb{E}} \left[ \frac{1}{2} \int_0^T \|\bar{\psi}_x(s)\|^2 ds + f \left( \bar{W} + \int_0^s \bar{\psi}_x(s) ds, x \right) \right] \Pi(dx)$$

$$\geq \int_{\mathbb{S}} \tilde{f}(x) \Pi(dx),$$

where the last inequality is from (8.12). Taking the infimum over all $\bar{\psi} \in \bar{A}$ in the above inequality implies that LHS ≤ RHS in (8.17).

We now consider the reverse inequality. Fix $\varepsilon \in (0, 1)$. For each $x \in \mathbb{S}$, let $\bar{\psi}_{x, \varepsilon} \in \bar{A}$ be an $\varepsilon$-optimal control in (8.12) such that

$$\tilde{f}(x) \geq \bar{\mathbb{E}} \left[ \frac{1}{2} \int_0^T \|\bar{\psi}_{x, \varepsilon}(s)\|^2 ds + f \left( \bar{W} + \int_0^s \bar{\psi}_{x, \varepsilon}(s) ds, x \right) \right] - \varepsilon. \quad (8.18)$$
We will now carefully select a countable sub-collection from \( \{ \tilde{\psi}_{x,\epsilon} \}_{x \in S} \) and use it to construct an \( \mathcal{F}_t \)-progressively measurable process \( \tilde{\psi}_\epsilon \in \tilde{A} \). From (8.18) we have
\[
\sup_{x \in S} \mathbb{E} \left[ \frac{1}{2} \int_0^T \| \tilde{\psi}_{x,\epsilon}(s) \|_2^2 \, ds \right] \leq 2 \| f \|_\infty + 1.
\]

Using this it is easy to check that \( \{ \int_0^\cdot \tilde{\psi}_{x,\epsilon}(s) \, ds \}_{x \in S} \) is tight in \( C \) and thus so is \( \{ \tilde{W}^{x,\epsilon} = \tilde{W} + \int_0^\cdot \tilde{\psi}_{x,\epsilon}(s) \, ds \}_{x \in S} \). Then there exists a compact subset \( K_\epsilon \subset C \) such that
\[
\sup_{x \in S} \mathbb{E} \left[ f \left( \tilde{W}^{x,\epsilon} / \in K_\epsilon \right) \right] \leq \varepsilon / \| f \|_\infty. \tag{8.19}
\]

Let \( \tilde{K}_\epsilon \) be a compact subset of \( S \) such that
\[
\Pi(S \setminus \tilde{K}_\epsilon) \leq \varepsilon / \| f \|_\infty.
\]

Since \( f \) is continuous, so is \( \tilde{f} \). In particular, \( f \) is uniformly continuous on \( K_\epsilon \times \tilde{K}_\epsilon \) and \( \tilde{f} \) is uniformly continuous on \( \tilde{K}_\epsilon \). So there exists some \( M_\epsilon \in \mathbb{N} \) and a finite partition \( \{ B_i : i = 1, \ldots, M_\epsilon \} \) of \( \tilde{K}_\epsilon \) such that
\[
\max_{i=1,\ldots,M_\epsilon} \sup_{\phi \in K_\epsilon, x_1, x_2 \in B_i} | f(\phi, x_1) - f(\phi, x_2) | + | \tilde{f}(x_1) - \tilde{f}(x_2) | \leq \varepsilon. \tag{8.20}
\]

Now for each \( i = 1, \ldots, M_\epsilon \), fix a \( y_i \in B_i \) and define
\[
\tilde{\psi}_{y_i,\epsilon}(s, \tilde{\omega}) \doteq \begin{cases} 
\tilde{\psi}_{y_i,\epsilon}(s, \tilde{W}(\tilde{\omega})), & \tilde{\omega} \in \tilde{X}^{-1}(B_i), \\
0, & \tilde{\omega} \notin \tilde{X}^{-1}(\tilde{K}_\epsilon).
\end{cases}
\]

Since \( \tilde{\psi}_{y_i,\epsilon} \) is \( \mathcal{F}_t \)-progressively measurable for each \( i = 1, \ldots, M_\epsilon \) and \( \tilde{X} \) is \( \mathcal{F}_0 \)-measurable, \( \tilde{\psi}_\epsilon \) is \( \mathcal{F}_t \)-progressively measurable. For \( x \in B_i, i \in \{1, \ldots, M_\epsilon\} \), we have
\[
\mathbb{E} f \left( \tilde{W} + \int_0^\cdot \tilde{\psi}_{y_i,\epsilon}(s, \tilde{W}) \, ds, x \right) = \mathbb{E} f \left( \tilde{W}^{y_i,\epsilon}, x \right) 
\leq \mathbb{E} \left[ f \left( \tilde{W}^{y_i,\epsilon}, x \right) 1_{\{ \tilde{W}^{y_i,\epsilon} / \in K_\epsilon \}} \right] + \varepsilon 
\leq \mathbb{E} \left[ f \left( \tilde{W}^{y_i,\epsilon}, y_i \right) 1_{\{ \tilde{W}^{y_i,\epsilon} / \in K_\epsilon \}} \right] + 2\varepsilon 
\leq \mathbb{E} f \left( \tilde{W}^{y_i,\epsilon}, y_i \right) + 3\varepsilon,
\]

154
where the first and third inequalities use (8.19), and the second inequality follows from (8.20). Using this, definition of $\tilde{K}_\varepsilon$ and the independence between $\bar{W}$ and $\bar{X}$, we have

$$\mathbb{E}f\left(\bar{W} + \int_0^T \bar{\psi}_\varepsilon(s) \, ds, \bar{X}\right)$$

$$= \sum_{i=1}^{M_\varepsilon} \mathbb{E}\left[1_{\{\bar{X} \in B_i\}} f\left(\bar{W} + \int_0^T \bar{\psi}_\varepsilon(s) \, ds, \bar{X}\right)\right] + \mathbb{E}\left[1_{\{\bar{X} \notin \tilde{K}_\varepsilon\}} f\left(\bar{W} + \int_0^T \bar{\psi}_\varepsilon(s) \, ds, \bar{X}\right)\right]$$

$$\leq \sum_{i=1}^{M_\varepsilon} \int_{B_i} \mathbb{E}f\left(\bar{W} + \int_0^T \bar{\psi}_{y_i,\varepsilon}(s, \bar{W}) \, ds, x\right) \Pi(dx) + \varepsilon$$

$$\leq \sum_{i=1}^{M_\varepsilon} \int_{B_i} \mathbb{E}f\left(\bar{W}_{y_i,\varepsilon}, y_i\right) \Pi(dx) + 4\varepsilon.$$ 

Also note that

$$\mathbb{E}\int_0^T \|\bar{\psi}_\varepsilon(s)\|^2 \, ds = \sum_{i=1}^{M_\varepsilon} \mathbb{E}\left[\int_0^T \|\bar{\psi}_{y_i,\varepsilon}(s, \bar{W})\|^2 \, ds\right]$$

$$= \sum_{i=1}^{M_\varepsilon} \int_{B_i} \mathbb{E}\int_0^T \|\bar{\psi}_{y_i,\varepsilon}(s)\|^2 \, ds \Pi(dx).$$

Combining above two displays, we have the following estimate of the RHS in (8.17)

$$\mathbb{E}\left[\frac{1}{2} \int_0^T \|\bar{\psi}_\varepsilon(s)\|^2 \, ds + f\left(\bar{W} + \int_0^T \bar{\psi}_\varepsilon(s) \, ds, \bar{X}\right)\right]$$

$$\leq \sum_{i=1}^{M_\varepsilon} \int_{B_i} \mathbb{E}\left[\frac{1}{2} \int_0^T \|\bar{\psi}_{y_i,\varepsilon}(s)\|^2 \, ds + f\left(\bar{W}_{y_i,\varepsilon}, y_i\right)\right] \Pi(dx) + 4\varepsilon.$$

It then follows from (8.18), (8.20) and definition of $\tilde{K}_\varepsilon$ that the above display can be bounded by

$$\sum_{i=1}^{M_\varepsilon} \int_{B_i} \tilde{f}(y_i) \Pi(dx) + 5\varepsilon \leq \sum_{i=1}^{M_\varepsilon} \int_{B_i} \tilde{f}(x) \Pi(dx) + 6\varepsilon \leq \int_S \tilde{f}(x) \Pi(dx) + 7\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary we have RHS $\leq$ LHS in (8.17) and this completes the proof.

In the proof of Theorem 8.3 we will need to work with a filtration that is larger than the one generated by the Brownian motions. So it will be convenient to write the above result in a general form, which allows for an arbitrary filtration, as in Proposition 8.6 below. Let $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ be as before, on which is given a standard $d$-dimensional Brownian motion $\bar{W}$ and an independent
\[ -\log E[\exp(-f(B, X))] \]
\[ = \inf_{\Upsilon} \inf_{\hat{\psi} \in \hat{\mathcal{A}}(\Upsilon)} \left\{ R(\Pi \| \rho) + \mathbb{E} \left[ \frac{1}{2} \int_0^T \| \hat{\psi}(s) \|^2 ds + f \left( \hat{W} + \int_0^s \hat{\psi}(s) ds, \hat{X} \right) \right] \right\}. \quad (8.21) \]

**Proposition 8.6.** Let \( f \in \mathcal{M}_b(C \times S) \). Then

\[ -\log E[\exp(-f(B, X))] \]
\[ = \inf_{\Upsilon} \inf_{\hat{\psi} \in \hat{\mathcal{A}}(\Upsilon)} \left\{ R(\Pi \| \rho) + \mathbb{E} \left[ \frac{1}{2} \int_0^T \| \hat{\psi}(s) \|^2 ds + f \left( \hat{W} + \int_0^s \hat{\psi}(s) ds, \hat{X} \right) \right] \right\}. \quad (8.22) \]

**Proof.** It suffices to prove that with \( \Upsilon, \hat{X}, \hat{W}, \Pi \) as above, such that \( R(\Pi \| \rho) < \infty \),

\[ \inf_{\hat{\psi} \in \hat{\mathcal{A}}(\Upsilon)} \mathbb{E} \left[ \frac{1}{2} \int_0^T \| \hat{\psi}(s) \|^2 ds + f \left( \hat{W} + \int_0^s \hat{\psi}(s) ds, \hat{X} \right) \right] \]
\[ = \inf_{\hat{\psi} \in \hat{\mathcal{A}}(\Upsilon)} \mathbb{E} \left[ \frac{1}{2} \int_0^T \| \hat{\psi}(s) \|^2 ds + f \left( \hat{W} + \int_0^s \hat{\psi}(s) ds, \hat{X} \right) \right]. \quad (8.22) \]

We first claim that it suffices to prove (8.22) for \( f \in \mathcal{C}_b(C \times S) \). The verification of this claim follows along the same lines as the proof of the claim in the proof of Lemma 8.5, and hence we only provide a sketch here. With \( \{f_n\} \subset \mathcal{C}_b(C \times S) \) introduced as below (8.17), it was shown in proof of Lemma 8.5 that \( \Lambda_\Pi(\bar{A}, f_n) \rightarrow \Lambda_\Pi(\bar{A}, f) \). Thus it suffices to show \( \Lambda_\Pi(\hat{A}(\Upsilon), f_n) \rightarrow \Lambda_\Pi(\hat{A}(\Upsilon), f) \) as \( n \rightarrow \infty \), where \( \Lambda_\Pi(\hat{A}(\Upsilon), g) \) is defined by the right side of (8.15), replacing \( \bar{A} \) with \( \hat{A}(\Upsilon) \). For each \( N < \infty \), define \( \hat{A}^N(\Upsilon) \) as in (8.14) with \( \bar{A} \) replaced by \( \hat{A}(\Upsilon) \). Then by the same stopping time argument below (8.17), it suffices to argue that for each \( N < \infty \), \( \Lambda_\Pi(\hat{A}^N(\Upsilon), f_n) \rightarrow \Lambda_\Pi(\hat{A}^N(\Upsilon), f) \) as \( n \rightarrow \infty \). However the proof of this is identical to the proof of Lemma 8.4. Indeed, it is easily seen that the form of the filtration (as long as \( \hat{W} \) is a Brownian motion with respect to the filtration and \( \hat{X} \) is measurable with respect to \( \hat{F}_0 \)) does not play any role in the proof. This completes the proof of the claim.
Henceforth we will assume that $f \in \mathbb{C}_b(C \times S)$. It is clear that LHS $\geq$ RHS in (8.22). For the reverse inequality, we will show that

$$
\inf_{\bar{\psi} \in \tilde{A}} \mathbb{E} \left[ \frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 \, ds + f \left( \bar{W} + \int_0^T \bar{\psi}(s) \, ds, \bar{X} \right) \right] 
\leq \mathbb{E} \left[ \frac{1}{2} \int_0^T \|\hat{\psi}(s)\|^2 \, ds + f \left( \hat{W} + \int_0^T \hat{\psi}(s) \, ds, \hat{X} \right) \right] 
$$

(8.23)

for each $\hat{\psi} \in \hat{A}(\Upsilon)$, and the proof is done in two steps.

**Step 1. Simple $\hat{\psi}$.** For simplicity we consider the case in which

$$
\hat{\psi}(s) = Y \mathbf{1}_{[t,T]}(s),
$$

where $t \in [0, T]$, $Y$ is $\hat{F}_t$-measurable, and $\|Y\| \leq N < \infty$ a.s. The proof for a general simple process is similar. Consider the map

$$
g(\phi, x, y) \doteq \mathbb{E} \left[ \frac{T-t}{2} \|y\|^2 + f \left( \phi^B + \int_0^T y \mathbf{1}_{[t,T]}(s) \, ds, x \right) \right],
$$

where $(\phi, x) \in \mathcal{C}([0, t] : \mathbb{R}^d) \times S$, $y \in K_N \doteq \{z \in \mathbb{R}^d : \|z\| \leq N\}$, and

$$
\phi^B(s) \doteq \begin{cases} 
\phi(s), & s \in [0, t], \\
\phi(t) + B(s-t), & s \in [t, T]. 
\end{cases}
$$

Note that $g$ is bounded, and that by the dominated convergence theorem it is also continuous in $(\phi, x, y)$. From a classical measurable selection result (see e.g. [10], Lemma 3.3) there exists a Borel measurable function $g_1 : \mathcal{C}([0, t] : \mathbb{R}^d) \times S \rightarrow K_N$ such that

$$
g(\phi, x, g_1(\phi, x)) \leq g(\phi, x, y)
$$

for all $(\phi, x) \in \mathcal{C}([0, t] : \mathbb{R}^d) \times S$ and $y \in K_N$. With the definition $\bar{W}_{[0,t]} \doteq \{\bar{W}(s)\}_{0 \leq s \leq t}$ and $\bar{Y} \doteq g_1(\bar{W}_{[0,t]}, \bar{X})$, we set

$$
\bar{\psi}(s) \doteq \bar{Y} \mathbf{1}_{[t,T]}(s) \in \bar{A}.
$$
Then
\[\mathbb{E} \left[ \frac{1}{2} \int_0^T \| \hat{\psi}(s) \|^2 ds + f \left( \hat{W} + \int_0^T \hat{\psi}(s) ds, \hat{X} \right) \right] \]
\[= \mathbb{E} \left\{ \mathbb{E} \left[ \frac{T - t}{2} \| Y \|^2 + f \left( \hat{W} + \int_0^T Y 1_{[t,T]}(s) ds, \hat{X} \right) \right] \mid \mathcal{F}_t \right\} \]
\[= \mathbb{E} g(W_{[0,t]}, \hat{X}, Y) \]
\[\geq \mathbb{E} g(W_{[0,t]}, \bar{X}, g_1(W_{[0,t]}, \bar{X})) \]
\[= \mathbb{E} g(W_{[0,t]}, \bar{X}, \bar{Y}) \]
\[\mathbb{E} \left\{ \mathbb{E} \left[ \frac{T - t}{2} \| Y \|^2 + f \left( \hat{W} + \int_0^T Y 1_{[t,T]}(s) ds, \hat{X} \right) \right] \mid \mathcal{F}_t \right\} \]
\[= \mathbb{E} \left[ \frac{1}{2} \int_0^T \| \hat{\psi}(s) \|^2 ds + f \left( \hat{W} + \int_0^T \hat{\psi}(s) ds, \hat{X} \right) \right]. \]

So (8.23) holds for simple \( \hat{\psi} \in \hat{A}(\U). \)

**Step 2. General \( \hat{\psi} \).** Next consider \( \hat{\psi} \in \hat{A}(\U). \) We can assume without loss of generality that \( \mathbb{E} \int_0^T \| \hat{\psi}(s) \|^2 ds < \infty. \) Then (cf. [45], Lemma 3.2.4) there exists a sequence of simple processes \( \{ \hat{\psi}_n \}_{n \in \mathbb{N}} \subset \hat{A}(\U) \) such that
\[
\lim_{n \to \infty} \mathbb{E} \int_0^T \| \hat{\psi}_n(s) - \hat{\psi}(s) \|^2 ds = 0.
\]
This implies that as \( n \to \infty, \)
\[
\mathbb{E} \sup_{0 \leq t \leq T} \left\| \int_0^t \hat{\psi}_n(s) ds - \int_0^t \hat{\psi}(s) ds \right\|^2 \leq T \mathbb{E} \int_0^T \| \hat{\psi}_n(s) - \hat{\psi}(s) \|^2 ds \to 0
\]
and therefore
\[
\left( \hat{W} + \int_0^T \hat{\psi}_n(s) ds, \hat{X} \right) \Rightarrow \left( \hat{W} + \int_0^T \hat{\psi}(s) ds, \hat{X} \right).
\]
Combining above displays and using Step 1, we have
\[
\inf_{\hat{\psi} \in \hat{A}} \mathbb{E} \left[ \frac{1}{2} \int_0^T \| \hat{\psi}(s) \|^2 ds + f \left( \hat{W} + \int_0^T \hat{\psi}(s) ds, \hat{X} \right) \right] \]
\[\leq \mathbb{E} \left[ \frac{1}{2} \int_0^T \| \hat{\psi}_n(s) \|^2 ds + f \left( \hat{W} + \int_0^T \hat{\psi}_n(s) ds, \hat{X} \right) \right] \]
\[\to \mathbb{E} \left[ \frac{1}{2} \int_0^T \| \hat{\psi}(s) \|^2 ds + f \left( \hat{W} + \int_0^T \hat{\psi}(s) ds, \hat{X} \right) \right], \]
as $n \to \infty$, where the convergence holds since $f \in C_b(\mathcal{C} \times \mathcal{S})$. Thus (8.23) holds for general $\hat{\psi} \in \hat{\mathcal{A}}(\Upsilon)$ and this completes the proof. \hfill\Box

### 8.5 Variational representation and tightness properties

In this section we will apply Proposition 8.6 to establish a variational representation for

$$-rac{1}{n} \log \mathbb{E} \left[ \exp \left( -nF(\mu^n) \right) \right],$$

where $F \in C_b(\mathcal{D})$.

Let $\Upsilon = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$ be a filtered probability space on which is given an $\mathbb{R}^d_+$-valued $\mathcal{F}_0$-measurable random variable $S^n = (S^n_i)_{i=1}^n$ with law $\Pi^n \in \mathcal{P}(\mathbb{R}^d_+)$, and an $nd$-dimensional standard $\mathcal{F}_t$-Brownian motion $\beta^n = (\beta^n_i)_{i=1}^n$. Denote by $\mathbb{E}$ the expectation under $\mathbb{P}$. We can disintegrate $\Pi^n$ as

$$\Pi^n(dx_1, \ldots, dx_n) = \Pi^n_1(dx_1)\Pi^n_2(x_1, dx_2) \cdots \Pi^n_n(x_1, \ldots, x_{n-1}, dx_n),$$

where $x^n = (x_i)_{i=1}^n \in \mathbb{R}^d_+$. Define the random measures

$$\nu^n_i = \nu^n_i(S^n, \cdot), \ i = 1, \ldots, n. \quad (8.24)$$

Let

$$\mathcal{A}_n(\Upsilon) = \{ \psi^n = (\psi^n_i)_{i=1}^n : \text{the process } \psi^n_i \text{ is } \mathbb{R}^d_+\text{-valued } \mathcal{F}_t\text{-progressively measurable}

\text{and } \mathbb{E} \sum_{i=1}^n \int_0^T \|\psi^n_i(s)\|^2 ds < \infty \},$$

For $\psi^n \in \mathcal{A}_n(\Upsilon)$, consider the following controlled processes

$$\tilde{\beta}^n = (\tilde{\beta}^n_i)_{i=1}^n, \quad \tilde{\beta}^n_i(t) = \beta^n_i(t) + \int_0^t \psi^n_i(s) ds, \ i = 1, \ldots, n, \quad (8.25)$$

$$\tilde{\mu}^n(t) = \frac{1}{n} \sum_{i=1}^n \delta_{\tilde{\beta}^n_i(t)} 1\{S^n_i > \int_0^t \langle \zeta, \tilde{\mu}^n(s) \rangle ds \}. \quad (8.26)$$
We then have the following variational representation formula which is the starting point of our proof to Theorem 8.3.

**Lemma 8.7.** Let $F \in \mathbb{C}_b(D)$. Then

$$
-\frac{1}{n} \log \mathbb{E} [\exp (-n F(\mu^n))] \\
= \inf_{\Upsilon} \inf_{\psi^n \in A_n(\Upsilon)} \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^{n} \left( R(\nu_i^n \| \theta) + \frac{1}{2} \int_0^T \| \psi_i^n(s) \|^2 ds \right) + F(\tilde{\mu}^n) \right].
$$

(8.27)

**Proof.** Note that $F(\mu^n) = F\left( \left\{ \frac{1}{n} \sum_{i=1}^{n} \delta_{B_i(t)} 1 \{ X_i > f_i^t \langle \zeta, \mu^n(s) \rangle ds \} \right\}_{t \in [0,T]} \right) \equiv \Psi(B^n, X^n)$ for some $\Psi \in \mathcal{M}_b(C^n \times \mathbb{R}_+^n)$, where $B^n \equiv (B_i)_{i=1}^{n}$ and $X^n \equiv (X_i)_{i=1}^{n}$. With the same measurable function $\Psi$, $F(\tilde{\mu}^n) = \Psi(\tilde{\beta}^n, S^n)$ a.s. Applying Proposition 8.6 with $C, S, f, \bar{W}$ and $\bar{X}$ replaced by $C^n, \mathbb{R}_+^n, n\Psi, \beta^n$ and $S^n$ respectively, we have

$$
-\frac{1}{n} \log \mathbb{E} [\exp (-n F(\mu^n))] \\
= \inf_{\Upsilon} \inf_{\psi^n \in A_n(\Upsilon)} \left\{ \frac{1}{n} R(\Pi^n \| \theta^n) + \mathbb{E} \left[ \frac{1}{2n} \sum_{i=1}^{n} \int_0^T \| \psi_i^n(s) \|^2 ds + F(\tilde{\mu}^n) \right] \right\}.
$$

Using chain rule for relative entropies (cf. [27], Theorem C.3.1) it follows that

$$
R(\Pi^n \| \theta^n) = \bar{E} \left[ \sum_{i=1}^{n} R(\nu_i^n \| \theta) \right],
$$

where $(\nu_i^n)_{i=1}^{n}$ is as defined in (8.24). The result follows on combining the above two displays. \(\square\)

For each $n \in \mathbb{N}$, let $\Upsilon$ (with the associated $\Pi^n$ and $S^n$) and $\psi^n \in A_n(\Upsilon)$ be such that

$$
\sup_{n \in \mathbb{N}} \bar{E} \left[ \frac{1}{n} \sum_{i=1}^{n} \left( R(\nu_i^n \| \theta) + \frac{1}{2} \int_0^T \| \psi_i^n(s) \|^2 ds \right) \right] \equiv C_0 < \infty.
$$

(8.28)

For $i = 1, \ldots, n$, define $\mathcal{R}^W$-valued random variables $\rho_i^n$ as

$$
\rho_i^n(A \times [0,t]) \equiv \int_0^t \delta_A(\psi_i^n(s)) ds, \ A \in \mathcal{B}(\mathbb{R}_+^n), \ t \in [0,T].
$$
Next, define a sequence $\{(Q^n, \nu^n)\}_{n \in \mathbb{N}}$ of $\mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R}_+)$-valued random variables as

$$Q^n(A) = \frac{1}{n} \sum_{i=1}^{n} \delta_{(\beta_i^n, \beta_i^n \cdot \rho_i^n, s_i^n)}(A), \ A \in \mathcal{B}(\Xi)$$

(8.29)

and

$$\nu^n(A) = \frac{1}{n} \sum_{i=1}^{n} \nu^n_i(A), \ A \in \mathcal{B}(\mathbb{R}_+).$$

(8.30)

We then have the following tightness result.

**Lemma 8.8.** The sequence of random variables $\{(Q^n, \nu^n)\}_{n \geq 1}$ is tight in $\mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R}_+)$.  

**Proof.** We first argue tightness of $\{Q^n\}$. Since $\{\beta_i^n\}$ are standard Brownian motions, $\{Q^n_2\}$ is tight.

Next we consider $Q^n_3$.

Next we consider $\check{Q}^n_3$. Note that

$$g_1(r) = \int_{\mathbb{R}_d \times [0,T]} \|y\|^2 r(dy \, dt), \ r \in \mathcal{R}^W,$$

is a tightness function on $\mathcal{R}^W$, namely it is bounded from below and has pre-compact level sets. This says that $G_1: \mathcal{P}(\mathcal{R}^W) \rightarrow [0, \infty]$ defined as

$$G_1(m) = \int_{\mathcal{R}^W} g_1(r) \, m(dr), \ m \in \mathcal{P}(\mathcal{R}^W)$$

is a tightness function on $\mathcal{P}(\mathcal{R}^W)$. Next note that

$$\check{E}G_1(Q^n_3) = \frac{1}{n} \sum_{i=1}^{n} \check{E}g_1(\rho_i^n) = \frac{1}{n} \sum_{i=1}^{n} \check{E} \int_{\mathbb{R}_d \times [0,T]} \|y\|^2 \rho_i^n(dy \, dt) = \frac{1}{n} \sum_{i=1}^{n} \check{E} \int_{0}^{T} \|\psi_i^n(s)\|^2 \, ds \leq 2C_0.$$

This proves that $\{Q^n_3\}$ is tight.

We now argue tightness of $\{Q^n_1\}$. Define a sequence $\{\check{Q}^n\}_{n \in \mathbb{N}}$ of $\mathcal{P}(\mathcal{C})$-valued random variables as

$$\check{Q}^n(A) = \frac{1}{n} \sum_{i=1}^{n} \delta_{u_i^n}(A), \ A \in \mathcal{B}(\mathcal{C}),$$

where

$$u_i^n(t) = \int_{\mathbb{R}_d \times [0,t]} y \rho_i^n(dy \, ds) = \int_{0}^{t} \psi_i^n(s) \, ds, \ t \in [0,T].$$

161
We claim that \( \{ \tilde{Q}_n \}_{n \in \mathbb{N}} \) is a tight sequence. To see this, note that \( g_2 : \mathcal{C} \to [0, \infty] \) defined as
\[
g_2(x) = \begin{cases} \int_0^T \| \dot{x}(s) \|^2 \, ds + \| x(0) \|^2, & \text{if } x \text{ is absolutely continuous,} \\ \infty, & \text{otherwise,} \end{cases}
\]
is a tightness function on \( \mathcal{C} \), from which it follows that \( G_2 : \mathcal{P}(\mathcal{C}) \to [0, \infty] \), defined as
\[
G_2(m) = \int_\mathcal{C} g_2(x) \, m(dx), \quad m \in \mathcal{P}(\mathcal{C}),
\]
is a tightness function on \( \mathcal{P}(\mathcal{C}) \). Also,
\[
\bar{E}(G_2(\tilde{Q}_n)) = \frac{1}{n} \sum_{i=1}^n \bar{E}g_2(u^n_i) = \frac{1}{n} \sum_{i=1}^n \bar{E} \int_0^T \| \psi^n_i(s) \|^2 \, ds \leq 2C_0.
\]
This proves tightness of \( \{ \tilde{Q}_n \}_{n \in \mathbb{N}} \). Define the sequence \( \{ \tilde{Q}^n_{(2,3)} \}_{n \in \mathbb{N}} \) of \( \mathcal{P}(\mathcal{C} \times \mathcal{C}) \)-valued random variables as
\[
\tilde{Q}^n_{(2,3)}(A) = \frac{1}{n} \sum_{i=1}^n \delta(\beta^n_i, u^n_i)(A), \quad A \in \mathcal{B}(\mathcal{C} \times \mathcal{C}).
\]
Then, from tightness of \( \{ Q^n_{(2)} \} \) and \( \{ \tilde{Q}^n \} \) it follows that \( \{ \tilde{Q}^n_{(2,3)} \} \) is tight. Next, noting that the map \( g_3 : \mathcal{C} \times \mathcal{C} \to \mathcal{C} \) defined as \( g_3(x, u) = x + u \) is continuous and that \( Q^n_{(1)} = \tilde{Q}^n_{(2,3)} \circ g_3^{-1} \), we get tightness of \( \{ Q^n_{(1)} \} \).

Finally, tightness of \( \{ Q^n_{(4)} \} \) and \( \{ \nu^n \} \) can be proved using a standard argument as in the proof of Sanov’s theorem (see e.g. [27], Theorem 2.5.1). However we provide details for sake of completeness.

Note that \( G_3(\cdot) \doteq R(\cdot \| \theta) \) is a convex tightness function on \( \mathcal{P}(\mathbb{R}_+) \) (cf. [27], Theorem 1.4.3), and that by Jensen’s inequality,
\[
\bar{E}G_3(\nu^n) = \bar{E}R(\nu^n \| \theta) \leq \bar{E} \left[ \frac{1}{n} \sum_{i=1}^n R(\nu^n_{i} \| \theta) \right] \leq C_0.
\]
So \( \{ \nu^n \} \) is tight, which implies that \( \{ \bar{E}\nu^n \} \) is also tight. Next, since \( \nu^n_i \) is the conditional distribution used to select \( S^n_i \), for any \( f \in M_b(\mathbb{R}^d) \),
\[
\bar{E} \int_{\mathbb{R}^d} f(x) \, Q^n_{(4)}(dx) = \frac{1}{n} \sum_{i=1}^n \bar{E}f(S^n_i) = \frac{1}{n} \sum_{i=1}^n \bar{E} \int_{\mathbb{R}^d} f(x) \, \nu^n_i(dx) = \bar{E} \int_{\mathbb{R}^d} f(x) \, \nu^n(dx).
\]
So \( \tilde{E}Q^n_{(4)} = \tilde{E}\nu^n \), from which we have the desired tightness of \( \{Q^n_{(4)}\} \).

The following lemma gives useful characterization of weak limit points of \( \{(Q^n, \nu^n) : n \in \mathbb{N}\} \).

**Lemma 8.9.** Suppose \( (Q^n, \nu^n) \) converges along a subsequence, in distribution, to \( (Q^*, \nu^*) \) given on some probability space \( (\Omega^*, \mathcal{F}^*, \mathbb{P}^*) \). Then \( Q^*_{(4)} = \nu^* \) a.s. and \( Q^* \in \mathcal{P}_\infty \) a.s. \( \mathbb{P}^* \).

**Proof.** The first statement, namely \( Q^*_{(4)} = \nu^* \) a.s. \( \mathbb{P}^* \), can be proved using a standard martingale argument as in the proof of Sanov’s theorem (see e.g. [27], Theorem 2.5.1).

In order to check that \( Q^* \in \mathcal{P}_\infty \) a.s. \( \mathbb{P}^* \), we need to verify that under \( Q^* \), properties (1)--(4) of Section 8.3.2 are satisfied (with \( \Theta \) replaced by \( Q^* \)). Without loss of generality assume that \( (Q^n, \nu^n) \) converges weakly to \( (Q^*, \nu^*) \) along the whole sequence. Recall the canonical variables \( (\tilde{b}, b, \rho, \sigma) \) and the canonical filtration \( \{\mathcal{G}_t\} \) introduced in Section 8.3.1. We can find a countable collection \( \{\eta_j\}_{j=1}^\infty \) of continuous nonnegative functions with compact support in \( \mathbb{R}^d \), such that, denoting

\[
\rho^n_{\eta_j}(\xi) = \int_{\mathbb{R}^d \times [0,t]} \eta_j(y) \rho(dy \, ds)[\xi], \quad t \in [0,T], \ j \in \mathbb{N}, \ \xi \in \Xi
\]

and defining the stochastic process

\[
\mathcal{E}(t) = (\tilde{b}_t, b_t, \rho^n_{\eta_j}, \sigma)_{j \in \mathbb{N}}, \quad t \in [0,T]
\]

with sample paths in \( \mathcal{D}_\infty \doteq \mathcal{D}([0,T] : \mathbb{R}^\infty) \), we have \( \mathcal{G}_t = \sigma\{\mathcal{E}(\cdot \land t)\} \) for \( t \in [0,T] \).

We first verify (1). It suffices to check that for all \( 0 \leq s \leq t \leq T, \ g \in \mathcal{C}_c^2(\mathbb{R}^d) \) and \( f \in \mathcal{C}_b(\mathcal{D}_\infty) \)

\[
\mathbb{E}^* \left| \mathbb{E}^{Q^n} f(\mathcal{E}(\cdot \land s)) \left( g(b_t) - g(b_s) - \frac{1}{2} \int_s^t \Delta g(b_u) \, du \right) \right|^2 = 0 \quad (8.31)
\]

Since \( Q^n \) converges weakly to \( Q^* \) the left side equals

\[
\lim_{n \to \infty} \frac{1}{n^2} \mathbb{E} \left| \mathbb{E}^{Q^n} f(\mathcal{E}(\cdot \land s)) \left( g(b_t) - g(b_s) - \frac{1}{2} \int_s^t \Delta g(b_u) \, du \right) \right|^2
= \lim_{n \to \infty} \frac{1}{n^2} \mathbb{E} \left| \sum_{i=1}^n f(\mathcal{\mathcal{E}}^n_i(\cdot \land s)) \left( g(\mathcal{\beta}^n_i(t)) - g(\mathcal{\beta}^n_i(s)) - \frac{1}{2} \int_s^t \Delta g(\mathcal{\beta}^n_i(u)) \, du \right) \right|^2,
\]

where \( \mathcal{E}^n_i \) is defined similarly to \( \mathcal{E} \) by replacing \( (\tilde{b}, b, \rho, \sigma) \) with \( (\tilde{\beta}^n_i, b^n_i, \rho^n_i, S^n_i) \). Conditioning on \( \mathcal{F}_s \) and using the fact that \( \beta^n \) is a standard \( \mathcal{F}_t \)-Brownian motion, we see that cross product terms
do not contribute when the above squared sum is written as a double sum. So the above limit is 0 which proves (8.31).

Consider now (2). It suffices to show that

$$ E^* E^{Q^*} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, dt) < \infty. $$

Noting that $Q^n \Rightarrow Q^*$ as $n \to \infty$ and

$$ E E^{Q^n} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, dt) = \frac{1}{n} \sum_{i=1}^n E \int_0^T \|\psi_n^i(t)\|^2 \, dt \leq 2C_0, $$

it suffices to show that the function

$$ Q \mapsto E^* \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, dt) $$

is a lower semi-continuous function from $\mathcal{P}(\Xi)$ to $[0, \infty]$. This is immediate from Fatou’s lemma on observing that the function

$$ \xi = (\tilde{b}, b, \rho, \tilde{\sigma}) \mapsto \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) $$

is a lower semi-continuous function from $\Xi$ to $[0, \infty]$. This proves (2).

Next we verify (3). It suffices to show that for each $t \in [0,T]$

$$ E^* E^{Q^*} \left( \left| \tilde{b}_t - b_t - \int_{\mathbb{R}^d \times [0,t]} y \rho(dy \, ds) \right| \wedge 1 \right) = 0. $$

(8.33)

Note that for each $t \in [0,T]$, the function

$$ \xi = (\tilde{b}, b, \rho, \tilde{\sigma}) \mapsto \left| \tilde{b}_t - b_t - \int_{\mathbb{R}^d \times [0,t]} y \rho(dy \, ds) \right| \wedge 1 $$

is a continuous and bounded function from $\Xi$ to $\mathbb{R}_+$. Also,

$$ E E^{Q^n} \left( \left| \tilde{b}_t - b_t - \int_{\mathbb{R}^d \times [0,t]} y \rho(dy \, ds) \right| \wedge 1 \right) = \frac{1}{n} \sum_{i=1}^n E \left( \left| \tilde{b}_t^i - b_t^i - \int_0^t \psi_i^s(s) \, ds \right| \wedge 1 \right) = 0. $$
Combining the above two observations and recalling that \( Q^n \Rightarrow Q^* \) we have (8.33).

Finally we consider (4). By lower semi-continuity and convexity of the function \( R(\cdot\|\theta) \), we have, on using the first statement of the lemma that

\[
\mathbb{E}^* R(Q^*_n(\cdot\|\theta)) = \mathbb{E}^* R(\nu^*\|\theta) \leq \liminf_{n \to \infty} \mathbb{E} R(\nu^n\|\theta) \leq \liminf_{n \to \infty} \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^{n} R(\nu^n_i\|\theta) \right] \leq C_0.
\]

So \( R(Q^*_n(\cdot\|\theta)) < \infty \) and consequently \( Q^*_n(\cdot\|\theta) \ll \theta \) a.s. \( P^* \).

Using Lemma 8.9, we can now argue tightness of \( \{ \tilde{\mu}^n \} \).

**Lemma 8.10.** The sequence of random variables \( \{ \tilde{\mu}^n \}_{n \geq 1} \) is tight in \( \mathcal{D} \).

**Proof.** We first argue tightness of \( \{ \tilde{\mu}^n(t) \} \) in \( \mathcal{M}(\mathbb{R}^d) \) for each \( t \in [0, T] \). Note that

\[
G_4(m) = \int_{\mathbb{R}^d} \|x\|^2 m(dx), \quad m \in \mathcal{M}(\mathbb{R}^d)
\]

is a tightness function on \( \mathcal{M}(\mathbb{R}^d) \) and

\[
\mathbb{E} G_4(\tilde{\mu}^n(t)) = \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^{n} \|\tilde{\beta}^n_i(t)\|^2 \mathbf{1}_{\{S^n_i > f^n_0(\zeta, \tilde{\mu}^n(s)) \, ds\}} \right] \leq 2E \left[ \frac{1}{n} \sum_{i=1}^{n} \|\beta^n_i(t)\|^2 \right] + 2E \left[ \frac{1}{n} \sum_{i=1}^{n} \left( \int_0^t \|\psi^n_i(s)\|^2 \, ds \right) \right]^2 \leq 2dt + 2tE \left[ \frac{1}{n} \sum_{i=1}^{n} \int_0^t \|\psi^n_i(s)\|^2 \, ds \right] \leq 2dT + 4TC_0.
\]

So \( \{ \tilde{\mu}^n(t) \} \) is tight in \( \mathcal{M}(\mathbb{R}^d) \) for each \( t \in [0, T] \).

Next we proceed to consider fluctuations of \( \tilde{\mu}^n \). For \( \delta \in [0, T] \), let \( \mathcal{T}^{\delta,n} \) be the collection of all \( [0, T-\delta]-\text{valued stopping times} \) \( \tau \) on \( \Upsilon \). In order to argue tightness of \( \{ \tilde{\mu}^n \} \) in \( \mathcal{D} \), by Aldous-Kurtz tightness criterion (cf. [47], Theorem 2.7), it suffices to show that

\[
\limsup_{\delta \to 0} \limsup_{n \to \infty} \sup_{\tau \in \mathcal{T}^{\delta,n}} \mathbb{E} d_{BL}(\tilde{\mu}^n(\tau + \delta), \tilde{\mu}^n(\tau)) = 0. \tag{8.35}
\]

165
Note that for $\delta \in [0, T]$,

$$d_{BL}(\hat{\mu}^n(\tau + \delta), \mu^n(\tau))$$

$$= \sup_{\|f\|_{BL} \leq 1} \left| \langle f, \hat{\mu}^n(\tau + \delta) \rangle - \langle f, \mu^n(\tau) \rangle \right|$$

$$\leq \sup_{\|f\|_{BL} \leq 1} \frac{1}{n} \sum_{i=1}^{n} \left| f(\tilde{\beta}_i^n(\tau + \delta)) \mathbf{1}_{\{S_i^n > \int_{0}^{\tau + \delta} \langle \zeta, \mu^n(s) \rangle \, ds \}} - f(\tilde{\beta}_i^n(\tau)) \mathbf{1}_{\{S_i^n > \int_{0}^{\tau} \langle \zeta, \mu^n(s) \rangle \, ds \}} \right|$$

$$\leq \sup_{\|f\|_{BL} \leq 1} \frac{1}{n} \sum_{i=1}^{n} \left| f(\tilde{\beta}_i^n(\tau + \delta)) - f(\tilde{\beta}_i^n(\tau)) \right| \mathbf{1}_{\{S_i^n > \int_{0}^{\tau + \delta} \langle \zeta, \mu^n(s) \rangle \, ds \}}$$

$$+ \sup_{\|f\|_{BL} \leq 1} \frac{1}{n} \sum_{i=1}^{n} \left| f(\tilde{\beta}_i^n(\tau)) \left( \mathbf{1}_{\{S_i^n > \int_{0}^{\tau + \delta} \langle \zeta, \mu^n(s) \rangle \, ds \}} - \mathbf{1}_{\{S_i^n > \int_{0}^{\tau} \langle \zeta, \mu^n(s) \rangle \, ds \}} \right) \right|$$

$$\equiv T_1^n + T_2^n.$$

It then suffices to show that for $j = 1, 2$,

$$\lim_{\delta \to 0} \sup_{n \to \infty} \sup_{\tau \in T_{\delta,n}} \mathbf{E} T_j^n = 0. \quad (8.36)$$

For $T_1^n$, we have

$$T_1^n \leq \frac{1}{n} \sum_{i=1}^{n} \left\| \tilde{\beta}_i^n(\tau + \delta) - \tilde{\beta}_i^n(\tau) \right\| \leq \frac{1}{n} \sum_{i=1}^{n} \left( \|\beta_i^n(\tau + \delta) - \beta_i^n(\tau)\| + \int_{\tau}^{\tau + \delta} \|\psi^n_i(s)\| \, ds \right).$$

So by Cauchy–Schwarz inequality,

$$\sup_{\tau \in T_{\delta,n}} \mathbf{E} T_1^n \leq \sup_{\tau \in T_{\delta,n}} \left( \mathbf{E} \frac{1}{n} \sum_{i=1}^{n} \|\beta_i^n(\tau + \delta) - \beta_i^n(\tau)\|^2 \right)^{\frac{1}{2}} + \sqrt{\delta} \left( \mathbf{E} \frac{1}{n} \sum_{i=1}^{n} \int_{0}^{T} \|\psi^n_i(s)\|^2 \, ds \right)^{\frac{1}{2}}$$

$$\leq \sqrt{d\delta} + \sqrt{2C_0\delta},$$

which implies that (8.36) holds for $T_1^n$.

Now we consider $T_2^n$. Note that

$$T_2^n \leq \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\int_{0}^{\tau} \langle \zeta, \tilde{\mu}^n(s) \rangle \, ds < S_i^n \leq \int_{0}^{\tau + \delta} \langle \zeta, \tilde{\mu}^n(s) \rangle \, ds \}} = Q_4^n \left( \left( \int_{0}^{\tau} \langle \zeta, \tilde{\mu}^n(s) \rangle \, ds, \int_{0}^{\tau + \delta} \langle \zeta, \tilde{\mu}^n(s) \rangle \, ds \right) \right).$$
Since
\[ \left| \int_0^{\tau + \delta} \langle \zeta, \tilde{\mu}^n(s) \rangle \, ds - \int_0^{\tau} \langle \zeta, \tilde{\mu}^n(s) \rangle \, ds \right| \leq \| \zeta \|_\infty \delta, \quad \left| \int_0^{\tau + \delta} \langle \zeta, \tilde{\mu}^n(s) \rangle \, ds \right| \leq \| \zeta \|_\infty T, \]
it follows that
\[ \mathcal{T}_2^n \leq \max_{k=0,\ldots,[T/\delta]} Q^n_{(4)}([k\| \zeta \|_\infty \delta, (k+2)\| \zeta \|_\infty \delta]) \]
So we have
\[ \sup_{\tau \in \mathcal{T}^n} \mathbf{E} T_2^n \leq \mathbf{E} \max_{k=0,\ldots,[T/\delta]} Q^n_{(4)}([k\| \zeta \|_\infty \delta, (k+2)\| \zeta \|_\infty \delta]). \quad (8.37) \]
By Lemma 8.8 and 8.9, we can assume without loss of generality that \((Q^n, \nu^n)\) converges weakly along the whole sequence to \((Q^*, \nu^*)\), given on some probability space \((\Omega^*, F^*, P^*)\), with \(Q^* \in \mathcal{P}_\infty\) a.s. \(P^*\). It then follows from (8.37), Fatou’s lemma, property (4) of \(P^*_\infty\) and Remark 8.1 that
\[ \limsup_{n \to \infty} \sup_{\tau \in \mathcal{T}^n} \mathbf{E} T_2^n \leq \limsup_{n \to \infty} \mathbf{E} \max_{k=0,\ldots,[T/\delta]} Q^n_{(4)}([k\| \zeta \|_\infty \delta, (k+2)\| \zeta \|_\infty \delta]) \leq \mathbf{E}^* \max_{k=0,\ldots,[T/\delta]} Q^*_{(4)}([k\| \zeta \|_\infty \delta, (k+2)\| \zeta \|_\infty \delta]) \to 0 \]
as \(\delta \to 0\). Hence (8.36) also holds for \(\mathcal{T}_2^n\) and we have tightness of \(\{\tilde{\mu}^n\}\) in \(\mathcal{D}\). \qed

The following lemma gives important characterization of weak limit points of \(\{\tilde{\mu}^n\}\). Recall the definition of \(\varpi_\Theta\) for \(\Theta \in \mathcal{P}_\infty\) given in Section 8.3.2.

**Lemma 8.11.** Suppose (8.28) holds and \((Q^n, \nu^n, \tilde{\mu}^n)\) converges along a subsequence, in distribution, to \((Q^*, \nu^*, \tilde{\mu}^*)\) given on some probability space \((\Omega^*, F^*, P^*)\). Then we have \(P^*\)-a.s., \(\tilde{\mu}^* = \varpi_{Q^*}\), i.e.
\[ \langle f, \tilde{\mu}^*(t) \rangle = E^{Q^*} \left[ f(\tilde{b}_t) 1_{\{ \sigma > \tilde{f}_0(\zeta, \tilde{\mu}^*(s)) \} \, ds} \right], \quad \forall f \in C_b(\mathbb{R}^d), t \in [0, T]. \quad (8.38) \]

**Proof.** Without loss of generality assume that \((Q^n, \nu^n, \tilde{\mu}^n)\) converges weakly to \((Q^*, \nu^*, \tilde{\mu}^*)\) along the whole sequence. By appealing to Skorokhod representation we can further assume without loss of generality that the convergence holds a.s. Noting that as \(n \to \infty\),
\[ \sup_{t \in [0,1]} d_{BL}(\tilde{\mu}^n(t), \tilde{\mu}^n(t-)) = \sup_{t \in [0,1]} \sup_{\|f\|_{BL} \leq 1} |\langle f, \tilde{\mu}^n(t) \rangle - \langle f, \tilde{\mu}^n(t-) \rangle| \leq \frac{1}{n} \to 0 \]
a.s., we have $\mu^* \in \mathcal{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$ a.s. $P^*$. So $\langle f, \tilde{\mu}^n(t) \rangle \to \langle f, \tilde{\mu}^*(t) \rangle$ for all $f \in \mathbb{C}_b(\mathbb{R}^d)$ and $t \in [0, T]$. From (8.25) and (8.26) we have

$$\langle f, \tilde{\mu}^n(t) \rangle = E^{Q^n} \left[ f(\tilde{b}_t) 1_{\{\tilde{\sigma} > f^t_0 \langle \zeta, \tilde{\pi}^n(s) \rangle ds \}} \right] , \forall f \in \mathbb{C}_b(\mathbb{R}^d), t \in [0, T].$$

(8.39)

It then suffices to argue that RHS of (8.39) converges in probability to RHS of (8.38) for fixed $f \in \mathbb{C}_b(\mathbb{R}^d)$ and $t \in [0, T]$. For this, consider the following function $g : \mathcal{P}(\Xi) \times \mathcal{D} \to \mathbb{R}$, defined as

$$g(\Theta, \pi) = E^{\Theta} \left[ f(\tilde{b}_t) 1_{\{\tilde{\sigma} > f^t_0 \langle \zeta, \pi(s) \rangle ds \}} \right] , (\Theta, \pi) \in \mathcal{P}(\Xi) \times \mathcal{D}.$$  

(8.40)

We claim that any $(\Theta, \pi) \in \mathcal{P}_\infty \times \mathcal{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$ is a continuity point of the function $g$. Once the claim is verified, (8.38) follows by continuous mapping theorem.

To prove the claim, let $(\Theta^n, \pi^n) \to (\Theta, \pi)$ in $\mathcal{P}(\Xi) \times \mathcal{D}$ with $(\Theta, \pi) \in \mathcal{P}_\infty \times \mathcal{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$. Note that

$$|g(\Theta^n, \pi^n) - g(\Theta, \pi)| \leq |g(\Theta^n, \pi^n) - g(\Theta^n, \pi)| + |g(\Theta^n, \pi) - g(\Theta, \pi)|.$$  

(8.41)

By property (4) of $\mathcal{P}_\infty$ and Remark 8.1, we have

$$\lim_{n \to \infty} |g(\Theta^n, \pi) - g(\Theta, \pi)| = 0.$$

For the first term on the right hand side of (8.41), we have

$$|g(\Theta^n, \pi^n) - g(\Theta^n, \pi)| \leq E^{\Theta^n} \left| f(\tilde{b}_t) \left( 1_{\{\tilde{\sigma} > f^t_0 \langle \zeta, \pi^n(s) \rangle ds \}} - 1_{\{\tilde{\sigma} > f^t_0 \langle \zeta, \pi(s) \rangle ds \}} \right) \right| \leq \|f\|_{\infty}^{\Theta^n} \left( \tilde{\sigma} \text{ is between } \int_0^t \langle \zeta, \pi^n(s) \rangle ds \text{ and } \int_0^t \langle \zeta, \pi(s) \rangle ds \right).$$

Since $\pi^n \to \pi$ in $\mathcal{D}$ and $\pi \in \mathcal{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$, we have as $n \to \infty$,

$$\left| \int_0^t \langle \zeta, \pi^n(s) \rangle ds - \int_0^t \langle \zeta, \pi(s) \rangle ds \right| \leq \int_0^T |\langle \zeta, \pi^n(s) \rangle - \langle \zeta, \pi(s) \rangle| ds \to 0.$$

Also note that

$$\left| \int_0^t \langle \zeta, \pi^n(s) \rangle ds \right| \leq \|\zeta\|_\infty T, \quad \left| \int_0^t \langle \zeta, \pi(s) \rangle ds \right| \leq \|\zeta\|_\infty T.$$
So for any $\delta > 0$,

$$
\limsup_{n \to \infty} |g(\Theta^n, \pi^n) - g(\Theta^n, \pi)| \\
\leq \|f\|_{\infty} \limsup_{n \to \infty} \Theta^n \left( \sigma \text{ lies between } \int_0^t \langle \zeta, \pi_n(s) \rangle \, ds \text{ and } \int_0^t \langle \zeta, \pi(s) \rangle \, ds \right) \\
\leq \|f\|_{\infty} \max_{k=0, \ldots, \lfloor T/\delta \rfloor} \Theta^n(4) \left( [k\|\zeta\|_{\infty}\delta, (k+1)\|\zeta\|_{\infty}\delta] \right) \\
\leq \|f\|_{\infty} \max_{k=0, \ldots, \lfloor T/\delta \rfloor} \Theta^n(4) \left( [k\|\zeta\|_{\infty}\delta, (k+1)\|\zeta\|_{\infty}\delta] \right).
$$

Letting $\delta \to 0$ in the above display and using property (4) of $\mathcal{P}_\infty$ and Remark 8.1 gives us

$$
\limsup_{n \to \infty} |g(\Theta^n, \pi^n) - g(\Theta^n, \pi)| = 0. \text{ This proves } g(\Theta^n, \pi^n) \to g(\Theta, \pi) \text{ as } n \to \infty \text{ and therefore the claim holds.}
$$

8.6 Laplace upper bound

In this section we prove the Laplace upper bound (8.7):

$$
\liminf_{n \to \infty} -\frac{1}{n} \log \mathbb{E} \left[ \exp \left( -nF(\mu^n) \right) \right] \geq \inf_{\pi \in D} \{F(\pi) + I(\pi)\}.
$$

Let $\varepsilon \in (0, 1)$ be arbitrary. For each $n \in \mathbb{N}$, let $\Upsilon$ (with the associated $\Pi^n$ and $S^n$) and $\psi^n \in A_n(\Upsilon)$ be an $\varepsilon$-optimal control in (8.27), namely

$$
-\frac{1}{n} \log \mathbb{E} \left[ \exp \left( -nF(\mu^n) \right) \right] \geq \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n \left( R(\nu_i^n \| \theta) + \frac{1}{2} \int_0^T \|\psi_i^n(s)\|^2 \, ds \right) + F(\tilde{\mu}^n) \right] - \varepsilon, \quad (8.42)
$$

where $(\nu_i^n)_{i=1}^n$ and $\tilde{\mu}^n$ are as defined in (8.24) and (8.26). The above inequality implies that

$$
\sup_{n \in \mathbb{N}} \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n \left( R(\nu_i^n \| \theta) + \frac{1}{2} \int_0^T \|\psi_i^n(s)\|^2 \, ds \right) \right] \leq 2\|F\|_{\infty} + 1,
$$

i.e. equation (8.28) holds with $C_0 = 2\|F\|_{\infty} + 1$. Then with $Q^n, \nu^n$ defined as in (8.29) and (8.30), we have from (8.42) and Jensen’s inequality that

$$
-\frac{1}{n} \log \mathbb{E} \left[ \exp \left( -nF(\mu^n) \right) \right] \geq \mathbb{E} \left[ R(\nu^n \| \theta) + \mathbb{E}Q^n \left( \frac{1}{2} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) \right) + F(\tilde{\mu}^n) \right] - \varepsilon.
$$

169
It follows from Lemmas 8.8 and 8.10 that \( \{(Q^n, \nu^n, \tilde{\mu}^n)\} \) is tight. Assume without loss of generality that \( (Q^n, \nu^n, \tilde{\mu}^n) \) converges along the whole sequence weakly to \( (Q^*, \nu^*, \tilde{\mu}^*) \), given on some probability space \( (\Omega^*, F^*, P^*) \). By Lemmas 8.9 and 8.11, we have \( Q^* \in \mathcal{P}_\infty \), \( Q^*_4 = \nu^* \), and that \( \tilde{\mu}^* \) solves \( (8.38) \) a.s. \( P^* \). It then follows from Lemma 8.2 that \( \varpi_{Q^*} = \tilde{\mu}^* \) a.s. \( P^* \). Using the lower semi-continuity of the function \( R(\cdot \| \theta) \) and the map in \( (8.32) \), we have from Fatou’s lemma that

\[
\liminf_{n \to \infty} \mathbb{E} \left( \frac{1}{2} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) \right) \geq \mathbb{E}^* \left( \frac{1}{2} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) \right) = \mathbb{E}^*(J(Q^*)) \geq \mathbb{E}^*(I(\varpi_{Q^*})).
\]

Also note that \( F \in \mathcal{C}_b(\mathcal{D}) \), which implies

\[
\lim_{n \to \infty} \mathbb{E} F(\tilde{\mu}^n) = \mathbb{E}^* F(\tilde{\mu}^*) = \mathbb{E}^* F(\varpi_{Q^*}).
\]

Combining above three displays we have

\[
\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{E} [\exp (-nF(\mu^n))] \geq \mathbb{E}^* (F(\varpi_{Q^*}) + I(\varpi_{Q^*})) - \varepsilon \geq \inf_{\pi \in D} (F(\pi) + I(\pi)) - \varepsilon.
\]

Since \( \varepsilon > 0 \) is arbitrary, this completes the proof of the Laplace upper bound.

\[ 8.7 \text{ Laplace lower bound} \]

In this section we prove the **Laplace lower bound** \( (8.8) \):

\[
\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E} [\exp (-nF(\mu^n))] \leq \inf_{\pi \in D} \{F(\pi) + I(\pi)\}.
\]

Let \( \varepsilon > 0 \) be arbitrary and let \( \pi^* \in \mathcal{D} \) be such that

\[
F(\pi^*) + I(\pi^*) \leq \inf_{\pi \in D} \{F(\pi) + I(\pi)\} + \varepsilon.
\]
Next let $\Theta \in \mathcal{P}_\infty$ be such that $\varpi_\Theta = \pi^*$ and $\mathcal{J}(\Theta) \leq I(\pi^*) + \varepsilon$. We claim that without loss of generality one can assume that

$$\rho(A \times [0, t]) = \int_0^t \delta_A(\psi(s)) \, ds, \quad t \in [0, T], \; A \in \mathcal{B}(\mathbb{R}^d), \; \text{a.s.} \; \Theta, \quad (8.43)$$

where $\psi$ is an $\mathbb{R}^d$-valued $\mathcal{G}_t$-progressively measurable process such that

$$\mathbb{E}^\Theta \int_0^T \lVert \psi(s) \rVert^2 \, ds < \infty.$$ 

For this, define $\tilde{\psi}$ and $\tilde{\rho}$ on $\Xi$ as follows. For $t \in [0, T], \; A \in \mathcal{B}(\mathbb{R}^d)$ and $\xi \in \Xi$,

$$\tilde{\psi}(t)[\xi] = \int_{\mathbb{R}^d} y \rho(dy|t), \quad \tilde{\rho}(A \times [0, t]) = \int_0^t \delta_A(\psi(s)) \, ds.$$ 

Then

$$\int_{\mathbb{R}^d \times [0, t]} y \tilde{\rho}(dy \, ds) = \int_0^t \tilde{\psi}(s) \, ds = \int_{\mathbb{R}^d \times [0, t]} y \rho(dy \, ds) \quad (8.44)$$

and

$$\int_{\mathbb{R}^d \times [0, t]} \lVert y \rVert^2 \tilde{\rho}(dy \, ds) = \int_0^t \lVert \psi(s) \rVert^2 \, ds \leq \int_{\mathbb{R}^d \times [0, t]} \lVert y \rVert^2 \rho(dy \, ds). \quad (8.45)$$

Thus with $\tilde{\Theta} = \Theta \circ (\tilde{b}, \tilde{b}, \tilde{\rho}, \tilde{\sigma})^{-1}$, it follows from (8.44) and (8.45) that $\tilde{\Theta} \in \mathcal{P}_\infty$, $\varpi_\tilde{\Theta} = \varpi_\Theta = \pi^*$ and $\mathcal{J}(\tilde{\Theta}) \leq \mathcal{J}(\Theta) \leq I(\pi^*) + \varepsilon$. This proves the claim. Henceforth we assume (8.43).

We now construct a filtered probability space $\Upsilon \triangleq (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbf{P})$ as follows. Let $\tilde{\Omega} \triangleq \Xi^{\otimes \infty}$, $\tilde{\mathcal{F}} \triangleq \mathcal{B}(\Xi^{\otimes \infty})$, $\tilde{\mathbf{P}} \triangleq \Theta^{\otimes \infty}$. For $\xi \triangleq (\xi_i)_{i \in \mathbb{N}} \in \tilde{\Omega}$, where $\xi_i \triangleq (\tilde{b}_i, b_i, \rho_i, \tilde{\sigma}_i)$, define canonical variables as

$$\tilde{b}_i(\xi) \triangleq \tilde{b}_i, \quad b_i(\xi) \triangleq b_i, \quad \rho_i(\xi) \triangleq \rho_i, \quad \tilde{\sigma}_i(\xi) \triangleq \tilde{\sigma}_i, \quad i \in \mathbb{N}.$$ 

Let $\mathcal{F}_t \triangleq \sigma \{\tilde{b}_i(s), b_i(s), \rho_i(A \times [0, s]), \tilde{\sigma}_i, \; i \in \mathbb{N}, \; s \leq t, \; A \in \mathcal{B}(\mathbb{R}^d)\}$. Define

$$\psi_i(t) \triangleq \int_{\mathbb{R}^d} y \rho_i(dy|t), \quad i \in \mathbb{N}, \; t \in [0, T].$$

Note that by (8.43), under $\mathbf{P}$, $\rho_i(A \times [0, t]) = \int_0^t \delta_A(\psi_i(s)) \, ds$ for $A \in \mathcal{B}(\mathbb{R}^d), \; t \in [0, T]$ and $i \in \mathbb{N}$. With above choices of $\Upsilon$ and canonical random variables, let $(\boldsymbol{\nu}_i)^n_{i=1}$ and $\tilde{\mu}^n$ be defined as above Lemma 8.7, with $\Pi^n = \Theta_n^{\otimes n}$ and $(\boldsymbol{S}^n, \tilde{\beta}^n, \beta^n)$ replaced with $(\tilde{\sigma}^n, \tilde{b}^n, b^n)$, where $\tilde{\sigma}^n \triangleq (\tilde{\sigma}_i)^n_{i=1}$,
\( \tilde{b}^n \equiv (b_i)_{i=1}^n \) and \( b^n \equiv (b_i)_{i=1}^n \). Note in particular that \( \nu_i^n = \Theta(4) \) for each \( i = 1, \ldots, n \). It follows from Lemma 8.7 that

\[
- \frac{1}{n} \log E e^{-nF(\mu^n)} \leq \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n \left( R(\nu_i^n \| \theta) + \frac{1}{2} \int_0^T \| \psi_i(s) \|^2 ds \right) + F(\tilde{\mu}^n) \right]. \tag{8.46}
\]

Note that

\[
\mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n \left( R(\nu_i^n \| \theta) + \frac{1}{2} \int_0^T \| \psi_i(s) \|^2 ds \right) \right] \\
\leq R(\Theta(4) \| \theta) + \frac{1}{2n} \sum_{i=1}^n \mathbb{E} \left[ \int_{\mathbb{R}^d \times [0,T]} \| y \|^2 \rho_i(dy ds) \right] \\
= R(\Theta(4) \| \theta) + \mathbb{E}^\Theta \left[ \frac{1}{2} \int_{\mathbb{R}^d \times [0,T]} \| y \|^2 \rho(dy ds) \right] \\
= J(\Theta), \tag{8.47}
\]

where the first equality holds since \( \rho_i \) are i.i.d. under \( \bar{P} \). The above calculation shows that (8.28) holds with \( C_0 \) replaced by \( J(\Theta) < \infty \). Define \((Q^n, \nu^n)\) as in (8.29) and (8.30) (with \((S^n, \tilde{\beta}^n, \beta^n)\) replaced by \((\tilde{\sigma}^n, \tilde{b}^n, b^n)\) and \((\rho_i^n)_{i=1}^n\) replaced by \((\rho_i^n)_{i=1}^n\)). It follows from Lemmas 8.8 and 8.10 that \( \{(Q^n, \nu^n, \tilde{\mu}^n)\} \) is a tight sequence of \( \mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R}^d) \times \mathcal{D} \)-valued random variables. Suppose \((Q^*, \nu^*, \tilde{\mu}^*)\) is a weak limit point of the sequence given on some probability space \((\Omega^*, \mathcal{F}^*, P^*)\). From Lemmas 8.9 and 8.11 it follows that \( Q^* \in \mathcal{P}(\Xi), Q^*_v = \nu^* \) and \( \tilde{\mu}^* \) solves (8.38) \( P^*\)-a.s. By law of large numbers \( Q^* = \Theta \) a.s. \( P^* \) and hence by Lemma 8.2, we have \( \tilde{\mu}^* = \pi \Theta = \pi^* \) a.s. \( P^* \).

Combining above observations with (8.46) and (8.47) we have

\[
\limsup_{n \to \infty} - \frac{1}{n} \log E e^{-nF(\mu^n)} \leq J(\Theta) + F(\pi^*) \leq I(\pi^*) + F(\pi^*) + \varepsilon \leq \inf_{\pi \in \mathcal{D}} \{ F(\pi) + I(\pi) \} + 2\varepsilon,
\]

where the last two inequalities follow from the choices of \( \Theta \) and \( \pi^* \). Since \( \varepsilon > 0 \) is arbitrary, the desired Laplace lower bound follows.

8.8 \( I \) is a rate function.

In this section we will prove that the function \( I \) defined in (8.5) is a rate function, namely for each fixed \( K < \infty \), the set \( F_K = \{ \pi \in \mathcal{D} : I(\pi) \leq K \} \) is compact.
Let \( \{\pi_n\}_{n \in \mathbb{N}} \subset F_K \). Then for each \( n \in \mathbb{N} \) there exists \( \Theta^n \in \mathcal{P}_\infty \) such that \( \pi_{\Theta^n} = \pi_n \) and \( J(\Theta^n) \leq K + \frac{1}{n} \). In particular

\[
\sup_{n \in \mathbb{N}} \left( R(\Theta^n_\theta) + E^{\Theta^n_\theta} \left[ \frac{1}{2} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) \right] \right) \leq K + 1.
\]

Using the above bound we now argue that \( \{\Theta^n\} \) is tight. The proof is similar to that of Lemma 8.8 and so we only provide a sketch. Note that \( \Theta^n_\theta \) is the \( d \)-dimensional Wiener measure for each \( n \) and so the tightness of \( \{\Theta^n_\theta\} \) is immediate. Next, recall tightness functions \( G_j \) for \( j = 1, 2, 3 \) in the proof of Lemma 8.8. Since \( G_1(\Theta^n_\theta) = E^{\Theta^n_\theta} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) \leq 2(K + 1) \), \( \{\Theta^n_\theta\} \) is tight. To argue tightness of \( \{\Theta^n_\theta\} \), define canonical variable \( u: \Xi \to \mathcal{C} \) as \( u(t) = \int_{\mathbb{R}^d \times [0,t]} y \rho(dy \, ds) \) for \( t \in [0,T] \). It follows from Cauchy–Schwarz inequality that \( G_2(\Theta^n_\theta \circ u^{-1}) = E^{\Theta^n_\theta} \int_0^T \|\dot{u}(s)\|^2 ds \leq E^{\Theta^n_\theta} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) \leq 2(K + 1) \). So \( \{\Theta^n_\theta \circ u^{-1}\} \) is tight and consequently \( \{\Theta^n \circ (b,u)^{-1}\} \) is also tight. This implies tightness of \( \{\Theta^n\} \) since \( \Theta^n \circ (b,u)^{-1} = \Theta^n \circ (b,u)^{-1} \circ g_3^{-1} \) and \( g_3: \mathbb{C} \times \mathbb{C} \to \mathbb{C} \) defined as \( g_3(x,u) = x + u \) is continuous. Finally, tightness of \( \{\Theta^n_\theta\} \) follows since \( G_3(\Theta^n_\theta) = R(\Theta^n_\theta) \leq K + 1 \). This proves the tightness of \( \{\Theta^n\} \).

Now suppose \( \Theta^n \) converges along a subsequence (labeled as \( \{n\} \) for simplicity) to \( \Theta \in \mathcal{P}(\Xi) \). Using arguments similar to those in the proof of Lemma 8.9, we now show that \( \Theta \in \mathcal{P}_\infty \) and \( J(\Theta) < \infty \). We need to verify that properties (1)–(4) of Section 8.3.2 are satisfied. It is clear that (1) holds. Using the lower semi-continuity of the function \( R(\cdot) \) and the map in (8.32), we have

\[
J(\Theta) = R(\Theta_\theta) + E^\Theta \left[ \frac{1}{2} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) \right] \leq \liminf_{n \to \infty} \left( R(\Theta^n_\theta) + E^{\Theta^n_\theta} \left[ \frac{1}{2} \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) \right] \right) \leq K. \tag{8.48}
\]

This implies (2), (4) and that \( J(\Theta) < \infty \). Finally, noting that the map (8.34) is bounded and continuous, we have

\[
E^\Theta \left( \left| \bar{b}_t - b_t - \int_{\mathbb{R}^d \times [0,t]} y \rho(dy \, ds) \right| \wedge 1 \right) = \lim_{n \to \infty} E^{\Theta^n_\theta} \left( \left| \bar{b}_t - b_t - \int_{\mathbb{R}^d \times [0,t]} y \rho(dy \, ds) \right| \wedge 1 \right) = 0
\]

for each \( t \in [0,T] \), which verifies (3). Thus we have shown that \( \Theta \in \mathcal{P}_\infty \) and \( J(\Theta) < \infty \).
By Lemma 8.2, there exists a unique solution $\varpi_\Theta$ of (8.4) that belongs to $\mathcal{C}([0,T] : \mathcal{M}({\mathbb{R}^d}))$. From (8.48) it follows that $I(\varpi_\Theta) \leq J(\Theta) \leq K$, and thus $\varpi_\Theta \in F_K$. Recall $\varpi_\Theta = \pi_n$. Thus to complete the proof, it suffices to show that $\varpi_{\Theta^n} \to \varpi_\Theta$ as $n \to \infty$.

We will now use arguments similar to those in Lemma 8.10 to show that $\varpi_\Theta$ is pre-compact in $\mathcal{C}([0,T] : \mathcal{M}({\mathbb{R}^d}))$. First we argue tightness of $\{\varpi_{\Theta^n}(t)\}$ in $\mathcal{M}({\mathbb{R}^d})$ for each $t \in [0,T]$. For this, recall that $G_4(m) = \int_{\mathbb{R}^d} \|x\|^2 \, m(dx)$ defined in the proof of Lemma 8.10 is a tightness function on $\mathcal{M}({\mathbb{R}^d})$. Also, for every $n \in \mathbb{N}$,

$$G_4(\varpi_{\Theta^n}(t)) = E^{\Theta^n} \left[ \|\tilde{b}(t)\|^2 \mathbf{1}_{\{\hat{\sigma} > f_0^1(\zeta, \varpi_{\Theta^n}(s)) \, ds\}} \right] \leq 2E^{\Theta^n} \left[ \|\tilde{b}(t)\|^2 \right] + 2TE^{\Theta^n} \left[ \int_{\mathbb{R}^d \times [0,T]} \|y\|^2 \rho(dy \, ds) \right] \leq 2dT + 4T(K + 1).$$

This proves the tightness of $\{\varpi_{\Theta^n}(t)\}$ for each $t \in [0,T]$. Next we proceed to consider fluctuations.

Fix $\delta > 0$ and consider $0 \leq t_1 \leq t_2 \leq T$ with $|t_1 - t_2| < \delta$. Note that

$$d_{BL}(\varpi_{\Theta^n}(t_1), \varpi_{\Theta^n}(t_2))$$

$$= \sup_{\|f\|_{BL} \leq 1} |(f, \varpi_{\Theta^n}(t_1)) - (f, \varpi_{\Theta^n}(t_2))|$$

$$\leq \sup_{\|f\|_{BL} \leq 1} E^{\Theta^n} \left| f(\tilde{b}(t_1)) \mathbf{1}_{\{\hat{\sigma} > f_0^1(\zeta, \varpi_{\Theta^n}(s)) \, ds\}} - f(\tilde{b}(t_2)) \mathbf{1}_{\{\hat{\sigma} > f_0^2(\zeta, \varpi_{\Theta^n}(s)) \, ds\}} \right|$$

$$\leq \sup_{\|f\|_{BL} \leq 1} E^{\Theta^n} \left| f(\tilde{b}(t_1)) - f(\tilde{b}(t_2)) \right| \mathbf{1}_{\{\hat{\sigma} > f_0^1(\zeta, \varpi_{\Theta^n}(s)) \, ds\}}$$

$$+ \sup_{\|f\|_{BL} \leq 1} E^{\Theta^n} \left| f(\tilde{b}(t_2)) \left( \mathbf{1}_{\{\hat{\sigma} > f_0^1(\zeta, \varpi_{\Theta^n}(s)) \, ds\}} - \mathbf{1}_{\{\hat{\sigma} > f_0^2(\zeta, \varpi_{\Theta^n}(s)) \, ds\}} \right) \right|$$

$$\leq \tilde{T}_1^n + \tilde{T}_2^n.$$
It then follows from Cauchy–Schwarz inequality that

\[
\sup_{|t_1 - t_2| < \delta} \mathcal{T}^n_1 \leq \sup_{|t_1 - t_2| < \delta} \left( \mathbb{E}^{\Theta^n} \left\| b(t_1) - b(t_2) \right\|^2 \right)^{\frac{1}{2}} + \sqrt{\delta} \left( \mathbb{E}^{\Theta^n} \int_{[0, T]} |y|^2 \rho(dy \, ds) \right)^{\frac{1}{2}} \\
\leq \sqrt{d\delta} + \sqrt{2(K + 1)\delta} \to 0,
\]
as \delta \to 0. For \(\mathcal{T}^n_2\), observing that

\[
\left| \int_0^{t_1} \langle \zeta, \varpi_{\Theta^n}(s) \rangle \, ds - \int_0^{t_2} \langle \zeta, \varpi_{\Theta^n}(s) \rangle \, ds \right| \leq \|\zeta\|_\infty \delta,
\]
we have

\[
\sup_{|t_1 - t_2| < \delta} \mathcal{T}^n_2 \leq \sup_{|t_1 - t_2| < \delta} \Theta^n \left( \int_0^{t_1} \langle \zeta, \varpi_{\Theta^n}(s) \rangle \, ds < \tilde{\sigma} \leq \int_0^{t_2} \langle \zeta, \varpi_{\Theta^n}(s) \rangle \, ds \right) \\
\leq \max_{k=0, \ldots, \lfloor T/\delta \rfloor} \Theta^n(4)([k\|\zeta\|_\infty \delta, (k + 2)\|\zeta\|_\infty \delta]).
\]

Since \(\Theta^n \to \Theta \in \mathcal{P}_\infty\), it then follows from Fatou’s lemma, property (4) of \(\mathcal{P}_\infty\) and Remark 8.1 that

\[
\limsup_{n \to \infty} \sup_{|t_1 - t_2| < \delta} \mathcal{T}^n_2 \leq \limsup_{n \to \infty} \max_{k=0, \ldots, \lfloor T/\delta \rfloor} \Theta^n(4)([k\|\zeta\|_\infty \delta, (k + 2)\|\zeta\|_\infty \delta]) \to 0,
\]
as \delta \to 0. Combining above two convergence results of \(\mathcal{T}^n_1\) and \(\mathcal{T}^n_2\) gives us pre-compactness of \(\{\varpi_{\Theta^n}\}\) in \(\mathcal{C}([0, T] : \mathcal{M}(\mathbb{R}^d))\).

Finally, let \(\pi^*\) be any limit point of \(\{\varpi_{\Theta^n}\}\) in \(\mathcal{C}([0, T] : \mathcal{M}(\mathbb{R}^d))\) along a subsequence (labeled as once again \(\{n\}\)). Then (along a further subsequence) we have \((\Theta^n, \varpi_{\Theta^n}) \to (\Theta, \pi^*)\) in \(\mathcal{P}(\Xi) \times \mathcal{C}([0, T] : \mathcal{M}(\mathbb{R}^d))\). Note that

\[
\langle f, \varpi_{\Theta^n}(t) \rangle = \mathbb{E}^{\Theta^n} \left[ f(h_t) \mathbf{1}_{\tilde{\sigma} > f_t'(\zeta, \varpi_{\Theta^n}(s)) \, ds} \right], \forall f \in \mathcal{C}_b(\mathbb{R}^d), t \in [0, T].
\]

Taking \(n \to \infty\) in the above display and using the property that \((\Theta, \pi^*) \in \mathcal{P}_\infty \times \mathcal{C}([0, T] : \mathcal{M}(\mathbb{R}^d))\) is a continuity point of the function \(g\) defined in (8.40), we have by continuous mapping theorem.
that the above display holds with \((\Theta^n, \varpi^n)\) replaced by \((\Theta, \pi^*)\). Hence \(\pi^* = \varpi_\Theta\) by Lemma 8.2.

This implies that \(\pi^n = \varpi_\Theta \rightarrow \varpi_\Theta \in F_K\) and completes the proof that \(I\) is a rate function. \(\square\)

### 8.9 Sketch of proof of Theorem 8.1

Fix \(T > 0\). Let \(\Gamma^n_T = \frac{1}{n} \sum_{i=1}^n \delta_{(B_i(-\wedge T), X_i)}\). It is clear that \(\mathcal{L}(\Gamma^n_T, \mu^n_T) = \mathcal{L}(Q^n_{(1, 4)}, \tilde{\mu}^n)\), where \(Q^n\) and \(\tilde{\mu}^n\) are as defined in (8.29) and (8.26) on some filtered probability space \(\Upsilon\), with \(\psi_i^n \equiv 0\) and \(\Pi^n = \theta^{\otimes n}\). Note that the bound in (8.28) holds trivially in this case. It follows from Lemmas 8.8–8.11 that \(\{(Q^n, \tilde{\mu}^n)\}_{n \in \mathbb{N}}\) is tight in \(\mathcal{P}(S_T) \times D_T\), and any weak limit point \((Q^*, \tilde{\mu}^*)\) is almost surely such that \(Q^*_{(1, 4)} = \mathcal{L}(B_1(- \wedge T)) \otimes \theta, \tilde{\mu}^* \in \mathcal{C}([0, T], \mathcal{M}(\mathbb{R}^d))\) and

\[
\langle f, \tilde{\mu}^*(t) \rangle = \mathbf{E}^{Q^*} \left[ f(\tilde{b}_t) \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \tilde{\mu}^*(s) \rangle \, ds\}} \right], \quad \forall f \in \mathcal{C}_b(\mathbb{R}^d), t \in [0, T].
\]

It is easily checked that above equation holds with \(\tilde{\mu}^*\) replaced by \(\mu_T = \{\mu(t)\}_{t \in [0, T]}\), where \(\mu(\cdot)\) is as defined below (8.3). By Lemma 8.2, we must have \(\tilde{\mu}^* = \mu_T\). This implies that \(\mu^n_T \Rightarrow \mu_T\) for each \(T > 0\). The result follows. \(\square\)
REFERENCES


