1

PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

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

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

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

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

AVIS

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

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

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

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

Diffusion Processes with Random Interactions

by

Xinhong Ding

M.Sc., M.Sc.

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

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

Carleton University

Ottawa, Ontario

November 1992

©

1992, Xinhong Ding

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Votre référence*

Our file *Notre référence*

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

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

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

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

ISBN 0-315-84041-2

Canada

Name

Xinhong Ding

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

Mathematics

SUBJECT TERM

0405

U·M·I

SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS

Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0455

EDUCATION

General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523

Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS

Language 0679
General 0289
Ancient 0290
Linguistics 0291
Modern 0291
Literature 0401
General 0294
Classical 0295
Comparative 0297
Medieval 0298
Modern 0316
African 0591
American 0305
Asian 0342
Canadian (English) 0355
Canadian (French) 0593
English 0311
Germanic 0312
Latin American 0315
Middle Eastern 0313
Romance 0314
Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY

Philosophy 0422
Religion 0318
General 0321
Biblical Studies 0319
Clergy 0320
History of 0322
Philosophy of 0469
Theology 0323

SOCIAL SCIENCES

American Studies 0323
Anthropology 0324
Archaeology 0326
Cultural 0327
Physical 0310
Business Administration 0272
General 0770
Accounting 0454
Banking 0338
Management 0385
Marketing 0501
Canadian Studies 0503
Economics 0505
General 0508
Agricultural 0509
Commerce-Business 0510
Finance 0511
History 0358
Labor 0366
Theory 0351
Folklore 0578
Geography 0578
Gerontology 0578
History 0578
General 0578

Ancient 0579
Medieval 0581
Modern 0582
Black 0328
African 0331
Asia, Australia and Oceania 0332
Canadian 0334
European 0335
Latin American 0336
Middle Eastern 0333
United States 0337
History of Science 0585
Law 0398
Political Science 0615
General 0616
International Law and 0617
Relations 0616
Public Administration 0814
Recreation 0452
Social Work 0626
Sociology 0627
Criminology and Penology 0938
Demography 0631
Ethnic and Racial Studies 0628
Individual and Family 0629
Studies 0630
Industrial and Labor 0700
Relations 0344
Public and Social Welfare 0709
Social Structure and 0999
Development 0453
Theory and Methods 0709
Transportation 0999
Urban and Regional Planning 0453
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES

Agriculture 0473
General 0285
Agronomy 0475
Animal Culture and 0476
Nutrition 0359
Animal Pathology 0478
Food Science and 0479
Technology 0817
Forestry and Wildlife 0777
Plant Culture 0746
Plant Pathology 0306
Plant Physiology 0287
Range Management 0308
Wood Technology 0309
Botany 0379
Cell 0329
Ecology 0353
Entomology 0369
Genetics 0792
Limnology 0416
Microbiology 0307
Molecular 0317
Neuroscience 0416
Oceanography 0433
Physiology 0821
Radiation 0778
Veterinary Science 0472
Zoology 0786
Biophysics 0760
General 0786
Medical 0760

EARTH SCIENCES

Biogeochemistry 0425
Geochemistry 0996

Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0412
Paleozoology 0985
Palynology 0427
Physical Geography 0344
Physical Oceanography 0414

HEALTH AND ENVIRONMENTAL SCIENCES

Environmental Sciences 0748
Health Sciences 0544
General 0544
Audiology 0330
Chemotherapy 0992
Dentistry 0547
Education 0350
Hospital Management 0749
Human Development 0758
Immunology 0542
Medicine and Surgery 0544
Mental Health 0547
Nursing 0549
Nutrition 0571
Obstetrics and Gynecology 0340
Occupational Health and 0354
Therapy 0381
Ophthalmology 0571
Pathology 0419
Pharmacology 0572
Pharmacy 0382
Physical Therapy 0573
Public Health 0574
Radiology 0575
Recreation 0575

Speech Pathology 0460
Toxicology 0383
Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences

Chemistry 0485
General 0749
Agricultural 0486
Analytical 0487
Biochemistry 0488
Inorganic 0738
Nuclear 0490
Organic 0491
Pharmaceutical 0494
Physical 0495
Polymer 0754
Radiation 0405
Mathematics 0405
Physics 0405
General 0986
Acoustics 0606
Astronomy and 0608
Astrophysics 0748
Atmospheric Science 0607
Atomic 0798
Electronics and Electricity 0759
Elementary Particles and 0609
High Energy 0610
Fluid and Plasma 0752
Molecular 0756
Nuclear 0611
Optics 0463
Radiation 0463
Solid State 0463
Statistics 0463

Applied Sciences

Applied Mechanics 0346
Computer Science 0984

Engineering 0537
General 0538
Aerospace 0539
Agricultural 0540
Automotive 0541
Biomedical 0542
Chemical 0543
Civil 0544
Electronics and Electrical 0348
Heat and Thermodynamics 0545
Hydraulic 0546
Industrial 0547
Marine 0794
Materials Science 0548
Mechanical 0743
Metallurgy 0551
Mining 0552
Nuclear 0549
Packaging 0765
Petroleum 0554
Sanitary and Municipal 0790
System Science 0428
Geotechnology 0796
Operations Research 0795
Plastics Technology 0994
Textile Technology 0994

PSYCHOLOGY

General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0340
Psychometrics 0451
Social 0451

**The undersigned hereby recommend to
The Faculty of Graduate Studies and Research
acceptance of the thesis,
Diffusion Processes With Random Interactions**

**submitted by
Xinhong Ding
in partial fulfilment of the requirements
for the degree of Doctor of Philosophy**

Chairman of the Department

Thesis Supervisor

External Examiner

**Carleton University
Department of Mathematics and Statistics**

Abstract

In this thesis we study the asymptotic behavior of certain systems of n randomly interacting diffusion processes in the limit when n goes to infinity. The first main objective of this thesis is to obtain a McKean-Vlasov type limit for the empirical measures associated to such systems in the almost sure sense with respect to the random interactions. The second main objective of this thesis is to study the equilibrium behaviors of such random systems in the limit when n goes to infinity. We establish the asymptotic behaviors of certain random Gibbs measures, which can be viewed as a mathematical rigorous version of the van Hemmen et al spin glass model. Finally, we consider a system in which the random interactions are assumed to be independent Brownian motions. Under proper conditions the limiting process has a unique invariant measure which is a zero mean Gaussian measure such that its variance solves the famous Sherrington-Kirkpatrick spin glass fixed point equation.

Acknowledgements

I express my greatest gratitude to Professor Donald A. Dawson for his guidance, patience, encouragement and financial support over the years. Also, I would like to thank all the teachers in our department and the Department of Mathematics at University of Ottawa (especially Professor M.Csörgő and Professor G.Ivanoff) who have taught me mathematics. I am very grateful to the Faculty of Graduate Studies and Research of Carleton University and the Department of Mathematics and Statistics of Carleton University for providing me the scholarships and the teaching assistant opportunities throughout the program. I thank the members of the Examining Board, in particular, Dr. G.Giroux from Universite de Sherbrooke, Dr. N.U.Ahmed from University of Ottawa, and Dr. L.Resnick from Carleton University, for their careful reading of the thesis and providing several corrections and suggestions. Among the friends I am particularly thankful to Jim, Ellie and Mrs Pekelsky for their encouragement and support in various ways. Finally, I thank my family for their understanding and moral support which made this work possible.

Table of Contents

Acceptance sheet	ii
Abstract	iii
Acknowledgements	iv
Introduction	vii
 Chapter 1. Preliminaries	 1
1.1. Markov diffusion processes	2
1.2. Weak convergence of probability measures	10
1.3. Large deviations and statistical mechanics	14
 Chapter 2. Spin Glass Problem	 22
2.1. Physical model	23
2.2. Dynamical mean field theory	29
2.3. Related mathematical work	34
 Chapter 3. Stochastic Models	 42
3.1. Interacting diffusion systems	42
3.2. Local empirical measures	48
3.3. A non-linear Markov process	52
 Chapter 4. McKean-Vlasov Type Limit	 61
4.1. Statement of the result	62
4.2. Weak convergence of $\mathcal{L}(X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)})$	63
4.3. Weak convergence of $\mathcal{L}(X_{\xi}^{(n)})$	75
4.4. A nonlinear neural network model	79

Chapter 5. Asymptotic Behavior of Random Gibbs Measures	83
5.1. Statement of the results	85
5.2. μ -a.s. large deviations	86
5.3. Varadhan type theorem	96
5.4. Proof of theorem 5.1	100
5.5. Proof of theorem 5.21	104
Chapter 6. Gaussian Limit	110
6.1. Statement of the result	114
6.2. Existence and uniqueness	116
6.3. McKean-Vlasov limit	120
6.4. Stationary Gaussian solution	127
Bibliography	131

Introduction

In this thesis we study the asymptotic behaviors of some randomly interacting diffusion processes on R in the limit when the number of processes tends to infinity. This research is motivated by spin glass type random models from physics and engineering, to which the theoretical studies (mainly mean field theory) at a physical level of rigor exposed many interesting but subtle and difficult problems and, therefore, created considerable research interest in recent years (see chapter 2 for a short introduction to the spin glass problem. A much more detailed review can be found in the book by Mézard et al [33]).

From an equilibrium statistical physics point of view, a typical spin glass model is usually described by the following n -particle random Hamiltonian:

$$H_n^{(\theta_{i,j})}(x_1, \dots, x_n) = -\frac{1}{2} \sum_{i,j=1}^n \theta_{i,j} x_i x_j \quad (0.1)$$

where $x_1, \dots, x_n$ denote n continuous spins in R and $\theta_{i,j}$ is a set of real valued random variables defined on some probability space. It is well known that the corresponding Gibbs distribution for the Hamiltonian (0.1) at inverse temperature $\beta = \frac{1}{T}$ is given by

$$P_{n,\beta}^{(\theta_{i,j})}(dx_1, \dots, dx_n) = \frac{1}{Z_{n,\beta}^{(\theta_{i,j})}} \exp\{-\beta H_n^{(\theta_{i,j})}(x)\} \prod_{i=1}^n \mu(dx_i), \quad (0.2)$$

where μ is a probability measure on R representing the distribution of a single spin without the interactions, and $Z_{n,\beta}^{(\theta_{i,j})}$ is a normalization factor, which is usually called

the partition function, given by

$$Z_{n,\beta}^{(\theta_{i,j})} = \int_{R^n} \exp\{-\beta H_n^{(\theta_{i,j})}(x)\} \prod_{i=1}^n \mu(dx_i) \quad (0.3)$$

The central problem in equilibrium statistical physics is to determine the free energy function per particle $f(\beta)$ in the thermodynamic limit $n \rightarrow \infty$

$$f(\beta) = \lim_{n \rightarrow \infty} \frac{1}{n} f_n^{(\theta_{i,j})}(\beta) \quad (0.4)$$

if this limit exists in some sense, where $f_n^{(\theta_{i,j})}(\beta)$ is the free energy of the n -particle system defined by

$$-\beta f_n^{(\theta_{i,j})}(\beta) = \log Z_{n,\beta}^{(\theta_{i,j})} \quad (0.5)$$

Unlike the usual mean field models in equilibrium statistical physics such as the Curie-Weiss model (see section 3 of chapter 1 for some standard results), the Gibbs distribution in a spin glass model depends on the realization of the random interactions $\{\theta_{i,j}\}$ and therefore becomes a random probability measure. Similarly the partition function $Z_{n,\beta}^{(\theta_{i,j})}$ becomes a random variable. This immediately raises the following fundamental questions on the equilibrium statistical mechanics of spin glass models: Does the thermodynamic limit (0.4) exist in some sense? If it does what is the limit $f(\beta)$ for a model with a random Hamiltonian such as (0.1)?

For general random interactions $\{\theta_{i,j}\}$, it is not always possible to give a mathematically satisfactory answer to this question. For example, it is believed that for the Sherrington-Kirkpatrick spin glass model [41] where the $\{\theta_{i,j}\}$ are i.i.d. Gaussian random variables with mean zero and variance $\frac{1}{n}$, the free energy per spin may not be the same for different samples of $\{\theta_{i,j}\}$. In that case, one first has to average the free energy $f_n^{(\theta_{i,j})}(\beta)$ of the n -particle system with respect to the Gaussian distribution of the random interactions and then take the thermodynamic limit $n \rightarrow \infty$ to obtain the free energy per spin. We will discuss this in some details in chapter 2.

However, for certain properly chosen random variables $\{\theta_{i,j}\}$, the free energy in the formal expression (0.4) does exist for almost all realizations $\{\theta_{i,j}\}$ and can be evaluated by solving a variational problem. One such example is the van Hemmen spin glass model [52], in which the random couplings $\{\theta_{i,j}\}$ assume the form

$$\theta_{i,j} = \frac{1}{n} c(\xi_i, \xi_j), \quad i, j = 1, \dots, n \quad (0.6)$$

where $\xi_1, \dots, \xi_n$ are n i.i.d. R^q -valued random vectors defined on some probability space, and $c(x, y)$ is a symmetric real valued function defined on $R^q \times R^q$. For simplicity, we assume that the random variables ξ_i take only p distinct values $\theta_1, \dots, \theta_p$ in R^q with probabilities $\lambda_1, \dots, \lambda_p$, respectively. The form of $\{\theta_{i,j}\}$ as in (0.6) represents a rather wide class of random models in practice including the Hopfield neural network model [24] in which $c(\xi_i, \xi_j)$ is defined according to the so-called Hebb's rule:

$$c(\xi_i, \xi_j) = \sum_{k=1}^q \xi_i^{(k)} \xi_j^{(k)} = \xi_i \cdot \xi_j \quad (0.7)$$

where $\xi_i \cdot \xi_j$ denotes the inner product in the vector space R^q . One of the applications of the Hopfield neural network is to act as an associative memory model in which the p vectors $\theta_1, \dots, \theta_p$ in R^q correspond to p stored patterns and under some appropriate dynamics the network system with the random interconnection $\{\theta_{i,j}\}$ defined by (0.6) and (0.7) will evolve towards one of these patterns as time goes on.

In order to study the equilibrium behavior of the spin glass model described by the Hamiltonian (0.1) with $\{\theta_{i,j}\}$ as in (0.6) van Hemmen et al introduced in [52] p local magnetization order parameters which, in our notation $m_{\xi,1}^{(n)}, \dots, m_{\xi,p}^{(n)}$, can essentially be defined as

$$m_{\xi,k}^{(n)} = \frac{1}{n} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n x_i, \quad k = 1, \dots, p \quad (0.8)$$

where $\xi = (\xi_1, \xi_2, \dots)$.

Denote by $Q_{n,\xi}(dm)$ the image measure on R^p of the product probability measure $\prod_{i=1}^n \mu(dx_i)$ on R^n under the mapping (0.8), i.e.,

$$Q_{n,\xi}(dm) = \left(\prod_{i=1}^n \mu(dx_i) \right) \{ (x_1, \dots, x_n) : (m_{\xi,1}^{(n)}, \dots, m_{\xi,p}^{(n)}) \in dm \}. \quad (0.9)$$

Then the right hand side of (0.5) can now be expressed as

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,\beta}^{(\theta, \cdot)} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{R^n} \exp \{ -\beta H_n^{(\theta, \cdot)}(x) \} \prod_{i=1}^n \mu(dx_i) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{R^n} \exp \left\{ \frac{1}{2} \beta n \sum_{k,l=1}^p c(\theta_k, \theta_l) \left(\frac{1}{n} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n x_i \right) \left(\frac{1}{n} \sum_{\substack{j=1 \\ \xi_j = \theta_l}}^n x_j \right) \right\} \prod_{i=1}^n \mu(dx_i) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{R^p} \exp \left\{ \frac{1}{2} \beta n \sum_{k,l=1}^p c(\theta_k, \theta_l) m_k m_l \right\} Q_{n,\xi}(dm) \end{aligned} \quad (0.10)$$

where $m = (m_1, \dots, m_p)$. This suggests, as was proposed by van Hemmen et al in [52] using a heuristic Laplace argument, that the limit in (0.10) exists for almost all realizations of the random interactions $\{\xi_i\}$ and the free energy per spin $f(\beta)$ in the expression (0.4) can be evaluated according to the following variational formula:

$$\begin{aligned} f(\beta) &= -\frac{1}{\beta} \sup_{m \in R^p} \left\{ \frac{1}{2} \beta \sum_{k,l=1}^p c(\theta_k, \theta_l) m_k m_l - I(m) \right\} \\ &= \inf_{m \in R^p} \left\{ \frac{1}{\beta} I(m) - \frac{1}{2} \sum_{k,l=1}^p c(\theta_k, \theta_l) m_k m_l \right\} \end{aligned} \quad (0.11)$$

where $I(m)$ is a function related to the large deviation rate of the random probability measures $(Q_{n,\xi})_{n=1}^\infty$.

While the van Hemmen et al's work is heuristically correct, the justification of (0.11) needs rigorous proof. From the probability theory point of view the heuristic arguments above can be abstracted into the following large deviation problem:

1. Show that the random probability measures $(Q_{n,\xi})_{n=1}^\infty$ form a large deviation system for almost all realizations of the random variable ξ with a non-random rate function $I(m)$.

2. Show that the Varadhan type asymptotic formula for evaluating certain integral holds in the almost sure sense, i.e. for almost all realizations of the random variables ξ

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{R^p} \exp\{n\Phi(m)\} Q_{n,\xi}(dm) = \inf_{m \in R^p} \{\Phi(m) - I(m)\} \quad (0.12)$$

where $\Phi(m)$ is a suitable continuous function defined on R^p .

One of the main objectives of this thesis is to carry out a rigorous proof for the heuristic calculations above. To do that we need to modify the classical large deviation theory so that one can study the large deviations in an almost sure sense. This was done in chapter 5. In particular, we consider the asymptotic behavior of the random Gibbs measures defined by

$$P_{n,\xi}(dx_1, \dots, dx_n) = \frac{1}{Z_{n,\xi}} \exp\left\{\frac{\beta}{2n} \sum_{i,j=1}^n c(\xi_i, \xi_j) x_i x_j\right\} \prod_{i=1}^n \mu(dx_i), \quad (0.13)$$

where $\xi = (\xi_1, \xi_2, \dots)$ is a sequence of i.i.d. $\{\theta_1, \dots, \theta_p\} \subset R^q$ -valued random variables assuming probabilities $\lambda_1, \dots, \lambda_p$ respectively; μ is a probability measure on R ; $c(x, y)$ is a symmetric continuous function defined on $R^q \times R^q$; and $Z_{n,\xi}$ is the normalizing constant

$$Z_{n,\xi} = \int_{R^n} \exp\left\{\frac{\beta}{2n} \sum_{i,j=1}^n c(\xi_i, \xi_j) x_i x_j\right\} \prod_{i=1}^n \mu(dx_i). \quad (0.14)$$

We show that the random probability measures $(Q_{n,\xi})_{n=1}^\infty$ defined by (0.9) form a large deviation system for almost all realizations of ξ with a non-random rate function

$$I(m) = \sum_{k=1}^p \lambda_k L^*\left(\frac{m_k}{\lambda_k}\right), \quad m = (m_1, \dots, m_p) \in R^p \quad (0.15)$$

where $L^*(x)$ is the Legendre transformation of $L(t)$

$$L^*(x) = \sup_{t \in R} \{tx - L(t)\}, \quad x \in R \quad (0.16)$$

and $L(t)$ is the logarithmic moment generating function of μ

$$L(t) = \log \int_R \exp(tx) \mu(dx), \quad x \in R \quad (0.17)$$

It follows from this result and a Varadhan type theorem that the free energy function $f(\beta)$ exists almost surely with respect to the random variables ξ and is given by a non-random variational formula. More specifically, for almost all ξ .

$$\begin{aligned} f(\beta) &= -\frac{1}{\beta} \sup_{m \in R^p} \left\{ \frac{1}{2} \beta \sum_{k,l=1}^p c(\theta_k, \theta_l) m_k m_l - \sum_{k=1}^p \lambda_k L^*\left(\frac{m_k}{\lambda_k}\right) \right\} \\ &= \inf_{m \in R^p} \left\{ \frac{1}{\beta} \sum_{k=1}^p \lambda_k L^*\left(\frac{m_k}{\lambda_k}\right) - \frac{1}{2} \sum_{k,l=1}^p c(\theta_k, \theta_l) m_k m_l \right\} \\ &= \inf_{m \in R^p} \left\{ \frac{1}{\beta} \sum_{k=1}^p \lambda_k L^*(m_k) - \frac{1}{2} \sum_{k,l=1}^p \lambda_k \lambda_l c(\theta_k, \theta_l) m_k m_l \right\}. \end{aligned} \quad (0.18)$$

For fixed β and $\lambda_1, \dots, \lambda_p$, denote

$$V(m_1, \dots, m_p) = \frac{1}{\beta} \sum_{k=1}^p \lambda_k L^*(m_k) - \frac{1}{2} \sum_{k,l=1}^p \lambda_k \lambda_l c(\theta_k, \theta_l) m_k m_l. \quad (0.19)$$

Then the infimum in (0.18) is realized among the m such that the gradient ∇V of $V(m_1, \dots, m_p)$ vanishes at m , i.e.,

$$\frac{\partial}{\partial m_k} V(m_1, \dots, m_p) = \frac{1}{\beta} \frac{d}{dm_k} L^*(m_k) - \sum_{l=1}^p \lambda_l c(\theta_k, \theta_l) m_l = 0, \quad k = 1, \dots, p. \quad (0.20)$$

where in the derivation of (0.20) we used the assumption that $c(x, y)$ is symmetric. According to the convex analysis [40], the derivative of the Legendre transformation of a convex function is the inverse of the derivative of the original function

$$\frac{d}{dt} L^*(t) = \left(\frac{d}{dt} L(t) \right)^{-1}. \quad (0.21)$$

Thus the infimum in (0.18) is realized among the m that satisfy the fixed-point equations

$$m_k = \frac{d}{dt} L(t) \Big|_{t=\beta \sum_{l=1}^p \lambda_l c(\theta_k, \theta_l) m_l}, \quad k = 1, \dots, p. \quad (0.22)$$

For Ising spin model, i.e., $x_i \in \{+1, -1\}$ with $\mu = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}$ the equations (0.22) can easily be reduced to

$$m_k = \tanh\left\{\beta \sum_{l=1}^p \lambda_l c(\theta_k, \theta_l) m_l\right\}, \quad k = 1, \dots, p \quad (0.23)$$

For $c(x, y)$ as in (0.7) the fixed point equations (0.23) have been studied in details by Amit et al in [1]. For general symmetric $c(x, y)$ the fixed point equations (0.23) have also been studied by van Hemmen et al [52]. As was shown there, for high temperature T (i.e., small β because $\beta = \frac{1}{T}$), the only solutions to (0.23) is $m_k = 0$ for all $k = 1, \dots, p$. However, bifurcation occurs at a critical temperature

$$T_c = 2^{-1} \lambda_{\max} \quad (0.24)$$

where $\lambda_{\max}$ is the largest positive eigenvalue of the $p \times p$ matrix $\{c(\theta_k, \theta_l)\}$. Under suitable conditions the first solutions to branch off as T decreases from T_c correspond to the p stored patterns. As temperature continuously decreases, further bifurcation is possible. The bifurcation temperature depends on the second largest eigenvalue of $\{c(\theta_k, \theta_l)\}$ and the solutions correspond to the metastable states.

Although the free energy per spin $f(\beta)$ does not depend on particular realizations of the random interactions in the thermodynamic limit, the random Gibbs measures $P_{n,\xi}(dx_1, \dots, dx_n)$ defined by (0.13) do depend on the realizations of the random interactions even in the thermodynamic limit. More precisely, we show that

1. For almost all ξ , $(\tilde{P}_{n,\xi})_{n=1}^{\infty}$ is relatively compact as a sequence of probability measures on $R^{\otimes \infty}$, where $\tilde{P}_{n,\xi}(dx_1, dx_2, \dots) = P_{n,\xi}(dx_1, \dots, dx_n) \prod_{i=n+1}^{\infty} \nu(dx_i)$ for some $\nu \in \mathbf{M}_1(R)$.
2. Any limit point P_ξ of $(\tilde{P}_{n,\xi})_{n=1}^{\infty}$ can be expressed as

$$P_\xi(dx_1, dx_2, \dots) = \int_{R^p} Q(dm) \prod_{k=1}^p \prod_{\substack{i=1 \\ \xi_i = \theta_k}}^{\infty} \frac{\exp\{x_i [Cm]_k\} \mu(dx_i)}{\int_R \exp\{x_i [Cm]_k\} \mu(dx_i)} \quad (0.25)$$

where $C = \{c(\theta_k, \theta_l)\}$ is the $p \times p$ matrix, the notation $[Cm]_k$ indicates the k th component of the vector Cm , and Q is a probability measure on R^p such that it concentrates on the subset of R^p where the function

$$\frac{1}{2\beta} \sum_{k,l=1}^p c(\theta_k, \theta_l) m_k m_l - \sum_{k=1}^p \lambda_k L^* \left(\sum_{l=1}^p c(\theta_k, \theta_l) m_l \right) \quad (0.26)$$

reaches its minimum.

The next step in dealing with any physical model is to study the dynamics. A natural stochastic dynamics with prescribed equilibrium distribution (0.2) is defined by the Itô stochastic differential equations in R

$$dx_i(t) = -[U'(x_i(t)) + \frac{\partial}{\partial x_i(t)} H_n^{(\theta, \beta)}(x_1(t), \dots, x_n(t))]dt + \sigma dw_i(t), \quad (0.27)$$

$$i = 1, \dots, n,$$

where $w_1, \dots, w_n$ are n independent 1-dimensional Wiener processes; $U(x)$ is a suitable potential function on R ; and σ is a positive constant. In order to expose the methods in obtaining the limit theorems for these interacting processes, let us first consider the continuous Curie-Weiss model where the Hamiltonian is not random. For example, Dawson in [5] considered one such system with a double well potential

$$dx_i(t) = -[U'(x_i(t)) + \frac{\theta}{n} \sum_{j=1}^n (x_i(t) - x_j(t))]dt + \sigma dw_i(t), \quad (0.28)$$

$$i = 1, \dots, n,$$

where θ is a positive constant and $U(x) = \frac{x^4}{4} - \frac{x^2}{2}$. In order to study the asymptotic behavior of the system (0.28) in the limit as n goes to infinity, it is natural to consider the empirical process X_n associated to (0.28), which is defined by

$$X_n(t) = \frac{1}{n} \sum_{i=1}^n \delta_{x_i(t)}, \quad t \geq 0, \quad (0.29)$$

where δ_x is the Dirac measure at x . It was proved in [5] that X_n is a measure-valued diffusion process with sample paths in $C([0, \infty), \mathbf{M}_1(R))$, where $\mathbf{M}_1(R)$ is the space

of probability measures on Borel subsets of R . As $n \rightarrow \infty$, the sequence of measure-valued process $(X_n(\cdot))_{n=1}^{\infty}$ converges in law in $C([0, \infty), \mathbf{M}_1(R))$ to a deterministic measure-valued function $X_{\infty}(\cdot)$, known as the McKean-Vlasov limit [31], which is the probability distribution of the unique solution of the following non-linear Markov diffusion process

$$dx(t) = -[U'(x(t)) + \theta(x(t) - m(t))]dt + \sigma dw(t), \quad (0.30)$$

where

$$m(t) = \int_R x X_{\infty}(t, dx) \quad (0.31)$$

and

$$X_{\infty} = \text{the distribution of } x(t) \quad (0.32)$$

One can then study the phase transition and fluctuations for the mean field model by considering the non-linear Markov diffusion process. It was proved in [5] that the associated McKean-Vlasov dynamics undergoes a bifurcation at a critical temperature σ_c . For $\sigma \geq \sigma_c$, the McKean-Vlasov dynamics has exactly one equilibrium μ_0 which is globally stable and corresponds to mean magnetization zero. For $\sigma < \sigma_c$, the stability of μ_0 will be lost and two new stable equilibria μ_+ and μ_- will appear which correspond to positive and negative mean magnetization, respectively. It was also established that the fluctuations at the critical point were non-Gaussian and occur at a time scale slower than the non-critical fluctuations.

Such systems, known as weakly interacting diffusion processes, were later studied by several mathematicians in more general setting (see, for examples, Dawson and Gärtner [6],[7], Gärtner [21], Léonard [29],[30], Oelschläger [34] and Sznitman [46]).

Inspired by the work on weakly interacting diffusions we consider the following system of Itô stochastic differential equations in R

$$dx_i^{(n)}(t) = b(x_i^{(n)}(t))dt + \frac{1}{n} \sum_{j=1}^n c(\xi_i, \xi_j) \Phi(x_i^{(n)}(t), x_j^{(n)}(t))dt + \sigma(x_i^{(n)}(t))dw_i(t), \quad (0.33)$$

$$i = 1, \dots, n,$$

where $w_1, \dots, w_n$ are n independent 1-dimensional Wiener processes; $b(x)$, $\sigma(x)$ and $\Phi(x, y)$ are suitable functions defined on R , R and $R \times R$, respectively; $\xi_1, \dots, \xi_n$ are n i.i.d. random variables defined on some probability space and take values in $\{\theta_1, \dots, \theta_p\} \subset R^q$. Under suitable conditions on the coefficient, for each $\xi_1, \dots, \xi_n$, the system of equations (0.33) determines uniquely a diffusion process $(x_{\xi,1}^{(n)}(t), \dots, x_{\xi,n}^{(n)}(t))$, $t \geq 0$, where $\xi = (\xi_1, \xi_2, \dots)$. The empirical measure $X_\xi^{(n)}$ associated to the n diffusions is defined by

$$X_\xi^{(n)} = \frac{1}{n} \sum_{i=1}^n \delta_{x_{\xi,i}^{(n)}} \quad (0.34)$$

which is viewed as a random probability measure on the path space $C([0, \infty), R)$.

Unlike the usual weakly interacting diffusions where the interactions can be expressed as some integral with respect to the empirical measure, the interactions in (0.33) depend on the p associated local empirical measures $X_{\xi,1}^{(n)}, \dots, X_{\xi,p}^{(n)}$ which can be defined by

$$X_{\xi,k}^{(n)} = \frac{1}{n_{\xi,\theta_k}} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \delta_{x_{\xi,i}^{(n)}}, \quad k = 1, \dots, p \quad (0.35)$$

where n_{ξ,θ_k} counts the total numbers of $i \in \{1, \dots, n\}$ such that ξ_i takes the value θ_k

$$n_{\xi,\theta_k} = \sum_{i=1}^n 1_{\theta_k}(\xi_i), \quad \text{where} \quad 1_x(y) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{if } y \neq x \end{cases} \quad (0.36)$$

In chapter 4 we first show that for almost all ξ , the laws $\mathcal{L}(X_{\xi,1}^{(n)}, \dots, X_{\xi,p}^{(n)})$ of the random probability measures $(X_{\xi,1}^{(n)}, \dots, X_{\xi,p}^{(n)})$ converge in probability, as $n \rightarrow \infty$, to a deterministic measure $(m_{\theta_1}, \dots, m_{\theta_p})$ in $\mathbf{M}_1(C([0, \infty), R))^{\otimes p}$, the p -fold product

space of $M_1(C([0, \infty), R))$. We then use some theorem on the propagation of chaos to show that for almost all ξ , the laws $(\mathcal{L}(X_\xi^{(n)})_{n=1}^\infty$ of $(X_\xi^{(n)})_{n=1}^\infty$ converge weakly as $n \rightarrow \infty$ in $M_1(M_1(C([0, \infty), R)))$ to the Dirac measure $\delta_{\sum_{k=1}^p \lambda_k m_{\theta_k}}$, written as

$$\mathcal{L}(X_\xi^{(n)}) \Rightarrow \delta_{\sum_{k=1}^p \lambda_k m_{\theta_k}} \text{ for all most all } \xi \quad (0.37)$$

where m_{θ_k} , $1 \leq k \leq p$, is the law on $C([0, \infty), R)$ of the unique solution x_ξ of the following stochastic differential equation under the condition that $\xi = \theta_k$:

$$dx_\xi(t) = b(x_\xi(t))dt + \sum_{l=1}^p \lambda_l c(\xi, \theta_l) \int_R \Phi(x_\xi(t), y) m_{\theta_l}(t, dy) dt + \sigma(x_\xi(t))dw(t), \quad (0.38)$$

$$m_{\theta_l}(t, dy) = P\{x_{\theta_l}(t) \in dy\}, \quad l = 1, \dots, p$$

As an example, this result is then applied to analyze a non-linear neural network model.

One of the most interesting yet difficult problem in spin glass models is the so called dynamical mean field theory proposed by Sompolinsky and Zippelius [43] in 1982 as an approach to study the relaxation dynamics of spin glasses. From mathematical point of view, this approach suggests one to consider the following n randomly coupled stochastic differential equations

$$dx_i^n(t) = b(x_i^n(t))dt + \frac{1}{\sqrt{n}} \sum_{j=1}^n J_{i,j} \varphi(x_j^n(t))dt + \sigma(x_i^n(t))dw_i(t), \quad (0.39)$$

$$i = 1, \dots, n$$

where $b(x)$, $\varphi(x)$ and $\sigma(x)$ are measurable functions defined on R ; $w_1, \dots, w_n$ are n independent R -valued Wiener processes; $\{J_{i,j}\}$ is a set of i.i.d. R -valued random variables with zero mean and variance 1. According to the dynamical mean field theory, in the thermodynamic limit of $n \rightarrow \infty$, the behavior of the system (0.39) should be described by one single equation of the following form:

$$dx(t) = b(x(t))dt + G(t)dt + \sigma(x(t))dw(t) \quad (0.40)$$

where $w(t)$ is the R -valued Wiener process; $G(t)$ is a Gaussian process with mean zero and variance $E[G(t)G(s)] = E[\varphi(x(t))\varphi(x(s))]$. Although these are very interesting non-linear stochastic differential equations, it is not clear how to construct such processes and to prove the McKean-Vlasov type limit theorem. In chapter 6 we consider a mathematically more tractable system described by the following coupled stochastic differential equations:

$$dx_i^n(t) = -x_i^n(t)dt + \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi(x_j^n(t))dw_{ij}(t) \quad (0.41)$$

$$i = 1, \dots, n$$

where $\{w_{ij}\}$ are n^2 independent 1-dimensional Brownian motions; φ is a continuous function on R . We show that, as $n \rightarrow \infty$, the empirical measure associated with the finite system (0.41) converges in probability to a deterministic measure which is the distribution of the process $\{x(t) : t \geq 0\}$ determined uniquely by the stochastic differential equation:

$$dx(t) = -x(t)dt + \sqrt{\int_R \varphi^2(x)\mu(t)(dx)}dw(t) \quad (0.42)$$

$$\mu(t) = \text{the distribution of } x(t)$$

where w is the 1-dimensional Brownian motion. For the particular function

$$\varphi(x) = \sqrt{2} \tanh(\beta x) \quad (0.43)$$

where β is a positive parameter, we show that for sufficiently large β the limiting process (0.42) has a unique invariant Gaussian measure

$$P_\sigma(dx) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{x^2}{2\sigma}\right\}dx \quad (0.44)$$

where σ is the unique positive solution of the Sherrington-Kirkpatrick spin glass fixed point equation

$$\sigma = \int_{-\infty}^{\infty} \tanh^2(\beta\sigma^{\frac{1}{2}}x) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)dx. \quad (0.45)$$

We now give the outline of the structure of this thesis. The thesis consists of six chapters. In chapter 1 we present some mathematical background materials including the basic notions on Markov diffusion processes, the weak convergence of probability measures on metric spaces, and the classical large deviation principle with its applications to statistical physics. In chapter 2 we give a short introduction to the spin glass problem and some related mathematical work. In chapter 3 we introduce the stochastic models of finite particle systems described by Itô stochastic differential equations. We show the existence and uniqueness for the solutions of the finite systems. The existence and uniqueness of the non-linear Markov diffusion processes associated to these models are also established. In chapter 4 we prove the McKean-Vlasov type limit theorem for the random empirical measure-valued process. The main idea is to introduce the local empirical measures associated with the finite systems, which can be viewed as order parameters from the physical point of view. In chapter 5 we study the asymptotic behavior of random Gibbs measures associated to finite systems using large deviation techniques. The main new concept there involves the notion of large deviations in the almost sure sense. Finally in chapter 6 we consider the system in which the random interactions are assumed to be a set of independent Brownian motions. We show that for sufficiently large parameter β the corresponding McKean-Vlasov limit has a unique invariant Gaussian probability measure such that its variance is determined by the Sherrington-Kirkpatrick spin glass equation.

Chapter 1

Preliminaries

In this chapter we collect some basic notions and results on stochastic processes, which will be used in the following chapters. In section 1 we present the basic theory of Markov diffusion processes and Itô stochastic differential equations. In section 2 we focus on the weak convergence of probability measures on metric spaces. In section 3 we introduce the classical large deviation theory and its applications to statistical mechanics.

We first introduce some notations used in this thesis.

R^d	d -dimensional Euclidean space; $R = R^1$.
$R^d \otimes R^d$	Space of $d \times d$ real matrices.
(S, d)	Complete and separable metric space.
$\mathcal{B}(S)$	Borel σ -algebra of subsets of S .
$B(S)$	Space of real valued bounded measurable functions on S .
$\ \cdot \ $	Supremum norm $\ f \ = \sup_{x \in S} f(x) $ for $f \in B(S)$.
$M(S)$	Space of finite signed measures on $(S, \mathcal{B}(S))$.
$\ \cdot \ ^\ast$	Total variational norm $\ \mu \ ^\ast = \mu (S)$ for $\mu \in M(S)$.
$C_b(S)$	Set of bounded continuous functions on S .
$C(R^d)$	Set of continuous functions on R^d .
$C^2(R^d)$	Set of two times continuously differentiable functions on R^d .

- $C_K(R^d)$ Set of real valued functions $f \in C_b(R^d)$ with compact support.
 $C_K^2(R^d)$ Set of functions $f \in C_K(R^d) \cap C^2(R)$.
 $C_K^\infty(R^d)$ Set of functions $f \in C_K(R^d)$ differentiable at any order.
 $\mathbf{M}_1(S)$ set of probability measures on $(S, \mathcal{B}(S))$.
 $\ll \gg$ quadratic variational process for martingales.
 $\langle \mu, f \rangle$ integral $\int f d\mu$ provided it makes sense.
 (x, y) the inner product in R^d .
 $c_1, c_2, \dots; k_1, k_2, \dots$ constants.

1.1 Markov diffusion processes

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $(\mathcal{F}_t)_{t \geq 0}$ be a family of sub σ -algebras of $\mathcal{F}$ which is assumed to satisfy the standard conditions:

1. it is increasing, i.e., $\mathcal{F}_s \subset \mathcal{F}_t$ if $0 \leq s \leq t$;
2. it is right-continuous, i.e., $\mathcal{F}_t = \bigcap_{\epsilon > 0} \mathcal{F}_{t+\epsilon}$ for every $t \geq 0$; and
3. it is complete, i.e., $\mathcal{F}_0$ contains all the subsets A of Ω such that there exist $B \in \mathcal{F}_0$ such that $A \subset B$ and $P(B) = 0$.

Let $X = \{X(t), t \geq 0\}$ be a stochastic process defined on $(\Omega, \mathcal{F})$ with values in S which is $\mathcal{F}_t$ -adapted, i.e, for each $t \geq 0$, $X(t)$ is measurable with respect to $\mathcal{F}_t$. Let $(P_x)_{x \in S}$ be a collection of probability measures on $(\Omega, \mathcal{F})$.

Definition 1.1 *The collection of the objects $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, X, (P_x)_{x \in S})$ given above is said to be a time homogeneous Markov process with state space $(S, \mathcal{B}(S))$ if the following conditions are satisfied:*

1. for any $t \geq 0$ and $\Gamma \in \mathcal{B}(S)$ the function

$$P_x\{X(t) \in \Gamma\} = p(t, x, \Gamma) \quad (1.1)$$

is measurable in the variable x with respect to the σ -algebra $\mathcal{B}(S)$.

$$2. p(0, x, S \setminus \{x\}) = 0.$$

3. for $0 \leq s \leq t$, $x \in S$ and $\Gamma \in \mathcal{B}(S)$ the equality

$$P_x\{X(t) \in \Gamma | \mathcal{F}_s\} = p(t-s, X(s), \Gamma) \quad (1.2)$$

holds almost surely with respect to the measure P_x .

The function $p(t, x, \Gamma)$ is called the transition function of the Markov process. To each transition function of a Markov process, there are associated two families of linear operators: $\{T_t, t \geq 0\}$ and $\{U_t, t \geq 0\}$. The first one acts on $(B(S), \|\cdot\|)$ according to the formula

$$(T_t f)(x) = \int_S f(y) p(t, x, dy), \quad x \in S. \quad (1.3)$$

and the second one acts on $(M(S), \|\cdot\|^*)$ according to the formula

$$(U_t \mu)(\Gamma) = \int_S p(t, x, \Gamma) \mu(dx), \quad \Gamma \in \mathcal{B}(S). \quad (1.4)$$

The two operators T_t and U_t are connected by the relationship

$$\int_S (T_t f)(x) \mu(dx) = \int_S f(x) (U_t \mu)(dx), \quad f \in B(S), \mu \in M(S). \quad (1.5)$$

It follows easily from (1.3) that the family of linear operators $\{T_t, t \geq 0\}$ is bounded which preserves non-negativity and forms a contractive semigroup, i.e., for $s, t \geq 0$ and $f \in B(S)$,

$$1. T_t f(x) \geq 0 \quad \text{if} \quad f \geq 0$$

$$2. T_0 f(x) = f(x)$$

$$3. \|T_t f\| \leq \|f\|$$

$$4. T_{t+s}f(x) = T_t(T_sf(x))$$

The infinitesimal generator A of the semigroup T_t is defined by

$$Af = \lim_{t \downarrow 0} \frac{T_t f - f}{t} \quad (1.6)$$

where the limit is understood in the sense of convergence in the norm

$$\| Af - \frac{1}{t}(T_t f - f) \| \rightarrow 0 \quad \text{as } t \downarrow 0 \quad (1.7)$$

and the domain $D(A)$ of A consists of all $f \in B(S)$ such that the limit (1.6) exists in $B(S)$. It is known that the domain $D(A)$ of A is everywhere dense in the space

$$B_0(S) = \{f \in B(S) : \lim_{t \downarrow 0} \| T_t f - f \| = 0\}. \quad (1.8)$$

If the transition function $p(t, x, \Gamma)$ is stochastically continuous in the sense that

$$\lim_{t \downarrow 0} p(t, x, S \setminus N) = 0 \quad (1.9)$$

for every $x \in S$ and every neighbourhood N of x , then the infinitesimal generator A of the semigroup T_t determines uniquely all finite-dimensional distributions of the Markov process [14].

Based on the ideas from the theory of semigroups, it can be proved that function $u(t, x) = T_t f(x)$ is a solution of the abstract Cauchy problem

$$\frac{\partial u(t, x)}{\partial t} = Au(t, x), \quad (1.10)$$

$$\lim_{t \downarrow 0} u(t, x) = f(x).$$

The solution of this problem is always unique in the class of functions $u(t, x)$ such that $\| u(t, \cdot) \| \leq K \exp(\alpha t)$, where K and α are arbitrary constants [14].

Similarly, the family of linear operators $\{T_t, t \geq 0\}$ defined by (1.4) also forms a contractive semigroup and one can define the infinitesimal generator A^* of the

semigroup U_t analogously. The semigroup U_t determines the evolution of the one-dimensional distribution of the Markov process, i.e., suppose that the Markov process starts at the random point $X(0)$ which has distribution μ

$$P\{X(0) \in \Gamma\} = \mu(\Gamma), \quad \Gamma \in \mathcal{B}(S) \quad (1.11)$$

then for any $t > 0$

$$P\{X(t) \in \Gamma\} = \int_S P\{X(0) \in dx\} P(t, x, \Gamma) = (U_t \mu)(\Gamma). \quad (1.12)$$

A measure μ on $(S, \mathcal{B}(S))$ is called an invariant measure of the Markov process if

$$U_t \mu = \mu \quad \text{for all } t \geq 0. \quad (1.13)$$

It can be verified that every invariant measure belongs to the domain of A^* and

$$A^* \mu = 0. \quad (1.14)$$

For further information about Markov processes and their generators see the books by Dynkin [14] or the book by Ethier and Kurtz [18].

We now consider a special class of Markov processes, the diffusion processes. These are the processes with state space R^d and related to a second order differential operator of the form

$$L = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(x) \frac{\partial}{\partial x_i} \quad (1.15)$$

where for each $x \in R^d$, $a(x) = (a_{ij}(x))_{i,j=1}^d$ is a symmetric non-negative definite $d \times d$ matrix, i.e.,

$$a_{ij}(x) = a_{ji}(x)$$

$$\sum_{i,j=1}^d \lambda_i a_{ij}(x) \lambda_j \geq 0 \quad \text{for all } (\lambda_1, \dots, \lambda_d) \in R^d$$

and $b(x) = (b_i(x))_{i=1}^d$ is a $d \times 1$ vector. We assume that the components $a_{ij}(x)$ and $b_i(x)$ are Borel measurable and bounded on each compact subset of R^d .

continuous $(\mathcal{F}_t, P)$ local martingale with corresponding quadratic variational processes $\langle\langle M \rangle\rangle_t$. In the following we always write $X(0) + V(t)$ as $A(t)$ so that $X(t) = A(t) + M(t)$. We have the following theorem which is mainly due to Aldous and Rebolledo [27].

Theorem 1.6 *Let $\{P_n\}$ be a sequence of probability measures on $C([0, \infty), R)$ which are the laws of the continuous semimartingales $X_n(t) = A_n(t) + M_n(t)$. Suppose the following conditions hold:*

1. *for each $t \geq 0$, the laws $\mathcal{L}(X_n(t))$ of $X_n(t)$ are tight in $M_1(R)$;*
2. *for each positive η, ϵ and T , there exists $\delta > 0$ such that for any stopping time τ_n bounded by T such that*

$$\limsup_{n \rightarrow \infty} \sup_{\theta \leq \delta} P(|A_n(\tau_n + \theta) - A_n(\tau_n)| > \epsilon) \leq \eta$$

and

$$\limsup_{n \rightarrow \infty} \sup_{\theta \leq \delta} P(|\langle M_n \rangle_{\tau_n + \theta} - \langle M_n \rangle_{\tau_n}| > \epsilon) \leq \eta$$

Then $\{P_n\}$ is tight in $M_1(C([0, \infty), R))$.

Proof: Refer to Joffe and Metivier [27].

The book by Ethier and Kurtz [18] contains a very nice and complete presentation on the theory of weak convergence of probability measures on metric spaces.

1.3 Large deviations and statistical mechanics

Throughout this section let the following objects be given:

(P_n) a sequence of probability measures on $\mathcal{B}(S)$.

(a_n) a sequences of positive numbers tending to infinity.

I a function that maps S to $[0, \infty]$.

2. $P_x\{\omega: X(0) = x\} = 1$ for every $x \in R^d$.
3. $\{P_x, x \in R^d\}$ has the Markov property, i.e., for every $t > s \geq 0$, $A \in \mathcal{B}_t$ and a Borel subset B of R^d , we have

$$P_x(A \cap \{X(t+s) \in B\}) = \int_A P_{y(s)}(X(t) \in B) P_x(dy) \quad (1.20)$$

for every $x \in R^d$.

4. for all $f \in C_K^2(R^d)$, the process M_f defined by

$$M_f(t) = f(X(t)) - f(X(0)) - \int_0^t Lf(X(s))ds \quad (1.21)$$

is a $((\mathcal{B}_t)_{t \geq 0}, P_x)$ -martingale for all $x \in R^d$.

There are several ways to construct a Markov diffusion process or a system of diffusion measures for a given differential operator L . The classical probability approach is to use Itô stochastic differential equations in which the Wiener process serves as a building block.

Let W be a d -dimensional Wiener process defined on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ such that $W(0) = 0$, P -a.s., and $W(t)$ is $\mathcal{F}_t$ -measurable for every $t \geq 0$. Let ξ be a R^d -valued random variable which is $\mathcal{F}_0$ -measurable and independent of the Wiener process.

Consider the following Itô stochastic differential equation on R^d :

$$dX(t) = b(X(t))dt + \sigma(X(t))dW(t) \quad (1.22)$$

or equivalently, in terms of its components,

$$dX_i(t) = b_i(X(t))dt + \sum_{j=1}^d \sigma_{ij}(X(t))dW_j(t), \quad i = 1, \dots, d \quad (1.23)$$

where $b: R^d \rightarrow R^d$ and $\sigma: R^d \rightarrow R^d \otimes R^d$ are Borel measurable functions.

Definition 1.4 We say that X is a solution of the equation (1.22) if

1. X is a $\mathcal{B}([0, \infty)) \times \mathcal{F}$ -measurable and $\mathcal{F}_t$ -adapted process defined on $(\Omega, \mathcal{F}, P)$ with sample paths in $C([0, \infty), R^d)$.
2. the following equality holds P -a.s.

$$X(t) = \xi + \int_0^t b(X(s))ds + \int_0^t \sigma(X(s))dW(s) \quad (1.24)$$

We say that the uniqueness holds for the solutions of (1.22) if, whenever X and Y are two solutions of (1.22), then

$$P\{X(t) = Y(t) \text{ for all } t \geq 0\} = 1. \quad (1.25)$$

The following theorem is one of the standard theorems on the existence and uniqueness of the stochastic differential equations (1.22).

Theorem 1.1 Suppose that $b : R^d \rightarrow R^d$ and $\sigma : R^d \rightarrow R^d \odot R^d$ satisfy the following conditions:

1. for every $R > 0$, there exists a constant K_R such that

$$|b(x) - b(y)|^2 + |\sigma(x) - \sigma(y)|^2 \leq K_R(|x - y|^2) \quad (1.26)$$

for every $|x| \leq R, |y| \leq R$.

2. there exists a constant K such that

$$x \cdot b(x) + |\sigma(x)|^2 \leq K(1 + |x|^2) \quad (1.27)$$

for every $x \in R^d$.

Then for any $\xi \in L^2(\Omega, \mathcal{F}_0, P)$, the equation (1.22) has a unique solution such that for any $T > 0$,

$$E\left\{\sup_{0 \leq t \leq T} |X(t)|^2\right\} \leq C(1 + E(|\xi|^2)) \quad (1.28)$$

where C is a constant depending only on K_R , K and T . More generally, if $E(|\xi|^{2m}) < \infty$ for some positive integer m , then

$$E\left\{\sup_{0 \leq t \leq T} |X(t)|^{2m}\right\} \leq C(1 + E(|\xi|^{2m})) \quad (1.29)$$

Proof: Refer to Ethier and Kurtz [18] or Ikeda and Watanabe [26].

Suppose the symmetric non-negative definite matrix $a(x) = (a_{ij}(x))_{i,j}^d$ can be written as

$$a(x) = \sigma(x)\sigma^*(x) \quad (1.30)$$

and the conditions in the theorem 1.1 are satisfied. Denote by $\{X_x(t), t \geq 0\}$ the unique solution of the corresponding stochastic differential equation with initial condition $X(0) = x$. Let P_x be the probability distribution of $\{X_x(t), t \geq 0\}$ on the path space. For $t > 0$, $x \in R^d$ and $\Gamma \in \mathcal{B}(R^d)$, define

$$P(t, x, \Gamma) = P_x\{X(t) \in \Gamma\} \quad (1.31)$$

Then it can be proved that the family $\{P_x : x \in R^d\}$ of probability measures on $(C([0, \infty), R^d), \mathcal{B})$ satisfies all the conditions in definition 1.3 and hence defines the system of diffusion measures generated by the operator $(L, C_K^2(R^d))$. The corresponding Markov diffusion process has transition function $P(t, x, \Gamma)$. The domain $D(A)$ of its infinitesimal generator A contains $C_K^2(R^d)$ and for $f \in C_K^2(R^d)$

$$Af(x) = Lf(x). \quad (1.32)$$

The proof of this and many other results on diffusion processes defined by Itô stochastic differential equations can be found in the book by Ethier and Kurtz [18] or in the book by Ikeda and Watanabe [26].

1.2 Weak convergence of probability measures

Let $\mathbf{M}_1(S)$ be the set of probability measures on the complete and separable metric space (S, d) . We view $\mathbf{M}_1(S)$ as a subset of the dual space of $C_b(S)$ and furnish it the inherited weak-star topology, i.e., the weakest topology such that the mapping

$$P \rightarrow \int_S f(x)P(dx), \quad P \in \mathbf{M}_1(S)$$

is continuous for each $f \in C_b(S)$.

Definition 1.5 *A sequence $\{P_n\}$ of probability measures on $(S, \mathcal{B}(S))$ is said to be weakly convergent to a probability measure P on $(S, \mathcal{B}(S))$ if for every $f \in C_b(S)$*

$$\lim_{n \rightarrow \infty} \int_S f(x)P_n(dx) = \int_S f(x)P(dx), \quad (1.33)$$

and we write $P_n \Rightarrow P$.

A sequence $\{X_n\}$ of S -valued random variables is said to converge in distribution to the S -valued random variable X if the distributions of X_n , denoted by $\mathcal{L}(X_n)$, converges weakly to $\mathcal{L}(X)$, or equivalently, if

$$\lim_{n \rightarrow \infty} Ef(X_n) = Ef(X), \quad (1.34)$$

for all $f \in C_b(S)$.

It is well known that the topology of weak convergence on $\mathbf{M}_1(S)$ is metrizable. One such metric is the Prohorov metric $\rho(P, Q)$ defined by

$$\rho(P, Q) = \inf\{\epsilon > 0 : P(F) \leq Q(F^\epsilon) + \epsilon \text{ for all } F \in \mathcal{C}\} \quad (1.35)$$

where $\mathcal{C}$ is the collection of closed subsets of S and

$$F^\epsilon = \inf\{x \in S : \inf_{y \in F} d(x, y) < \epsilon\}. \quad (1.36)$$

The other metric is $\|P - Q\|_{BL}^*$, the dual norm on the dual space $(BL(S, d))^*$, defined by

$$\|P - Q\|_{BL}^* = \sup_{f \in BL(S, d), \|f\|_{BL} \leq 1} \{|\langle P, f \rangle - \langle Q, f \rangle|\} \quad (1.37)$$

where $BL(S, d)$ is the Banach space of the real valued functions $f \in C_b(S)$ such that f is Lipschitz with the corresponding norm

$$\|f\|_{BL} = \sup_{x \in S} |f(x)| + \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}. \quad (1.38)$$

Furthermore, $(M_1(S), \rho)$ and $(M_1(S), \|\cdot\|_{BL}^*)$ are both complete and separable metric spaces [18] [13].

In order to characterize the compact sets in $M_1(S)$, it is necessary to introduce the notion of tightness.

Definition 1.6 A sequence $\{P_n\}$ of probability measures on $(S, \mathcal{B}(S))$ is said to be tight, if for every $\varepsilon > 0$ there exists a compact subset $K_\varepsilon \subset S$ such that

$$\inf_n P_n(K_\varepsilon) \geq 1 - \varepsilon \quad (1.39)$$

Now the compact sets in $M_1(S)$ can be characterized as follows.

Theorem 1.2 (Prohorov) Let S be a complete and separable metric space and let $\{P_n\} \subset M_1(S)$. Then $\{P_n\}$ is relatively compact in $M_1(S)$ if and only if $\{P_n\}$ is tight.

Proof: Refer to Ethier and Kurtz [18].

The way of proving the weak convergence of a sequence $\{P_n\} \subset M_1(S)$ of probability measures towards a probability measure $P \in M_1(S)$ usually consists of the following two steps:

1. Prove $\{P_n\}$ is relatively compact in $M_1(S)$ by first showing it is tight and then by invoking the Prohorov theorem.
2. Show that P is the unique limit of every convergent subsequence.

It is, therefore, technically important to be able to show the tightness of a given sequence of probability measures $\{P_n\}$ in $M_1(S)$. One of the general results for the tightness of probability measures in $M_1(C([0, \infty), S))$ can be stated as follows.

Theorem 1.3 *Let (S, d) be a complete and separable metric space and $X(t)$ be the usual projective map on $C([0, \infty), S)$. A sequence $\{P_n\}$ of probability measures in $M_1(C([0, \infty), S))$ is tight if and only if the following two conditions are satisfied*

1. *For every $\epsilon > 0$ and $t \geq 0$, there exists a compact set $K_\epsilon \subset S$ such that*

$$\liminf_{n \rightarrow \infty} P_n\{X(t) \in K_\epsilon\} \geq 1 - \epsilon \quad (1.40)$$

2. *For every $\epsilon > 0$ and $T \geq 0$, there exists $\delta > 0$ such that*

$$\limsup_{n \rightarrow \infty} P_n\{w(X, \delta, T) \geq \epsilon\} \leq \epsilon \quad (1.41)$$

where for $x \in C([0, \infty), S)$, $\delta > 0$ and $T > 0$, $w(x, \delta, T)$ is defined by

$$w(x, \delta, T) = \inf_{\{t_i\}} \max_{i=1, \dots, n} \sup_{s, t \in [t_{i-1}, t_i]} d(x(s), x(t)) \quad (1.42)$$

where $\{t_i\}$ ranges over all partitions of the form $0 = t_0 < t_1 < \dots < t_{n-1} < T < t_n$ with $\min_{1 \leq i \leq n} (t_i - t_{i-1}) > \delta$ and $n \geq 1$.

Proof: Refer to Ethier and Kurtz [18].

We now present several useful criteria for checking tightness of a given sequence of probability measures $\{P_n\}$ in $M_1(S)$. The first one is on the tightness on the countably many product metric spaces.

Theorem 1.4 *Let (S_k, d_k) , $k = 1, 2, \dots$, be metric spaces, and define the metric space (S, d) by letting*

$$S = \prod_{k=1}^{\infty} S_k$$

$$d(x, y) = \sum_{k=1}^{\infty} 2^{-k} [d_k(x_k, y_k) \wedge 1]$$

for all $x, y \in S$. Let $\{P_n\}$ be a sequence of probability measures on S . For each $k = 1, 2, \dots$, let $\pi_k : S \rightarrow S_k$ be the projection given by $\pi_k(x) = x_k$. Then $\{P_n\}$ is tight in $\mathbf{M}_1(S)$ if and only if $\{P_n \circ \pi_k^{-1}\}$ is tight in $\mathbf{M}_1(S_k)$ for $k = 1, 2, \dots$.

Proof: Refer to Ethier and Kurtz [18].

The next result is on the tightness of probability measures in $\mathbf{M}_1(\mathbf{M}_1(S))$.

Theorem 1.5 *Let S be a complete and separable metric space and $\{P_n\}$ be a sequence of probability measures in $\mathbf{M}_1(\mathbf{M}_1(S))$. Let I_n be the intensity measure of P_n defined by*

$$\langle I_n, f \rangle = \int_{\mathbf{M}_1(S)} \langle m, f \rangle P_n(dm), \quad f \in C_b(S) \quad (1.43)$$

Then $\{P_n\}$ is tight in $\mathbf{M}_1(\mathbf{M}_1(S))$ if and only if $\{I_n\}$ is tight in $\mathbf{M}_1(S)$.

Proof: See, for example, [29].

The final tightness result is on the probability measures on the path space $C([0, \infty), R)$, which are induced by some continuous semimartingales.

A stochastic process $X = \{X(t) : t \geq 0\}$ defined on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ with sample paths in $C([0, \infty), R)$ is called a continuous semimartingale if $X(t)$ can be represented by

$$X(t) = X(0) + V(t) + M(t), \quad t \geq 0, \quad (1.44)$$

where $X(0)$ is a $\mathcal{F}_0$ -measurable random variable, V is a continuous $\mathcal{F}_t$ -adapted process with bounded variation on bounded intervals and $V(0) = 0$, and M is a

continuous $(\mathcal{F}_t, P)$ local martingale with corresponding quadratic variational processes $\langle\langle M \rangle\rangle_t$. In the following we always write $X(0) + V(t)$ as $A(t)$ so that $X(t) = A(t) + M(t)$. We have the following theorem which is mainly due to Aldous and Rebolledo [27].

Theorem 1.6 *Let $\{P_n\}$ be a sequence of probability measures on $C([0, \infty), R)$ which are the laws of the continuous semimartingales $X_n(t) = A_n(t) + M_n(t)$. Suppose the following conditions hold:*

1. *for each $t \geq 0$, the laws $\mathcal{L}(X_n(t))$ of $X_n(t)$ are tight in $M_1(R)$;*
2. *for each positive η, ϵ and T , there exists $\delta > 0$ such that for any stopping time τ_n bounded by T such that*

$$\limsup_{n \rightarrow \infty} \sup_{\theta \leq \delta} P(|A_n(\tau_n + \theta) - A_n(\tau_n)| > \epsilon) \leq \eta$$

and

$$\limsup_{n \rightarrow \infty} \sup_{\theta \leq \delta} P(|\langle M_n \rangle_{\tau_n + \theta} - \langle M_n \rangle_{\tau_n}| > \epsilon) \leq \eta$$

Then $\{P_n\}$ is tight in $M_1(C([0, \infty), R))$.

Proof: Refer to Joffe and Metivier [27].

The book by Ethier and Kurtz [18] contains a very nice and complete presentation on the theory of weak convergence of probability measures on metric spaces.

1.3 Large deviations and statistical mechanics

Throughout this section let the following objects be given:

(P_n) a sequence of probability measures on $\mathcal{B}(S)$.

(a_n) a sequences of positive numbers tending to infinity.

I a function that maps S to $[0, \infty]$.

Definition 1.7 (S, P_n, a_n) is said to form a large deviation system with rate function I , if the following conditions are satisfied:

1. the set $K(s) = \{x \in X : I(x) \leq s\}$ is compact for each $s \geq 0$;
2. for every closed subset F of S

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log P_n(F) \leq - \inf_{x \in F} I(x); \quad (1.45)$$

3. for every open subset G of S

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \log P_n(G) \geq - \inf_{x \in G} I(x); \quad (1.46)$$

As a classical example, we state the following Cramèr theorem. Let μ be a probability measure on R^d such that

$$L(t) = \int_{R^d} \exp(t, x) \mu(dx) < \infty, \quad \text{for all } t \in R^d. \quad (1.47)$$

For $n \geq 1$, let $\mu^{\otimes n}$ be the n -fold product measure on R^{dn} and Let P_n be the image measure of $\mu^{\otimes n}$ under the mapping $R^{dn} \ni (x_1, \dots, x_n) \mapsto \frac{1}{n} \sum_{i=1}^n x_i \in R^d$.

Theorem 1.7 (Cramèr) (R^d, P_n, n) forms a large deviation system with rate function

$$I(x) = L^*(x) \quad x \in R^d, \quad (1.48)$$

where $L^*(x)$ is the Legendre transformation of $L(t)$

$$L^*(x) = \sup_{t \in R^d} \{(x, t) - L(t)\}. \quad (1.49)$$

Proof: See, for example, [45].

One of the most important applications of large deviations is to derive an asymptotic formula for certain integrals in functional space. It was initially formulated by Varadhan [54].

Theorem 1.8 Assume that (S, P_n, a_n) forms a large deviation system with rate function I . Let Φ be a continuous function that maps S into R and satisfies the following condition

$$\lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(\int_{\{x: \Phi(x) \geq L\}} \exp[a_n \Phi(x)] P_n(dx) \right) = -\infty. \quad (1.50)$$

Then

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \int_S \exp[a_n \Phi(x)] P_n(dx) = \sup_{x \in X} [\Phi(x) - I(x)]. \quad (1.51)$$

Proof: See [54].

We now consider the relationship between the large deviation theory and the equilibrium statistical mechanics.

A finite, say n , statistical mechanical model is usually defined by the object $(S^{\otimes n}, H_n(x), \mu_n)$, where $S^{\otimes n}$ is a configuration space of the physical quantities $x = (x_1, \dots, x_n)$, $H_n(x)$, $x \in S^{\otimes n}$, is the Hamiltonian which specifies the interaction energy among $x_1, \dots, x_n$, and μ_n is a probability measure on $S^{\otimes n}$ which specifies the distribution of $x_1, \dots, x_n$ without interactions. The relationship among these n physical quantities $x_1, \dots, x_n$ is usually described by the Gibbs distribution

$$P_{n,\beta}(dx) = \frac{1}{Z_{n,\beta}} \exp\{-\beta H_n(x)\} \mu_n(dx) \quad (1.52)$$

where $\beta > 0$ denotes the inverse temperature, and

$$Z_{n,\beta} = \int_{S^{\otimes n}} \exp\{-\beta H_n(x)\} \mu_n(dx) \quad (1.53)$$

is called the partition function.

The first fundamental task in the equilibrium statistical mechanics is to compute the free energy function $f(\beta)$ which is defined by

$$f(\beta) = -\frac{1}{\beta} \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,\beta} \quad (1.54)$$

The following theorem shows that under proper conditions the classical large deviation principles can be used in determining the free energy.

Theorem 1.9 *Suppose that the following three conditions are satisfied*

1. *X is a complete separable metric space.*
2. *For each n there exists a measurable mapping $\xi_n : S^{\otimes n} \rightarrow X$ such that (X, Q_n, n) forms a large deviation system with rate function $I(x)$, where*

$$Q_n(dx) = \mu_n\{x : \xi_n(x) \in dx\}.$$

3. *There exists a continuous function $\Phi : X \rightarrow R$ satisfying the condition (1.50) such that*

$$-H_n(x) = n\Phi(\xi_n(x)) + o(n) \quad (1.55)$$

uniformly for $x \in S^{\otimes n}$. Then the free energy $f(\beta)$ exists and is given by the variational formula

$$-\beta f(\beta) = \sup_{x \in X} \{\beta \Phi(x) - I(x)\} \quad (1.56)$$

Proof: Since Q_n is the distribution of ξ_n with respect to the probability measure μ_n . The free energy can be expressed as the $n \rightarrow \infty$ limit of the integral

$$\begin{aligned} -\beta f(\beta) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,\beta} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{S^{\otimes n}} \exp\{-\beta H_n(x)\} \mu_n(dx) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{S^{\otimes n}} \exp\{\beta[n\Phi(\xi_n(x)) + o(n)]\} \mu_n(dx) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_X \exp\{\beta n\Phi(m)\} Q_n(dm) \end{aligned} \quad (1.57)$$

Now $(X, Q_n(dm), n)$ forms a large deviation system with rate function $I(x)$, one can use the Varadhan theorem 1.8 to complete the proof. $\square$

For example, in the continuous Curie-Weiss ferromagnetic mean field model

$$S = X = R$$

$$H_n(x) = -\frac{\theta}{2n} \sum_{i,j=1}^n x_i x_j$$

$$\mu_n = \prod_{i=1}^n \mu(dx_i)$$

where $x = (x_1, \dots, x_n)$ denotes n spins in R , θ is a positive constant, and μ is a probability measure on R satisfying the condition

$$\log \int_R \exp(tx) \mu(dx) < \infty \quad \text{for all } t \in R. \quad (1.58)$$

The corresponding Gibbs distribution has the form:

$$P_{n,\beta}(dx) = \frac{1}{Z_{n,\beta}} \exp\{-\beta H_n(x)\} \prod_{i=1}^n \mu(dx_i) \quad (1.59)$$

Let $\xi_n : R^n \rightarrow R$ be the mean magnetization defined by

$$\xi_n = \frac{1}{n} \sum_{i=1}^n x_i \quad (1.60)$$

and let

$$\Phi(m) = \frac{1}{2} \theta m^2, \quad m \in R. \quad (1.61)$$

Then the free energy of the Curie-Weiss model can be expressed as

$$\begin{aligned} -\beta f(\beta) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,\beta} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{R^n} \exp\{-\beta H_n(x)\} \mu_n(dx) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{R^n} \exp\left\{\frac{1}{2} \beta \theta n \left(\frac{1}{n} \sum_{i=1}^n x_i\right)^2\right\} \prod_{i=1}^n \mu(dx_i) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_R \exp\{\beta n \Phi(m)\} Q_n(dm) \end{aligned} \quad (1.62)$$

Then according to the classical Cramér theorem 1.7 and the theorem 1.9 we have

$$\begin{aligned} f(\beta) &= -\frac{1}{\beta} \sup_{m \in R} \{\beta \Phi(m) - L^*(m)\} \\ &= \inf_{m \in R} \left\{ \frac{1}{\beta} L^*(m) - \Phi(m) \right\} \end{aligned} \quad (1.63)$$

where $L^*(m)$ is given by (1.49).

The other main application of large deviations to statistical mechanics is to find the asymptotic behavior of the distribution of the process ξ_n under the probability measure $P_{n,\beta}$. Since the process ξ_n often represents the physical order parameter, the information about the asymptotic behavior of this process usually provides important insight into the model. Let F be the function on X defined by

$$F(m) = \{\beta\Phi(m) - I(m)\} \quad (1.64)$$

and let

$$K_\beta = \{m \in X : F(m) \text{ attains its maximum at } m\} \quad (1.65)$$

Denote

$$Q_{n,\beta}(dm) = P_{n,\beta}\{x : \xi_n(x) \in dm\} \quad (1.66)$$

Then we have the following result.

Theorem 1.10 *Suppose that the conditions in theorem 1.9 are satisfied. Then the sequence of probability measures $\{Q_{n,\beta}\}$, is relatively compact in $\mathbf{M}_1(X)$. Let Q_β be any limit point of the sequence $Q_{n,\beta}$. Then*

$$Q_\beta(K_\beta) = 1$$

Proof: Refer to [17].

As a consequence of theorem 1.10, if the set K_β contains only a single element, say x_β , then

$$Q_{n,\beta} \Rightarrow \delta_{x_\beta} \quad \text{as } n \rightarrow \infty \quad (1.67)$$

Again, consider the continuous Curie-Weiss Ising spin model. Let K_β be the set of points m in R where the free energy (1.63) attains its minimum there. Then it is known that [17] there exists a critical temperature β_c such that

$$K_\beta = \begin{cases} \{0\} & \text{if } 0 < \beta < \beta_c \\ \{+m_\beta, -m_\beta\} & \text{if } \beta > \beta_c \end{cases} \quad (1.68)$$

where for $\beta > \beta_c$, m_β is the unique positive solution of the fixed point equation

$$m = \frac{d}{dt} L(t) \Big|_{t=m_\beta} = \frac{\int_R x \exp\{\beta \theta m x\} \mu(dx)}{\int_R \exp\{\beta \theta m x\} \mu(dx)} \quad (1.69)$$

Let $Q_{n,\beta}$ be the distribution of the order parameter ξ_n with respect to the probability measure $P_{n,\beta}$. Then it follows from the structure of K_β that

$$Q_{n,\beta} \Rightarrow \begin{cases} \delta_0 & \text{if } 0 < \beta < \beta_c \\ \frac{1}{2}\delta_{+m_\beta} + \frac{1}{2}\delta_{-m_\beta} & \text{if } \beta > \beta_c \end{cases} \quad (1.70)$$

as $n \rightarrow \infty$. Thus the Curie-Weiss model has a phase transition at the critical inverse temperature $\beta = \beta_c$.

We end this chapter by stating three frequently used technical results in this thesis. The first one is the Itô's formula from stochastic calculus. The second one is the Gronwall's inequality from analysis. The third one is the Burkholder-Davis-Gundy inequality from martingale theory.

Itô's formula Suppose that $X = (X_1, \dots, X_d)$ is a R^d -valued process defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ such that for $i = 1, \dots, d$,

$$X_i(t) = X_i(0) + V_i(t) + M_i(t), \quad t \geq 0, \quad (1.71)$$

where $X_i(0)$ is a $\mathcal{F}_0$ -measurable random variable, V_i is a continuous $\mathcal{F}_t$ -adapted process with bounded variation on bounded intervals and $V_i(0) = 0$, and M_i is a continuous $(\mathcal{F}_t, P)$ -local martingale. Let $f \in C^2(R^d)$. Then

$$\begin{aligned} f(X(t)) &= f(X(0)) + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(X(s)) dV_i(s) + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(X(s)) dM_i(s) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(X(s)) d\langle M_i, M_j \rangle_s, \end{aligned} \quad (1.72)$$

Gronwall's inequality Suppose that $f(t), t \in [0, T]$, is a non-negative function for which the inequality

$$f(t) \leq K + \alpha \int_0^t f(s) ds, \quad t \in [0, T] \quad (1.73)$$

holds for some $K, \alpha > 0$, where

$$\int_0^t f(s) ds < \infty \quad (1.74)$$

Then

$$f(t) \leq K \exp(\alpha t) \quad (1.75)$$

for every $t \in [0, T]$.

Burkholder-Davis-Gundy inequality For every $p \in (0, \infty)$, there exist two constants c_p and C_p such that, for all continuous local martingales M vanishing at zero,

$$c_p E[\langle M \rangle_t^{\frac{p}{2}}] \leq E[(\sup_{t \geq 0} |M(t)|)^p] \leq C_p E[\langle M \rangle_t^{\frac{p}{2}}]. \quad (1.76)$$

The proof of the first two results can be found in the book by Ethier and Kurtz [18].

The last result can be found in the book by Revuz and Yor [38].

Chapter 2

Spin Glass Problem

In this chapter we give a short introduction to the spin glass problem and some related mathematical work. Our attention will mainly focus on the infinite-ranged mean field model proposed by Sherrington and Kirkpatrick in 1975 [41] and the dynamical mean field theory proposed by Sompolinsky and Zippelius in 1982 [43]. While these theories have made significant contributions to the understanding of spin glass phenomena, there remain several important open questions. These questions are interesting from the mathematical point of view, and the objective of this chapter is to expose these problems as motivations for this thesis, and at the same time hope to raise further research interest in this field. For a detailed review on the spin glass problem we refer to the book by Mezard, Parisi and Virasoro [33] and the references cited there.

This chapter consists of three sections. In section 1, we introduce the physical models of the spin glasses. We discuss the mean field model and its replica solutions. In section 2, we present the dynamical mean field theory at a physical level of rigor. In section 3, we discuss some of the mathematical work related to the spin glass problem.

2.1 Physical model

Spin glasses are magnetic materials which exhibit a magnetic characteristic somewhat similar to that displayed by the material glass. In particular this involves the notion of “frozen disorder”. Although disorder also occurs in liquids, the atoms are mobile there while in a glass their positions are frozen. There is also a distinguishing difference between the spins in a normal magnet (such as paramagnet or ferromagnet) and those in a spin glass. In a normal magnet the magnetic moments usually exhibit some kind of order but the magnetic moments in the spin glass are believed to be frozen into thermal equilibrium orientations with no average long-range order.

The basic reason for the occurrence of the spin glass phase is that the interactions among the spins in the nearest-neighbourhood are random and their signs fluctuate. The latter is important, because a positive interaction favours the parallel alignment of spins while a negative interaction favours the opposite. Because of this random interaction, it becomes difficult in these materials to find an overall spin alignment that can simultaneously satisfy all the exchange bonds. This leads to the so called frustration phenomenon. Therefore, two basic features of the spin glass behavior can be characterized as having randomness in the interactions, and frustration due to the sign fluctuation of the random interactions.

In order to understand the collective behavior of this randomly interacting spin system Edwards and Anderson in 1975 [15] proposed a simplified spin glass model based on the Hamiltonian:

$$H = - \sum_{(i,j)} J_{i,j} S_i S_j, \quad (2.1)$$

where $S_i \in \{+1, -1\}$ denotes the Ising spin at location i and the summation (i, j) is over the nearest-neighbours; $\{J_{i,j}\}$ is a set of Gaussian random variables with mean zero. They proposed that the system may undergo a phase transition into a state, in which the local spins are frozen in random directions, which can be described by

the "EA order parameter"

$$q_{EA} = [\langle S_i^2 \rangle_T]_J \quad (2.2)$$

where $\langle \cdot \rangle_T$ denotes the thermal average and $[\cdot]_J$ denotes the average over the random interactions. It is generally accepted among the physicists today that this model is the simplest theoretical model which possesses the essential features of the spin glasses and can be studied analytically.

The usual approach to investigate the different phases of a physical system starts with the so called mean field approximation in which a spin is allowed to interact with all of the other spins. The mean field theory of spin glasses was proposed by Sherrington and Kirkpatrick in 1975 [41]. In the simplest case, their infinite-range Ising spin glass model (abbreviated as SK model) consists of n Ising spins $S_1, \dots, S_n$ with Hamiltonian:

$$H_n^{\{J_{i,j}\}}(S) = -\frac{1}{\sqrt{n}} \sum_{i,j=1}^n J_{i,j} S_i S_j \quad (2.3)$$

where $S = (S_1, \dots, S_n)$ and $\{J_{i,j}\}$ is a set of i.i.d. standard Gaussian random variables with probability density

$$P(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (2.4)$$

The central problem is to determine the average free energy per spin $f(\beta)$ at inverse temperature $\beta = \frac{1}{T}$ in the thermodynamic limit $n \rightarrow \infty$

$$f(\beta) = -\frac{1}{\beta} \lim_{n \rightarrow \infty} \frac{1}{n} E[\log Z_{n,\beta}^{\{J_{i,j}\}}] \quad (2.5)$$

where E denotes the average over the random interactions $\{J_{i,j}\}$ and $Z_{n,\beta}^{\{J_{i,j}\}}$ is the usual partition function defined as

$$\begin{aligned} Z_{n,\beta}^{\{J_{i,j}\}} &= \sum_S \exp\{-\beta H_n^{\{J_{i,j}\}}(S)\} \\ &= \sum_S \exp\left\{\beta \frac{1}{\sqrt{n}} \sum_{i,j=1}^n J_{i,j} S_i S_j\right\} \end{aligned} \quad (2.6)$$

where the summation $\sum_S$ is taken over all configurations of $S = (S_1, \dots, S_n)$. The direct computation of (2.5) is not an easy task because it involves the average over $\log Z_{n,\beta}^{\{J_{i,j}\}}$ rather than over $Z_{n,\beta}^{\{J_{i,j}\}}$. The general framework in physical literature to deal with this problem is to use the so called replica method. In this method one first introduces the partition function $Z_{n,\beta}^{m,\{J_{i,j}\}}$ of m identical replicas of the same system, interacting with the same random Hamiltonian

$$Z_{n,\beta}^{m,\{J_{i,j}\}} = \sum_{S^1} \cdots \sum_{S^m} \exp\left\{\beta \frac{1}{\sqrt{n}} \sum_{i,j=1}^n J_{i,j} \sum_{k=1}^m S_i^k S_j^k\right\} \quad (2.7)$$

where $S^k = (S_1^k, \dots, S_n^k)$ for $k = 1, \dots, m$ and the spin variables S_i^k now carry two indices: the upper one indicates the k th replica and goes from 1 to m , and the lower one indicates the site in the k th replica and goes from 1 to n . By using the identity

$$\log Z = \lim_{m \rightarrow 0} \frac{Z^m - 1}{m} \quad (2.8)$$

the average over $\log Z_{n,\beta}^{\{J_{i,j}\}}$ can be obtained from the relatively simpler average over $Z_{n,\beta}^{m,\{J_{i,j}\}}$

$$\begin{aligned} E[\log Z_{n,\beta}^{\{J_{i,j}\}}] &= \lim_{m \rightarrow 0} \frac{E[(Z_{n,\beta}^{\{J_{i,j}\}})^m] - 1}{m} \\ &= \lim_{m \rightarrow 0} \frac{E[Z_{n,\beta}^{m,\{J_{i,j}\}}] - 1}{m} \end{aligned} \quad (2.9)$$

The averaging free energy per spin $f(\beta)$ is then determined by

$$f(\beta) = -\frac{1}{\beta} \lim_{n \rightarrow \infty} \frac{1}{n} \lim_{m \rightarrow 0} \frac{E[Z_{n,\beta}^{m,\{J_{i,j}\}}] - 1}{m} \quad (2.10)$$

Thus the basic steps in the replica method consist in computing the averaged partition function $E[Z_{n,\beta}^{m,\{J_{i,j}\}}]$ for generic integer m , performing the analytic continuation from the integers up to the point $m = 0$, and handling the two limits.

For a positive integer m , the average with respect to the $\{J_{i,j}\}$ of the quantity $Z_{n,\beta}^{m,\{J_{i,j}\}}$ can be easily done and the result can be expressed as

$$E[Z_{n,\beta}^{m,\{J_{i,j}\}}] = \sum_{S^1} \cdots \sum_{S^m} \exp\left\{\frac{\beta^2}{2n} \sum_{i,j=1}^n \left(\sum_{k=1}^m S_i^k S_j^k\right)^2\right\}$$

$$\begin{aligned}
&= \sum_{S^1} \cdots \sum_{S^m} \exp\left\{\frac{\beta^2}{2n} \left[\sum_{\substack{k,l=1 \\ k \neq l}}^m \left(\sum_{i=1}^n S_i^k S_i^l \right)^2 + mn^2 \right]\right\} \\
&= \exp\left\{\frac{\beta^2 mn}{2}\right\} \sum_{S^1} \cdots \sum_{S^m} \exp\left\{\frac{\beta^2}{2n} \sum_{\substack{k,l=1 \\ k \neq l}}^m \left(\sum_{i=1}^n S_i^k S_i^l \right)^2\right\} \quad (2.11)
\end{aligned}$$

Using the identity on Gaussian integral

$$\exp(\lambda a^2) = \frac{1}{\sqrt{2\pi}} \int_R \exp\left\{-\frac{1}{2}x^2 + (2\lambda)^{\frac{1}{2}}ax\right\} dx \quad (2.12)$$

Eq.(2.11) can be put into the following integral form

$$\begin{aligned}
E[Z_{n,\beta}^{m,\{J_i, J\}}] &= \beta \left(\frac{n}{2\pi}\right)^{\frac{m(m-1)}{2}} \exp\left\{\frac{\beta^2 nm}{2}\right\} \int_{R^{m(m-1)}} \exp\left\{-n \frac{\beta^2}{2} \sum_{\substack{k,l=1 \\ k \neq l}}^m q_{k,l}^2\right\} \\
&\quad \sum_{S^1} \cdots \sum_{S^m} \exp\left\{\beta^2 \sum_{\substack{k,l=1 \\ k \neq l}}^m q_{k,l} \sum_{i=1}^n S_i^k S_i^l\right\} \prod_{\substack{k,l=1 \\ k \neq l}}^m dq_{k,l} \\
&= \int_{R^{m(m-1)}} \exp\left\{-n[\Phi_m(\{q_{k,l}\}) - \frac{m(m-1)}{2} \frac{1}{n} \log\left(\frac{\beta}{2\pi} n\right)]\right\} \prod_{\substack{k,l=1 \\ k \neq l}}^m dq_{k,l} \quad (2.13)
\end{aligned}$$

where

$$\begin{aligned}
\Phi_m(\{q_{k,l}\}) &= \beta \left(\frac{n}{2\pi}\right)^{\frac{m(m-1)}{2}} - \frac{\beta^2 m}{2} + \frac{\beta^2}{2} \sum_{\substack{k,l=1 \\ k \neq l}}^m q_{k,l}^2 \\
&\quad - \log \sum_{S_1=\pm 1} \cdots \sum_{S_m=\pm 1} \exp\left\{\beta^2 \sum_{\substack{k,l=1 \\ k \neq l}}^m q_{k,l} S_k S_l\right\} \quad (2.14)
\end{aligned}$$

Thus the averaged free energy per spin $f(\beta)$ can be expressed as

$$\begin{aligned}
f(\beta) &= -\frac{1}{\beta} \lim_{n \rightarrow \infty} \frac{1}{n} \lim_{m \rightarrow 0} \frac{1}{m} \int_{R^{m(m-1)}} \exp\left\{-n[\Phi_m(\{q_{k,l}\}) - \frac{m(m-1)}{2} \frac{1}{n} \log\left(\frac{\beta}{2\pi} n\right)]\right\} \prod_{\substack{k,l=1 \\ k \neq l}}^m dq_{k,l} \quad (2.15)
\end{aligned}$$

If we assume that the limit $m \rightarrow 0$ and the thermodynamic limit $n \rightarrow \infty$ can be interchanged (a discussion on the condition that this interchange is correct can be found in [53]), then (2.15) can be written as

$$f(\beta) = -\frac{1}{\beta} \lim_{m \rightarrow 0} \frac{1}{m} \lim_{n \rightarrow \infty} \frac{1}{n} \int_{R^{m(m-1)}} \exp\{-n[\Phi_m(Q) + o(n)]\} dQ \quad (2.16)$$

where we have used Q to denote the $m \times m$ symmetric matrix $\{q_{k,l}\}$ with $q_{k,l} = 0$ if $k \neq l$. The method of steepest descents suggests that the expression (2.16) can be written as,

$$f(\beta) = \frac{1}{\beta} \lim_{m \rightarrow 0} \frac{1}{m} \inf_Q \Phi_m(Q) \quad (2.17)$$

where $\inf_Q \Phi_m(Q)$ is the result of applying the saddle-point method to the integral in (2.15). For positive integer m (different from zero) the minimum of Φ can be found and the matrix has the form [53]:

$$Q_{k,l} = \begin{cases} q & \text{if } k \neq l \\ 0 & \text{if } k = l \end{cases} \quad (2.18)$$

which is usually called the replica symmetric solution of the SK spin glass mean field model. Now the big question is how to extend the minimum solution from integer m to all real numbers up to $m = 0$. The original SK mean field theory simply treats m in (2.17) as a real variable. In this way one finally derives the following equation for the free energy

$$f(\beta) = -\frac{1}{4}\beta^2(1-q)^2 + \int_{-\infty}^{\infty} \log[2 \cosh(\beta\sqrt{q}x)] \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}) dx \quad (2.19)$$

where q satisfies the fixed point equation

$$q = \int_{-\infty}^{\infty} \tanh^2(\beta\sqrt{q}x) \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}) dx. \quad (2.20)$$

Equation (2.20) is the basic fixed point equation for the SK spin glass model. It is known that (see chapter 6 for a proof) when $\beta \leq \beta_c = 1$ the only solution of (2.20) is $q = 0$. When $\beta > \beta_c$ there is a unique positive $q > 0$ which corresponds the spin glass phase. Although the mean field theory of the SK spin glass reproduced many of the desired features of the Edward-Anderson spin glass model, it was soon discovered that at low temperature the solution is not stable and also yields the negative entropy which contradicts basic physical principles. There are several reasons for the failure

of the SK mean field solution based on the replica method. As we can see from the above brief exposition, the interchange of the limit $m \rightarrow 0$ and the thermodynamic limit $n \rightarrow \infty$ is doubtful, the saddle-point method used in the integral is unjustified. The most serious problem lies in the continuous extension from the integer m to real number. For more details on what is wrong in the replica approach we refer the paper by Van Hemmen and Palmer [53].

These subtle problems raised considerable research interests among physicists during the past eighteen years and led to many important theoretical discoveries on the nature of the spin glass phase. One important attempt to improve the SK mean field theory was the paper by Thouless, Anderson and Palmer in 1977 [50]. Among other things, they demonstrated there exist many solutions for the SK mean field, which implies the existence of many equivalent or near-equivalent spin configurations. The other two important papers concerning the nature of the order parameter in spin glass was published by Parisi in 1979 and 1983 [35][36]. He not only formalized a new concept called replica symmetric breaking, but also provided an ansatz for implementing the symmetric breaking. Finally, Mezard et al [33] discovered that the family of spin glass solutions had a hierarchy of its own, which implies that the equilibrium states (of which there are many) of spin glass phase exhibit a very rich structure. They possess the so called ultrametric property which is related to a branching process [33]. We will not go into the details of the development along these lines. Instead, we shall focus on the dynamical approach proposed by Sompolinsky and Zippelius in 1982 [43] to study the relaxation dynamics of spin glasses. This dynamical approach is attractive because it provides a means for calculating average thermodynamic quantities without using the unphysical replica trick and also yields the time-dependent features of the spin glass state.

2.2 Dynamical mean field theory

In this section we try to use Itô stochastic differential equations to formulate the dynamical mean field theory. Due to the nature of this difficult problem some of the arguments in the following exposition are only heuristic.

In the dynamical mean field approach, one starts with a soft spin version of the SK model associated with the Hamiltonian

$$H(x_1, \dots, x_n) = \sum_{i=1}^n U(x_i) - \frac{1}{\sqrt{n}} \sum_{i,j=1}^n J_{i,j} x_i x_j \quad (2.21)$$

where the n soft spins $x_1, \dots, x_n$ are allowed to vary continuously in R , $\{J_{i,j}\}$ is a set of i.i.d. random variables with a standard Gaussian distribution

$$P\{J_{i,j} \in dx\} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx. \quad (2.22)$$

$U(x)$ is a proper potential function, for instance, one can take $U(x) = ax^4 - bx^2$ for some positive constants a and b . The dynamics is governed by the following Itô stochastic differential equations:

$$dx_i(t) = -\frac{\delta H(x_1(t), \dots, x_n(t))}{\delta x_i(t)} dt + dw_i(t), \quad i = 1, \dots, n \quad (2.23)$$

where $w_1, \dots, w_n$ are n independent 1-dimensional Wiener processes. The physical quantity of particular interest is the averaged two-spin correlation function

$$C_{i,j}(t, s) = E_J(E_w[x_i(t)x_j(s)]) \quad (2.24)$$

where the expectation E_w is taken with respect to the probability measure which defines the Wiener processes and the expectation E_J is taken with respect to the probability measure (2.22). The basic idea in physical literatures is to define a generating functional which can generate most of the physical quantities of interest including the averaged two-spin correlation function. For each given $\{J_{i,j}\}$, the

generating functional is usually defined by

$$\Phi_n^{(J, \gamma)}(\theta, \hat{\theta}) = \int_{C([0, \infty), R^n)} \int_{C([0, \infty), R^n)} \exp\left\{\sum_{i=1}^n \int_0^\infty [\theta_i(t)x_i(t) + \hat{\theta}_i(t)i\dot{x}_i(t)]dt\right\} \exp\{V^{(J, \gamma)}(x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n)\} dx_1 \dots dx_n d\dot{x}_1 \dots d\dot{x}_n \quad (2.25)$$

where $\theta = \{(\theta_1(t), \dots, \theta_n(t)), t \geq 0\}$, $\hat{\theta} = \{(\hat{\theta}_1(t), \dots, \hat{\theta}_n(t)), t \geq 0\}$ and

$$\begin{aligned} V^{(J, \gamma)}(x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n) &= \frac{1}{2} \sum_{i=1}^n \int_0^\infty \dot{x}_i^2(t) dt \\ &- \sum_{i=1}^n \int_0^\infty i\dot{x}_i(t) \left[\dot{x}_i(t) + \frac{\delta H(x_1(t), \dots, x_n(t))}{\delta x_i(t)} \right] dt \\ &+ \frac{1}{2} \sum_{i=1}^n \int_0^\infty \frac{\delta^2 H(x_1(t), \dots, x_n(t))}{\delta x_i^2(t)} dt \end{aligned} \quad (2.26)$$

Most quantities of physical interest can now be generated from this generating functional, for instance, the mean of the spin and the two-spin correlation are determined by

$$E[x_i(t)] = \frac{\delta \log \Phi_n^{(J, \gamma)}(\theta, \hat{\theta})}{\delta \theta_i(t)} \Big|_{\theta=\hat{\theta}=0} \quad (2.27)$$

$$E[x_i(t)x_i(s)] = \frac{\delta^2 \log \Phi_n^{(J, \gamma)}(\theta, \hat{\theta})}{\delta \theta_i(t) \delta \theta_i(s)} \Big|_{\theta=\hat{\theta}=0} \quad (2.28)$$

We now show that the quantity $\Phi_n^{(J, \gamma)}(\theta, \hat{\theta})$ is related to a Gibbs type measure on functional space. To derive the formalism, we start with the fact that heuristically the Wiener processes $w_1, \dots, w_n$ have the joint distribution on the path space $C([0, \infty), R^n)$ of the following form

$$P_w(dw_1, \dots, dw_n) = \frac{1}{Z} \exp\left\{-\frac{1}{2} \sum_{i=1}^n \int_0^\infty |\dot{w}_i(t)|^2 dt\right\} dw_1 \dots dw_n \quad (2.29)$$

where Z is the normalization constant

$$Z = \int_{C([0, \infty), R^n)} \exp\left\{-\frac{1}{2} \sum_{i=1}^n \int_0^\infty |\dot{w}_i(t)|^2 dt\right\} dw_1 \dots dw_n \quad (2.30)$$

The expression (2.29) above is only heuristic because the quantities on the right hand side of (2.29) make no mathematical sense. In particular, dw_i denotes the translation invariant measure on the path space $C([0, \infty), R)$ which does not exist. $\dot{w}_i(t)$ denotes the derivative of $w_i(t)$ which also fails to exist P_w -a.s. We now write $w_1, \dots, w_n$ in terms of $x_1, \dots, x_n$, using the equations of motion (2.23) and change the integration variables from the $w_1, \dots, w_n$ to the $x_1, \dots, x_n$, then for each $J_{i,j}$ the $x_1, \dots, x_n$ have the joint distribution

$$P_x^{(J_{i,j})}(dx_1, \dots, dx_n) = \frac{1}{Z} \exp\left\{-\frac{1}{2} \sum_{i=1}^n \int_0^\infty \left[\dot{x}_i(t) + \frac{\delta H(x_1(t), \dots, x_n(t))}{\delta x_i(t)}\right]^2 dt\right\} J(x_1, \dots, x_n) dx_1 \cdots dx_n \quad (2.31)$$

where $J(x_1, \dots, x_n)$ is the functional Jacobian which is shown to be [47]

$$J(x_1, \dots, x_n) = \exp\left\{-\frac{1}{2} \sum_{i=1}^n \int_0^\infty \frac{\delta^2 H(x_1(t), \dots, x_n(t))}{\delta x_i^2(t)} dt\right\} \quad (2.32)$$

and the normalization factor Z can now be written as

$$Z = \int_{C([0, \infty), R^n)} \exp\left\{-\frac{1}{2} \sum_{i=1}^n \int_0^\infty \left[\dot{x}_i(t) + \frac{\delta H(x_1(t), \dots, x_n(t))}{\delta x_i(t)}\right]^2 dt\right\} J(x_1, \dots, x_n) dx_1 \cdots dx_n \quad (2.33)$$

By using a Gaussian integral identity [44] (2.29) can be expressed as

$$\begin{aligned} Z &= \int_{C([0, \infty), R^n)} \int_{C([0, \infty), R^n)} \exp\left\{-\frac{1}{2} \sum_{i=1}^n \int_0^\infty \dot{\hat{x}}_i^2(t) dt\right. \\ &\quad \left.+ \sum_{i=1}^n \int_0^\infty i \hat{x}_i(t) \left[\dot{x}_i(t) + \frac{\delta H(x_1(t), \dots, x_n(t))}{\delta x_i(t)}\right] dt\right\} \\ &\quad J(x_1, \dots, x_n) dx_1 \cdots dx_n d\hat{x}_1 \cdots d\hat{x}_n \\ &= \int_{C([0, \infty), R^n)} \int_{C([0, \infty), R^n)} \exp\left\{-V^{(J_{i,j})}(x_1, \dots, x_n, \hat{x}_1 \cdots \hat{x}_n)\right\} \\ &\quad dx_1 \cdots dx_n d\hat{x}_1 \cdots d\hat{x}_n \end{aligned} \quad (2.34)$$

where $V^{(J_{i,j})}(x_1, \dots, x_n, \hat{x}_1 \cdots \hat{x}_n)$ is given by (2.26). The expressions (2.26) and (2.34) suggest that the quantity Z plays the role of the partition function for the

$2n$ dynamical variables $x_1(\cdot), \dots, x_n(\cdot), \hat{x}_1(\cdot), \dots, \hat{x}_n(\cdot)$ and $V^{(J_{i,j})}$ plays the role of a nonequilibrium Hamiltonian. Thus we may consider the quantity $\Phi_n^{(J_{i,j})}(\theta, \hat{\theta})$ defined by (2.25) as the associated generating functional of the following Gibbs type measure on the functional space $C([0, \infty), R^{2n})$

$$\frac{1}{Z} \exp\{-V^{(J_{i,j})}(x_1, \dots, x_n, \hat{x}_1, \dots, \hat{x}_n)\} dx_1 \cdots dx_n d\hat{x}_1 \cdots d\hat{x}_n \quad (2.35)$$

The correlation generated by $\Phi_n^{(J_{i,j})}(\theta, \hat{\theta})$ still depends on the random variables $J_{i,j}$. Since one is interested in the averaged correlation it is natural to define the averaged generating functional by

$$\Phi_n(\theta, \hat{\theta}) = E_J[\Phi_n^{(J_{i,j})}(\theta, \hat{\theta})] \quad (2.36)$$

Note that $\Phi_n^{(J_{i,j})}(0,0) = Z$ does not depend on the random variables $J_{i,j}$, the average with respect to $\{J_{i,j}\}$ can be done directly on $\Phi_n^{(J_{i,j})}(\theta, \hat{\theta})$. A straightforward integration then yields

$$\begin{aligned} \Phi_n(\theta, \hat{\theta}) &= \int_{C([0, \infty), R^n)} \int_{C([0, \infty), R^n)} \exp\{L(x, \hat{x}, \theta, \hat{\theta}) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \int_0^\infty \int_0^\infty i\hat{x}_i(t)i\hat{x}_i(s)x_i(t)x_i(s)dt ds\} \\ &\quad dx_1 \cdots dx_n d\hat{x}_1 \cdots d\hat{x}_n \end{aligned} \quad (2.37)$$

where

$$\begin{aligned} L(x, \hat{x}, \theta, \hat{\theta}) &= \sum_{i=1}^n \int_0^\infty \left\{ \theta_i(t)x_i(t) + \hat{\theta}_i(t)i\hat{x}_i(t) - \frac{1}{2}\hat{x}_i^2(t) \right. \\ &\quad \left. + i\hat{x}_i(t)[\dot{x}_i(t) - U'(x_i(t))] - \frac{1}{2} \frac{\delta^2 H(x_1(t), \dots, x_n(t))}{\delta x_i^2(t)} \right\} dt \end{aligned} \quad (2.38)$$

We now write the interaction term in (2.37) as squares of sum of the quantities

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n i\hat{x}_i(t)i\hat{x}_i(s)x_i(t)x_i(s) &= \frac{1}{4n} \sum_{i=1}^n (i\hat{x}_i(t)i\hat{x}_i(s) + x_i(t)x_i(s))^2 \\ &\quad - \frac{1}{4n} \sum_{i=1}^n i\hat{x}_i(t)i\hat{x}_i(s) - x_i(t)x_i(s) \end{aligned} \quad (2.39)$$

Then, by using a Gaussian transformation to decouple the square and introducing four auxiliary fields $Q_r(t, s)$, $r = 1, \dots, 4$ (see [43] for details). This will yields

$$\begin{aligned} \Phi(\theta, \hat{\theta}) &= \int \prod_{r=1}^4 DQ_r(t, s) \exp\left\{-n \int_0^\infty \int_0^\infty [Q_1(t, s)Q_2(t, s) + Q_3(t, s)Q_4(t, s)] dt ds\right. \\ &\quad \left. + \log \int_{C([0, \infty), R^n)} \int_{C([0, \infty), R^n)} \exp\{L(x, \hat{x}, \theta, \hat{\theta}, Q)\} \right. \end{aligned} \quad (2.40)$$

where $Q = \{Q_r(\cdot, \cdot)\}_{r=1}^4$ and

$$\begin{aligned} L(x, \hat{x}, \theta, \hat{\theta}, Q) &= L(x, \hat{x}, \theta, \hat{\theta}) + \int_0^\infty \int_0^\infty \sum_{i=1}^n [Q_1(t, s) i\hat{x}_i(t) i\hat{x}_i(s) + Q_2(t, s) i x_i(t) x_i(s) \\ &\quad + Q_3(t, s) i\hat{x}_i(t) x_i(s) + Q_4(t, s) i\hat{x}_i(s) x_i(t)] dt ds \end{aligned} \quad (2.41)$$

In the limit $n \rightarrow \infty$, the interaction over Q_r can be performed by using the method of steepest descent, which amounts to substituting for Q_r their stationary point values Q_r^s . After some heavy algebraic computation and physical arguments one ends up with the following one-site generating functional (compare with (2.37), the interaction term there has gone)

$$\Phi_{one}(\theta, \hat{\theta}) = \int_{C([0, \infty), R)} \int_{C([0, \infty), R)} \exp\{L_{one}(x, \hat{x}, \theta, \hat{\theta})\} dx d\hat{x} \quad (2.42)$$

where

$$\begin{aligned} L_{one}(x, \hat{x}, \theta, \hat{\theta}) &= - \int_0^\infty \int_0^\infty [\delta(t-s) + \frac{1}{2} C(t, s)] \hat{x}(t) \hat{x}(s) dt ds \\ &\quad + \int_0^\infty \left\{ i\hat{x}(t) [\dot{x}_i(t) + U'(x(t))] - \frac{1}{2} \frac{\delta^2 H}{\delta x^2(t)} \right\} dt \\ &\quad + \int_0^\infty (\theta(t)x(t) + \hat{\theta}(t)\hat{x}(t)) dt \end{aligned} \quad (2.43)$$

where $C(t, s)$ is the averaged correlation function between $x(t)$ and $x(s)$ determined self-consistently from (2.43).

From (2.43) the corresponding stochastic differential equation of the single-site motion has the form

$$dx(t) = -U'(x(t)) + \xi(t) + dw(t) \quad (2.44)$$

where $w(t)$ is the 1-dimensional Wiener process and $\xi(t)$ is a new Gaussian noise with zero mean and variance

$$E[\xi(t)\xi(s)] = C(t, s) = E[x(t)x(s)] \quad (2.45)$$

which is independent of the $w(t)$. From this limit equation one can then study the relaxational dynamics of the spin glasses. In particular, the relaxational dynamics is not exponential. Instead, it has the form of $t^{-\nu}$ at all temperatures below a critical temperature T_c , where ν depends on the temperature [43].

2.3 Related mathematical work

From the stochastic differential equation point of view, the dynamical mean field approach to the spin glass problem suggests one to consider the following stochastic differential equations in R :

$$dx_i^n(t) = b(x_i^n(t))dt + \frac{1}{\sqrt{n}} \sum_{j=1}^n J_{i,j} \varphi(x_j^n(t))dt + \sigma(x_i^n(t))dw_i(t) \quad (2.46)$$

$i = 1, \dots, n$

where $w_1, \dots, w_n$ are n independent 1-dimensional Wiener processes; $\{J_{i,j}\}$ is a set of i.i.d. R -valued random variables with zero mean and finite variance which is independent of the Wiener processes; $b(x)$, $\varphi(x)$ and $\sigma(x)$ are measurable functions defined on R . In the limit of $n \rightarrow \infty$ one expects that the system (2.46) becomes independent with its marginal behavior described by a single stochastic differential equation of the following form

$$dx(t) = b(x(t))dt + G(t)dt + \sigma(x(t))dw(t) \quad (2.47)$$

where w is the 1-dimensional Brownian motion; $G(t)$ is a Gaussian process with zero mean and variance

$$E[G(t)G(s)] = E[\varphi(x(t))\varphi(x(s))]. \quad (2.48)$$

In general it is still an open problem to give a mathematically satisfactory proof of the above conjecture.

S.Geman in [22] considered the following system of differential equations

$$\frac{dx_{n,i}(t)}{dt} = \alpha x_{n,i}(t) + \frac{1}{\sqrt{n}} \sum_{j=1}^n J_{i,j} x_{n,j}(t) \quad (2.49)$$

$$x_{n,i}(0) = 1$$

$$i = 1, \dots, n$$

where α is a constant; $\{J_{i,j}\}$ are i.i.d. random variables with mean zero and variance σ^2 . He proved the following theorem

Theorem 2.1 *Suppose there exists a constant β such that*

$$E(|J_{1,1}|^k) \leq k^{\beta k}, \quad \forall k \geq 2. \quad (2.50)$$

Then for each fixed $m \geq 1$ and each fixed $T > 0$, $(x_{n,1}(\cdot), \dots, x_{n,m}(\cdot))$ converges weakly, as $n \rightarrow \infty$, on $C([0, T], R^m)$ to i.i.d. Gaussian processes $(x_{\infty,1}(\cdot), \dots, x_{\infty,m}(\cdot))$

$$(x_{n,1}(\cdot), \dots, x_{n,m}(\cdot)) \Rightarrow (x_{\infty,1}(\cdot), \dots, x_{\infty,m}(\cdot)), \quad n \rightarrow \infty$$

where the limit Gaussian processes have mean

$$\mu(t) = e^{\alpha t} \quad (2.51)$$

and covariance function

$$C(t, s) = e^{\alpha(t+s)} \sum_{k=1}^{\infty} \frac{t^k s^k}{(k!)^2} \quad (2.52)$$

Although the equation (2.49) represents only a very special case of equation (2.46), its proof is highly nontrivial.

Another related work on these random system is the continuous van Hemmen's spin glass model studied by Bonilla et al [3]. The system can be described by the

following stochastic differential equations:

$$dx_i(t) = \mu(x_i(t) - x_i^j(t))dt - \frac{1}{N} \sum_{j=1}^n (J_0 + J\xi_i\eta_j + J\xi_j\eta_i)(x_i(t) - x_j(t))dt + \sigma dw_i(t) \quad (2.53)$$

$$i = 1, \dots, n$$

where $w_1, \dots, w_n$ are n independent 1-dimensional Wiener processes; ξ_i and η_i are R -valued i.i.d. random variables; μ , J_0 , J and σ are constants. Let $P_n(t, x; \{\xi_i, \eta_i\})$ be the n -particle probability density associated with the system (2.53). Suppose that at time $t = 0$ the n -particle probability density $P_n(t, x; \{\xi_i, \eta_i\})$ is a product of one-particle densities. Then it was proposed (but did not prove) in [3] that in the limit as $n \rightarrow \infty$ the system (2.53) becomes independent with the one-particle probability density given by the solution of the following non-linear Fokker-Plank equation:

$$\frac{\partial P}{\partial t} = \frac{1}{2} \frac{\partial^2 P}{\partial x^2} - \frac{\partial}{\partial x} \left\{ \left[\left(A - \frac{\theta^2}{2} x^2 \right) x + \bar{x} + \alpha(\xi \bar{x} \eta + \eta \bar{x} \xi) \right] P \right\} \quad (2.54)$$

$$1 = \int_{-\infty}^{\infty} P(t; x, \xi, \eta) dx$$

$$\bar{x} = \int x P(t; x, \xi, \eta) dx d\rho(\xi) d\rho(\eta)$$

$$\bar{x}\xi = \int x\xi P(t; x, \xi, \eta) dx d\rho(\xi) d\rho(\eta)$$

$$\bar{x}\eta = \int x\eta P(t; x, \xi, \eta) dx d\rho(\xi) d\rho(\eta)$$

where x, t and P have been transformed so that

$$J_0 t \rightarrow t$$

$$\left(\frac{J_0}{T} \right)^{\frac{1}{2}} x \rightarrow x$$

$$\left(\frac{J_0}{T} \right)^{\frac{1}{2}} P(t; x, \xi, \eta) \rightarrow P(t; x, \xi, \eta)$$

$$A = \frac{\mu}{J_0} - 1$$

$$\theta = \frac{(2T\mu)^{\frac{1}{2}}}{J_0}$$

$$\alpha = \frac{J}{J_0}$$

Starting from this nonlinear Fokker-Plank equation, Bonilla et al studied some interesting phenomena on this model, such as phase transition, bifurcation and metastability.

The spin glass problem also raised some new topics in large deviation theory. As was exposed in the introduction of this thesis one needs to consider the large deviation problem associated with a family of random probability measures. One related work in this direction was done by Comets [4] in which he formulated and proved a large deviation theorem for the conditional distribution of an i.i.d. random field. Based on that result he then studied the asymptotic behavior of the following Gibbs measures

$$Q_n^y(dx) = Z_n^y \exp\{-nv(< M, r_{n,\omega} >)\} \otimes_1^\infty \mu(dx) \quad (2.55)$$

where $r_{n,\omega}$ is the empirical measure defined by

$$r_{n,\omega} = \frac{1}{n} \sum_{i=1}^n \delta_{(x_i, y_i)} \in \mathbf{M}_1(X \times Y) \quad (2.56)$$

where X and Y are Polish spaces; $x_i \in X$ denotes the i th particle and $y_i \in Y$ represents the randomness of the interaction in the particle. $M : X \times Y \rightarrow l^2$ is a bounded continuous function, here l^2 denotes the Hilbert space of square summable real sequences with scalar product $t \cdot s$ for t, s in l^2 ; $v : l^2 \rightarrow \mathbf{R}$ is a twice continuously differentiable function with bounded derivatives on bounded sets in l^2 ; μ is a probability measure on X ; and Z_n^y is the usual normalization factor.

For $t \in l^2$, define

$$L(t) = \log \int \exp\{t \cdot M(x_1, y_1)\} \mu(dx_1) \nu(dy_1). \quad (2.57)$$

where ν is the distribution of the random variables y_i , and let L^* be the Legendre transformation of L

$$L^*(m) = \sup_{t \in l^2} \{t \cdot m - L(t)\}, \quad m \in l^2 \quad (2.58)$$

Then it was proved in [4] that

Theorem 2.2 *For almost every y ,*

$$1. \lim_{n \rightarrow \infty} \frac{1}{n} Z_n^y = - \inf_{m \in l^2} \{v(m) + L^*(m)\}$$

2. *The law of the empirical order parameter*

$$M_n(\omega) = \int M dr_{n,\omega}$$

under $(Q_n^y)_{n=1}^\infty$ obeys a large deviation principle with rate function

$$v + L^* + f$$

where $f = - \inf_{m \in l^2} \{v(m) + L^(m)\}$*

3. *The sequence $(Q_n^y)_{n=1}^\infty$ is tight on $\mathbf{M}_1(X)$, and any limit point is a mixture*

$$\int_S \bigotimes_1^\infty r_m^{y_i} \tau(dm)$$

for some $\tau \in \mathbf{M}_1(S)$ which depends on y , where

$$S = \{m \in l^2 : v(m) + L^*(m) + f = 0\}$$

and

$$r_m^{y_i}(dx_i) = \exp\{t \cdot M(x_i, y_i) - \log \int \exp\{t \cdot M(x_1, y_1)\} \mu(dx_1)\} \mu(dx_i)$$

and $t = t(m)$ satisfies $\frac{d}{dt} L(t) = m$.

We end this chapter by presenting two other interesting spin glass models that lead to some novel large deviation computation. These two models are called respectively Random Energy Model (REM) and Generalized Random Energy Model (GREM) which were introduced by Derrida [8][9]. Before introducing these two models let us first look at the SK spin glass model from a different point of view. In the SK spin glass model, the Hamiltonian has the form

$$H(S) = - \sum_{1 \leq i < j \leq n} \frac{J_{i,j}}{\sqrt{n}} S_i S_j, \quad (2.59)$$

where $S = (S_1, \dots, S_n)$, $S_i \in \{+1, -1\}$, and $\{J_{i,j}\}$ are independent Gaussian random variables with zero mean and variance 1.

One may view $H(S)$ as a Gaussian random variable with mean zero

$$E[H(S)] = 0 \quad (2.60)$$

and variance

$$\begin{aligned} E[H(S)H(\hat{S})] &= \frac{1}{n} \sum_{1 \leq i < j \leq n} \sum_{1 \leq k < l \leq n} S_i S_j \hat{S}_k \hat{S}_l E(J_{i,j} J_{k,l}) \\ &= \frac{1}{n} \sum_{1 \leq i < j \leq n} S_i S_j \hat{S}_i \hat{S}_j \\ &= \frac{1}{n} \left[\frac{1}{2} \left(\sum_{i=1}^n S_i \hat{S}_i \right)^2 - \frac{n}{2} \right] \end{aligned} \quad (2.61)$$

where we have used the fact that $S_i^2 = 1$.

If we let

$$q_n(S, \hat{S}) = \frac{1}{n} \sum_{i=1}^n S_i \hat{S}_i \quad (2.62)$$

Then

$$\begin{aligned} E[H(S)H(\hat{S})] &= \frac{1}{2n} [(nq_n(S, \hat{S}))^2 - n] \\ &= \frac{n}{2} [q_n^2(S, \hat{S}) - \frac{1}{n}] \end{aligned} \quad (2.63)$$

In particular

$$E(H(S)^2) = \frac{n-1}{2} \quad (2.64)$$

Let Z_n be the partition function defined by

$$Z_n = \sum_S \exp\{-\beta H(S)\} \quad (2.65)$$

where β is the inverse temperature. Then Z_n can be viewed as a summation of some random variables. Note that these random variables are not independent.

The REM and GREM are simplifications of the SK model. In the REM, the partition function is defined by

$$Z_n = \sum_{i=1}^{2^n} \exp\{\beta \sqrt{n} X_i\} \quad (2.66)$$

where $X_1, \dots, X_{2^n}$ are 2^n i.i.d. Gaussian random variables with zero mean and variance 1. In the GREM, the partition function is defined by

$$Z_{n,m} = \sum_{k_1=1}^{\alpha_1^n} \cdots \sum_{k_m=1}^{\alpha_m^n} \exp\{\beta \sqrt{n} \sum_{j=1}^m \sqrt{p_j} \epsilon_{k_1, \dots, k_j}^j\} \quad (2.67)$$

where n, m are integers; p_j, α_j are positive numbers such that

$$\sum_{i=1}^n a_i = 1, \quad \sum_{i=1}^n \log \alpha_i = \log 2; \quad (2.68)$$

and $\epsilon_{k_1, \dots, k_j}^j$ are $\alpha_1^n + \alpha_1^n \alpha_2^n + \alpha_1^n \alpha_2^n \alpha_3^n + \cdots + \alpha_1^n \cdots \alpha_m^n$ independent Gaussian random variables with zero mean and variance 1.

The problem is to find the free energies associated with these two models in the thermodynamic limit as $n \rightarrow \infty$, i.e., the asymptotic behavior of the following two quantities

$$\frac{1}{n} \log Z_n(\beta) = \frac{1}{n} \log \sum_{i=1}^{2^n} \exp\{\beta \sqrt{n} X_i\} \quad (2.69)$$

and

$$\frac{1}{n} \log Z_{n,m}(\beta) = \frac{1}{n} \log \sum_{k_1=1}^{\alpha_1^n} \cdots \sum_{k_m=1}^{\alpha_m^n} \exp\{\beta \sqrt{n} \sum_{j=1}^m \sqrt{p_j} \epsilon_{k_1, \dots, k_j}^j\} \quad (2.70)$$

The above problems were studied by Eisele [16] and Picco [37]. The basic results can be stated as follows:

Theorem 2.3 $\forall \beta > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log Z_n(\beta) = \begin{cases} \frac{\beta^2}{2} + \frac{\beta_c^2}{2} & \text{if } 0 \leq \beta \leq \beta_c \\ \beta \beta_c & \text{if } \beta > \beta_c. \end{cases} \quad (2.71)$$

almost surely and in $L^p(\Omega, \mathcal{F}, P)$ for any $1 \leq p < \infty$, where

$$\beta_c = \sqrt{2 \log 2} \quad (2.72)$$

Proof: See [37] and further references there.

Now, for any $k = 1, 2, \dots$ let $\xi(k)$ be the subset of R^k defined by

$$\xi(k) = \{(x_1, \dots, x_k) : \sum_{i=1}^j \frac{x_i^2}{2} \leq \sum_{i=1}^j \log \alpha_i, \forall j = 1, \dots, k\} \quad (2.73)$$

and let $\|x\|^2 = \sum_{i=1}^n x_i^2$ be the Euclidean norm of R^n .

Theorem 2.4 Let m^* be the n -dimensional vector $m^* = (m_i^*) = (\beta \sqrt{p_i})_{i=1}^n$ and let m be the point of the compact convex subset of $R^n : \xi(n)$ at minimal distance from m^* . Then for any $\beta > 0$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,m}(\beta) = \frac{\|m^*\|^2}{2} - \frac{\|m - m^*\|^2}{2} + \log 2 \quad (2.74)$$

almost surely and in $L^p(\Omega, \mathcal{F}, P)$ for any $1 \leq p < \infty$.

Proof: See [37].

Chapter 3

Stochastic Models

In this chapter we introduce the stochastic models which will be studied in the following chapters. In section 1 we establish the existence and uniqueness of a system of n randomly interacting diffusion processes. In section 2 we introduce the local empirical measures associated with such random system. In section 3 we consider a non-linear Markov diffusion process.

3.1 Interacting diffusion system

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_t)_{t \geq 0}$ be a family of sub σ -algebras of $\mathcal{F}$ which is assumed to satisfy the standard conditions (refer back to chapter 1). Let $w_1, \dots, w_n$ be n independent 1-dimensional Wiener processes defined on this probability space such that they are adapted to the family of sub σ -algebras $\mathcal{F}_t$. Let $\xi_1, \xi_2, \dots$ be a sequence of i.i.d. $\Theta \subset R^q$ valued random variables defined on some probability space with distribution λ . Throughout this chapter we assume that the set Θ consists of p distinct patterns in R^q :

$$\Theta = \{\theta_1, \dots, \theta_p\} \tag{3.1}$$

and the k th pattern occurs with the probability λ_k , i.e.

$$\lambda = \sum_{k=1}^p \lambda_k \delta_{\theta_k}. \tag{3.2}$$

where $\lambda_1, \dots, \lambda_p \in (0, 1)$, $\sum_{k=1}^p \lambda_k = 1$ and δ_x denotes the Dirac measure at x . Let $\mathcal{B}(\Theta)$ be the collection of all subsets of Θ . Then $(\Theta, \mathcal{B}(\Theta), \lambda)$ forms a probability space. Let $(\Xi, \mathcal{B}(\Xi), \varphi)$ be the infinite product probability space of $(\Theta, \mathcal{B}(\Theta), \lambda)$, i.e.

$$\begin{aligned}\Xi &= \Theta \times \Theta \times \dots \\ \mathcal{B}(\Xi) &= \mathcal{B}(\Theta) \times \mathcal{B}(\Theta) \times \dots \\ \varphi &= \lambda \otimes \lambda \otimes \dots\end{aligned}\tag{3.3}$$

The stochastic model we will study is given by the following Itô equations in R :

$$\begin{aligned}dx_i^{(n)}(t) &= b(x_i^{(n)}(t))dt + \frac{1}{n} \sum_{j=1}^n c(\xi_i, \xi_j) \Phi(x_i^{(n)}(t), x_j^{(n)}(t))dt + \sigma(x_i^{(n)}(t))dw_i(t), \\ i &= 1, \dots, n.\end{aligned}\tag{3.4}$$

where $b : R \rightarrow R$, $\sigma : R \rightarrow R$, and $\Phi : R \times R$ are measurable functions.

For each given realization $\xi = (\xi_1, \xi_2, \dots) \in \Xi$, the random evolution of the system (3.4) defines a Markov diffusion process in R^n with associated generator

$$\begin{aligned}L_\xi^{(n)}f(x) &= \sum_{i=1}^n \left\{ \frac{1}{2} \sigma^2(x_i) \frac{\partial^2}{\partial x_i^2} f(x) + b(x_i) \frac{\partial}{\partial x_i} f(x) \right\} \\ &+ \frac{1}{n} \sum_{i,j=1}^n c(\xi_i, \xi_j) \Phi(x_i, x_j) \frac{\partial}{\partial x_i} f(x)\end{aligned}\tag{3.5}$$

where $x = (x_1, \dots, x_n) \in R^n$ and $f \in C_K^2(R^n)$. To ensure that the above system exists and is unique, we make the following hypotheses on the functions $b(x), \sigma(x)$ and $\Phi(x, y)$:

H_1 . There exists a positive constant K such that for all $x, y \in R$

$$|b(x) - b(y)|^2 + |\sigma(x) - \sigma(y)|^2 \leq K|x - y|^2.$$

$$xb(x) + |\sigma(x)|^2 \leq K(1 + |x|^2).$$

H_2 . There exists a positive constant K such that for all $x_1, y_1, x_2, y_2, x, y \in R$

$$|\Phi(x_1, y_1) - \Phi(x_2, y_2)| \leq K(|x_1 - x_2| + |y_1 - y_2|).$$

$$|\Phi(x, y)| \leq K(|x| + |y|).$$

H_3 $x^{(n)}(0) = (x_i^{(n)}(0))_{i=1}^n \in L^2(\Omega, \mathcal{F}_0, P)$ is independent of $w_1, \dots, w_n$.

Theorem 3.1 Suppose that the hypotheses H_1 to H_3 are satisfied. Then

1. for each $\xi \in \Xi$, equations (3.4) defines a unique system of Markov diffusion processes $x_\xi^{(n)} = \{(x_{\xi,1}^{(n)}(t), \dots, x_{\xi,n}^{(n)}(t)); t \geq 0\}$ on R^n .
2. the domain $D(A_\xi^{(n)})$ of the infinitesimal generator $A_\xi^{(n)}$ contains $C_K^2(R^n)$ and for $f \in C_K^2(R^n)$, $A_\xi^{(n)}f(x) = L_\xi^{(n)}f(x)$.
3. if $\sup_{n \geq 1} E(|x_{\xi,i}^{(n)}(0)|^2) < \infty$, φ -a.s. for each $i \geq 1$, then for any $T > 0$, $\sup_{n \geq 1} E(\sup_{t \leq T} |x_\xi^{(n)}(t)|^2) < \infty$, φ -a.s.

Proof: 1. Let $\alpha : R^n \times (R^q)^\infty \rightarrow R^n$ and $\beta : R^n \rightarrow R^n \otimes R^n$ be the continuous functions defined by

$$\alpha(x, \xi) = \begin{pmatrix} b(x_1) + \frac{1}{n} \sum_{j=1}^n c(\xi_1, \xi_j) \Phi(x_1, x_j) \\ \vdots \\ b(x_n) + \frac{1}{n} \sum_{j=1}^n c(\xi_n, \xi_j) \Phi(x_n, x_j) \end{pmatrix} \quad (3.6)$$

$$\beta(x) = \begin{pmatrix} \sigma(x_1) & & 0 \\ & \ddots & \\ 0 & & \sigma(x_n) \end{pmatrix} \quad (3.7)$$

Then equation (3.4) can be written as an Itô equation in R^n

$$dx(t) = \alpha(x(t), \xi)dt + \beta(x(t))dw(t) \quad (3.8)$$

where $x(t) = (x_1(t), \dots, x_n(t))^t$ and $w(t) = (w_1(t), \dots, w_n(t))^t$. Here A^t denotes the transpose of a vector A .

Let $c = \max_{1 \leq k, l \leq p} c(\theta_k, \theta_l)$. For $x \in R^n$, by the hypotheses above

$$\begin{aligned}
 & x \cdot \alpha(x, \xi) + |\beta(x)|^2 \\
 &= \sum_{i=1}^n (x_i b(x_i) + |\sigma(x_i)|^2) + \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n c(\xi_i, \xi_j) x_i \Phi(x_i, x_j) \\
 &\leq \sum_{i=1}^n K(1 + |x_i|^2) + \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n c |x_i| K(|x_i| + |x_j|) \\
 &\leq Kn + K \sum_{i=1}^n |x_i|^2 + cK \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (|x_i|^2 + |x_i| |x_j|) \\
 &\leq c_1 + c_2 \sum_{i=1}^n |x_i|^2 \\
 &\leq c_3(1 + |x|^2).
 \end{aligned} \tag{3.9}$$

where c_1 , c_2 , and c_3 are some positive constants.

Similarly, for $x, y \in R^n$, by the hypotheses above

$$\begin{aligned}
 & |\alpha(x, \xi) - \alpha(y, \xi)|^2 + |\beta(x) - \beta(y)|^2 \\
 &= \sum_{i=1}^n |(b(x_i) - b(y_i)) + \frac{1}{n} \sum_{j=1}^n c(\xi_i, \xi_j) (\Phi(x_i, x_j) - \Phi(y_i, y_j))|^2 \\
 &+ \sum_{i=1}^n |\sigma(x_i) - \sigma(y_i)|^2 \\
 &\leq 2 \sum_{i=1}^n [|b(x_i) - b(y_i)|^2 + |\sigma(x_i) - \sigma(y_i)|^2] \\
 &+ \frac{2}{n} \sum_{i=1}^n \sum_{j=1}^n |c(\xi_i, \xi_j)|^2 |\Phi(x_i, x_j) - \Phi(y_i, y_j)|^2 \\
 &\leq 2 \sum_{i=1}^n K |x_i - y_i|^2 + \frac{2}{n} \sum_{i=1}^n \sum_{j=1}^n c^2 2K^2 (|x_i - y_i|^2 + |x_j - y_j|^2) \\
 &\leq c_4 |x - y|^2
 \end{aligned} \tag{3.10}$$

for some constant c_4 . This shows that $\alpha(x, \xi)$ and $\beta(x)$ satisfy the conditions 1 and 2 of theorem 1.1 and therefore the equation (3.4) has a unique solution.

2. We now show that the domain of the infinitesimal generator of the solution process contains $C_K^2(R^n)$ and coincides with $L_\xi^{(n)}$ on $C_K^2(R^n)$. For each $\xi \in \Xi$, let $\{x_{\xi,x}^{(n)}(t), t \geq 0\}$ be the unique diffusion process determined by equation (3.4) with $x_{\xi,x}^{(n)}(0) = x$ and let $T_\xi^{(n)}(t)$ be the corresponding semigroup defined by

$$T_\xi^{(n)}(t)f(x) = \int_{R^n} f(y)P_\xi^{(n)}(t, x, dy). \quad (3.11)$$

where $P_\xi^{(n)}(t, x, A)$ is the transition function of the Markov diffusion process $x_{\xi,x}^{(n)}(t)$

$$P_\xi^{(n)}(t, x, A) = P(x_{\xi,x}^{(n)}(t) \in A), \quad A \in \mathcal{B}(R^n). \quad (3.12)$$

For $f \in C_K^2(R^n)$, by Itô's formula

$$f(x_{\xi,x}^{(n)}(t)) - f(x) - \int_0^t L_\xi^{(n)}f(x_{\xi,x}^{(n)}(s))ds \quad (3.13)$$

is a $(\mathcal{F}_t, P)$ -martingale. Denote $g = L_\xi^{(n)}f$. Then the martingale property implies

$$T_\xi^{(n)}(t)f(x) - f(x) - \int_0^t T_\xi^{(n)}(s)g(x)ds = 0. \quad (3.14)$$

It follows that

$$\begin{aligned} \lim_{t \rightarrow 0} \left\| \frac{1}{t}(T_\xi^{(n)}(t)f - f) - g \right\| &= \lim_{t \rightarrow 0} \left\| \frac{1}{t} \int_0^t (T_\xi^{(n)}(s)g - g)ds \right\| \\ &\leq \lim_{t \rightarrow 0} \frac{1}{t} \int_0^t \|T_\xi^{(n)}(s)g - g\| ds \\ &= 0. \end{aligned} \quad (3.15)$$

Hence $f \in D(A_\xi^{(n)})$ and $A_\xi^{(n)}f = L_\xi^{(n)}f$ for $f \in C_K^2(R^n)$.

3. For each $i = 1, \dots, n$

$$\begin{aligned} E(\sup_{t \leq T} |x_{\xi,i}^{(n)}(t)|^2) &\leq 4E(|x_{\xi,i}^{(n)}(0)|^2) + 4T \int_0^T E(\sup_{s \leq t} |b(x_{\xi,i}^{(n)}(s))|^2)dt \\ &\quad + 4\frac{c^2}{n} \sum_{j=1}^n T \int_0^T E(\sup_{s \leq t} |\Phi(x_{\xi,i}^{(n)}(s), x_{\xi,j}^{(n)}(s))|^2)dt \\ &\quad + 4E(\sup_{t \leq T} |M_{\xi,i}^{(n)}(t)|^2) \end{aligned} \quad (3.16)$$

where

$$M_{\xi,i}^{(n)}(t) = \int_0^t \sigma(x_{\xi,i}^{(n)}(s)) dw_i(s). \quad (3.17)$$

According to the basic theory on stochastic integral we know that $M_{\xi,i}^{(n)}(t)$ is at least a local martingale, i.e., there exists a sequence of stopping times $\tau_1 \leq \tau_2 \leq \dots$ with $\tau_m \rightarrow \infty$ such that for each m the process $M_{\xi,i}^{(n)}(t \wedge \tau_m)$ is a martingale. Thus Doob's martingale inequality gives

$$\begin{aligned} E(\sup_{t \leq T} |M_{\xi,i}^{(n)}(t \wedge \tau_m)|^2) &\leq 4E(|M_{\xi,i}^{(n)}(T \wedge \tau_m)|^2) \\ &= 4E(|\int_0^{T \wedge \tau_m} \sigma(x_{\xi,i}^{(n)}(s)) dw_i(s)|^2) \\ &\leq 4T \int_0^T E(|\sigma(x_{\xi,i}^{(n)}(t \wedge \tau_m))|^2) dt \\ &\leq 4TK \int_0^T E(1 + |x_{\xi,i}^{(n)}(t \wedge \tau_m)|^2) dt \end{aligned} \quad (3.18)$$

It follows that

$$\begin{aligned} E(\sup_{t \leq T} |x_{\xi,i}^{(n)}(t \wedge \tau_m)|^2) &\leq 4E(|x_{\xi,i}^{(n)}(0)|^2) + 4TK \int_0^T E[(1 + \sup_{s \leq t} |x_{\xi,i}^{(n)}(s \wedge \tau_m)|^2)] dt \\ &\quad + 8 \frac{c^2}{n} \sum_{j=1}^n TK^2 \int_0^T E[\sup_{s \leq t} (|x_{\xi,i}^{(n)}(s \wedge \tau_m)|^2 + |x_{\xi,j}^{(n)}(s \wedge \tau_m)|^2)] dt \\ &\quad + 4TK \int_0^T E(1 + |x_{\xi,i}^{(n)}(t \wedge \tau_m)|^2) dt \\ &\leq c_1 + c_2 E(|x_{\xi,i}^{(n)}(0)|^2) + c_3 \int_0^T E(\sup_{s \leq t} |x_{\xi,i}^{(n)}(s \wedge \tau_m)|^2) dt \end{aligned} \quad (3.19)$$

for some positive constants c_1, c_2 and c_3 , where in deriving the last inequality above we used the fact that $x_{\xi,i}^{(n)}$ has the same distribution for different i . The Gronwall's inequality then yields

$$E(\sup_{t \leq T} |x_{\xi,i}^{(n)}(t \wedge \tau_m)|^2) \leq [c_1 + c_2 E(|x_{\xi,i}^{(n)}(0)|^2)] \exp(c_3 T) \quad (3.20)$$

Letting $m \rightarrow \infty$ and using the fact that the constants c_1, c_2 and c_3 do not depend on n we have the final result. $\square$

3.2 Local empirical measures

In order to study the asymptotic behavior of the system (3.4) in the limit as n tends to infinity it is natural to consider the empirical measures associated to each finite system

$$X_\xi^{(n)} = \frac{1}{n} \sum_{i=1}^n \delta_{x_{\xi,i}^{(n)}} \quad (3.21)$$

where $x_{\xi,1}^{(n)}, \dots, x_{\xi,n}^{(n)}$ is a configuration of the unique solution of (3.4). There are two different approaches to study this random object $X_\xi^{(n)}$. In the first approach $X_\xi^{(n)} = \{X_\xi^{(n)}(t) : t \geq 0\}$ is considered as a continuous measure-valued process, i.e., $X_\xi^{(n)} \in C([0, \infty), \mathbf{M}_1(R))$. In the second approach $X_\xi^{(n)}$ is viewed as a random probability measure on the path space $C([0, \infty), R)$, i.e., $X_\xi^{(n)} \in \mathbf{M}_1(C([0, \infty), R))$. In this thesis $X_\xi^{(n)}$ is always viewed as a random probability measure on the path space $C([0, \infty), R)$.

For each $\xi \in \Xi$, let $X_\xi^{(n)}(t)$ be the marginal probability measure in $\mathbf{M}_1(R)$ of $X_\xi^{(n)}$ at time t , which is defined as following. Let Z_t , $t \geq 0$, be the projective map on $C([0, \infty), R)$ defined by

$$\begin{aligned} Z_t : C([0, \infty), R) &\rightarrow R \\ Z_t(x(\cdot)) &= x(t) \end{aligned} \quad (3.22)$$

Then the marginal probability measure $X_\xi^{(n)}(t)$ corresponding to $X_\xi^{(n)}$ at time t is the probability measure in $\mathbf{M}_1(R)$ defined by

$$X_\xi^{(n)}(t) = X_\xi^{(n)} \circ Z_t^{-1} \quad (3.23)$$

According to Itô's lemma, for $f \in C_K^2(R)$, we have

$$\begin{aligned} d \langle X_\xi^{(n)}(t), f \rangle &= \frac{1}{n} \sum_{i=1}^n \left\{ \frac{1}{2} \sigma^2(x_{\xi,i}^{(n)}(t)) f''(x_{\xi,i}^{(n)}(t)) + b(x_{\xi,i}^{(n)}(t)) f'(x_{\xi,i}^{(n)}(t)) \right\} dt \\ &+ \frac{1}{n^2} \sum_{i,j=1}^n c(\xi_i, \xi_j) \Phi(x_{\xi,i}^{(n)}(t), x_{\xi,j}^{(n)}(t)) f'(x_{\xi,i}^{(n)}(t)) dt \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{n} \sum_{i=1}^n \sigma(x_{\xi,i}^{(n)}(t)) f'(x_{\xi,i}^{(n)}(t)) dw_i(t) \\
& = \langle X_{\xi}^{(n)}(t), Lf \rangle dt \\
& + \frac{1}{n^2} \sum_{i,j=1}^n c(\xi_i, \xi_j) \Phi(x_{\xi,i}^{(n)}(t), x_{\xi,j}^{(n)}(t)) f'(x_{\xi,i}^{(n)}(t)) dt \\
& + \frac{1}{\sqrt{n}} dM_{n,\xi}^f(t)
\end{aligned} \tag{3.24}$$

where L is the usual diffusion operator

$$Lf(x) = \frac{1}{2} \sigma^2(x) f''(x) + b(x) f'(x) \tag{3.25}$$

and $M_{n,\xi}^f(t)$ is a martingale term with quadratic variational process given by

$$\frac{d}{dt} \ll M_n^f \gg_t = \langle X_{\xi}^{(n)}(t), (f')^2 \sigma^2 \rangle \tag{3.26}$$

Unlike the usual mean field models of weakly interacting diffusions, the second term in the right-hand side of (3.24) cannot be expressed as an integral with respect to the empirical measure $X_{\xi}^{(n)}(t)$, and therefore it is not clear from (3.24) what the limit process will look like as $n \rightarrow \infty$. Motivated by the physical notion of order parameters for this particle spin glass model, we introduce the p local empirical measures $X_{\xi,\theta_1}^{(n)}, \dots, X_{\xi,\theta_p}^{(n)}$ as following

$$X_{\xi,\theta_k}^{(n)} = \frac{1}{n_{\xi,\theta_k}} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \delta_{x_{\xi,i}^{(n)}}, \quad k = 1, \dots, p \tag{3.27}$$

where n_{ξ,θ_k} counts the total numbers of $i \in \{1, \dots, n\}$ such that ξ_i takes the value θ_k , i.e.,

$$n_{\xi,\theta_k} = \sum_{i=1}^n 1_{\theta_k}(\xi_i). \tag{3.28}$$

Thus $X_{\xi,\theta_k}^{(n)}$ is the local empirical measure of the configuration $x_{\xi,1}^{(n)}, \dots, x_{\xi,n}^{(n)}$ associated with the k th pattern θ_k . Note that $X_{\xi,\theta_1}^{(n)}, \dots, X_{\xi,\theta_p}^{(n)}$ are random probability measures on $C([0, \infty), R)$ which depend on the realization of ξ and are related to the original

empirical measure $X_\xi^{(n)}$ through the following convex combination

$$X_\xi^{(n)} = \sum_{k=1}^p \frac{n_{\xi, \theta_k}}{n} X_{\xi, \theta_k}^{(n)}. \quad (3.29)$$

We now show, heuristically, that the local empirical measures $X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)}$ can be expressed in a form so that it is possible to obtain the limit process as $n \rightarrow \infty$.

Denote

$$Y_\xi^{(n)} = (X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)}) \quad (3.30)$$

Then, for each $\xi \in \Xi$, $Y_\xi^{(n)}$ can be viewed as a random measure on $(\Omega, \mathcal{F}, P)$ with values in $\mathbf{M}_1(C([0, \infty), R))^{\odot p} = \mathbf{M}_1(C([0, \infty), R)) \times \dots \times \mathbf{M}_1(C([0, \infty), R))$, the p -fold product space of $\mathbf{M}_1(C([0, \infty), R))$. Let $Y_\xi^{(n)}(t) = (X_{\xi, \theta_1}^{(n)}(t), \dots, X_{\xi, \theta_p}^{(n)}(t))$, $t \geq 0$, be the corresponding marginal probability measure of $Y_\xi^{(n)}$ at time t . Then for each $k = 1, \dots, p$ and each $f \in C_K^2(R)$, by Itô's formula

$$\begin{aligned} d \langle X_{\xi, \theta_k}^{(n)}(t), f \rangle &= \frac{1}{n_{\xi, \theta_k}} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \left\{ \frac{1}{2} \sigma^2(x_{\xi, i}^{(n)}(t)) f''(x_{\xi, i}^{(n)}(t)) + b(x_{\xi, i}^{(n)}(t)) f'(x_{\xi, i}^{(n)}(t)) \right\} dt \\ &+ \sum_{l=1}^p \frac{n_{\xi, \theta_l}}{n} c(\theta_k, \theta_l) \frac{1}{n_{\xi, \theta_k}} \sum_{i, \xi_i = \theta_k} \frac{1}{n_{\xi, \theta_l}} \sum_{j, \xi_j = \theta_l} \Phi(x_{\xi, i}^{(n)}(t), x_{\xi, j}^{(n)}(t)) f'(x_{\xi, i}^{(n)}(t)) dt \\ &+ \frac{1}{n_{\xi, \theta_k}} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \sigma(x_{\xi, i}^{(n)}(t)) f'(x_{\xi, i}^{(n)}(t)) dw_i(t) \\ &= \langle X_{\xi, k}^{(n)}(t), Lf \rangle dt \\ &+ \sum_{l=1}^p \frac{n_{\xi, \theta_l}}{n} c(\theta_k, \theta_l) \int_R f'(x) \left[\int_R \Phi(x, y) X_{\xi, \theta_l}^{(n)}(t, dy) \right] X_{\xi, \theta_k}^{(n)}(t, dx) dt \\ &+ \frac{1}{\sqrt{n_{\xi, \theta_k}}} dM_{\xi, \theta_k}^f(t) \\ &= \langle X_{\xi, k}^{(n)}(t), L_{\xi, \theta_k}^{(n)}(Y_\xi^{(n)}(t)) f \rangle dt + \frac{1}{\sqrt{n_{\xi, \theta_k}}} dM_{\xi, \theta_k}^f(t) \end{aligned} \quad (3.31)$$

where the operator $L_{\xi, \theta_k}^{(n)}(Y_{\xi}^{(n)}(t))$ is defined by

$$L_{\xi, \theta_k}^{(n)}(Y_{\xi}^{(n)}(t))f(x) = Lf(x) + \sum_{l=1}^p \frac{n_{\xi, \theta_l}}{n} c(\theta_k, \theta_l) \int_R \Phi(x, y) X_{\xi, \theta_l}^{(n)}(t, dy) f'(x). \quad (3.32)$$

and $M_{\xi, \theta_k}^f(t)$ is a martingale term with quadratic variational process satisfying

$$\frac{d}{dt} \ll M_{\xi, \theta_k}^f \gg_t = \langle X_{\xi, \theta_k}^{(n)}(t), (f')^2 \sigma^2 \rangle \quad (3.33)$$

Since by the strong law of large number

$$\frac{n_{\xi, \theta_k}}{n} = \frac{1}{n} \sum_{i=1}^n 1_{\theta_k}(\xi_i) \rightarrow \lambda_k \quad \text{w-a.s. as } n \rightarrow \infty. \quad (3.34)$$

one would expect that for almost all ξ , the random vector

$$Y_{\xi}^{(n)} = (X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)})$$

will converge, as $n \rightarrow \infty$, in probability to a deterministic measure

$$m = (m_{\theta_1}, \dots, m_{\theta_p})$$

in $\mathbf{M}_1(C([0, \infty), R))^{\otimes p}$ such that the marginal measure $m(t)$ of m satisfies the following non-linear equations

$$\frac{d}{dt} \langle m_{\theta_k}(t), f \rangle = \langle m_{\theta_k}(t), L_{\theta_k}(m(t))f \rangle, \quad k = 1, \dots, p \quad (3.35)$$

where $f \in C_K^2(R)$ and the operators $L_{\theta_k}(m(t))$, $k = 1, \dots, p$, are defined by

$$L_{\theta_k}(m(t))f(x) = Lf(x) + \sum_{l=1}^p \lambda_l c(\theta_k, \theta_l) \int_R \Phi(x, y) m_{\theta_l}(t, dy) f'(x). \quad (3.36)$$

According to (3.29), the empirical measure $X_{\xi}^{(n)}$ is a convex combination of the local empirical measures $X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)}$ with the coefficients $\frac{n_{\xi, \theta_k}}{n}$. It is natural to think that in the limit, as $n \rightarrow \infty$, the empirical measure $X_{\xi}^{(n)}$ will converge weakly, for almost all ξ , to the convex combination of the corresponding limit of the local empirical measures, i.e.,

$$X_{\xi}^{(n)} \Rightarrow \sum_{k=1}^p \lambda_k m_{\theta_k} \quad \text{w-a.s., } n \rightarrow \infty \quad (3.37)$$

These heuristic arguments will be made precise in the next chapter. We end this section by proving the following fact.

Proposition 3.1 *Let $\xi \in \Xi$, and $f \in C_K^2(R)$. For each $i = 1, \dots, n$, suppose that*

$$E(|x_{\xi,i}^{(n)}(0)|^2) < \infty, \quad \varphi\text{-a.s.} \quad (3.38)$$

then for each $k = 1, \dots, p$, the process $\{M_{n,i,\xi,\theta_k}^{f,X_{\xi}^{(n)}}(t) : t \leq T\}$ defined by

$$M_{n,i,\xi,\theta_k}^{f,X_{\xi}^{(n)}}(t) = f(x_{\xi,i}^{(n)}(t)) - f(x_{\xi,i}^{(n)}(0)) - \int_0^t L_{\xi,\theta_k}^{(n)}(X_{\xi}^{(n)})f(x_{\xi,i}^{(n)}(s))ds \quad (3.39)$$

is a $(\mathcal{F}_t, P)$ -martingale with corresponding quadratic variational process given by

$$\ll M_{n,i,\xi,\theta_k}^{f,X_{\xi}^{(n)}} \gg_t = \int_0^t \sigma^2(x_{\xi,i}^{(n)}(s))(f'(x_{\xi,i}^{(n)}(s)))^2 ds \quad (3.40)$$

Proof: Applying Itô's formula to $f(x_{\xi,i}^{(n)}(t))$ yields

$$M_{n,i,\xi,\theta_k}^{f,X_{\xi}^{(n)}}(t) = \int_0^t \sigma(x_{\xi,i}^{(n)}(s))f'(x_{\xi,i}^{(n)}(s))dw_i(s) \quad (3.41)$$

By the assumption on the function σ and 3. of theorem 3.1

$$\begin{aligned} & E\left(\int_0^T |\sigma(x_{\xi,i}^{(n)}(t))f'(x_{\xi,i}^{(n)}(t))|^2 dt\right) \\ & \leq \int_0^T k_1 E(1 + |x_{\xi,i}^{(n)}(t)|^2) dt \\ & \leq c_1 T [1 + E(\sup_{t \leq T} |x_{\xi,i}^{(n)}(t)|^2)] \text{ for some constant } c_1 \\ & < \infty \end{aligned} \quad (3.42)$$

This shows $M_{n,i,\xi,\theta_k}^{f,X_{\xi}^{(n)}}$ is a martingale. The rest is obvious.

3.3 A non-linear Markov process

In this section we consider the following non-linear Markov process described by the following Itô equation in R

$$dx_{\xi}(t) = b(x_{\xi}(t))dt + \sum_{l=1}^p \lambda_l c(\xi, \theta_l) \int_R \Phi(x_{\xi}(t), x) m_{\theta_l}(t, dx) dt + \sigma(x_{\xi}(t))dw(t), \quad (3.43)$$

$$m_{\theta_l}(t, dx) = P\{x_{\theta_l}(t) \in dx\},$$

$$l = 1, \dots, p$$

where w is the 1-dimensional $\mathcal{F}_t$ -Wiener process defined on the probability space $(\Omega, \mathcal{F}, P)$ and ξ is a Θ -valued random variable defined on some probability space with the distribution λ .

Let F be the collection of continuous functions f on R which are Lipschitz continuous and satisfy the linear growth condition, i.e.,

$$F = \{f \in C(R) : \|f\|_F < \infty\} \quad (3.44)$$

where $\|\cdot\|_F$ is defined by

$$\|f\|_F = \sup_{x \in R} \frac{|f(x)|}{1 + |x|} + \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}. \quad (3.45)$$

Then $(F, \|\cdot\|_F)$ is a normed space and we denote by $(F^*, \|\cdot\|_{F^*})$ the corresponding dual space, where the dual norm $\|\cdot\|_{F^*}$ is defined by

$$\|\mu\|_{F^*} = \sup_{f \in F, \|f\|_F \leq 1} \{|\langle f, \mu \rangle|\}, \quad \mu, \nu \in F^*. \quad (3.46)$$

Let $\mathbf{M}_1^{(1)}(R)$ be the collection of all probability measures in $\mathbf{M}_1(R)$ which have finite first absolute moment

$$\mathbf{M}_1^{(1)}(R) = \{\mu \in \mathbf{M}_1(R) : \int_R |x| \mu(dx) < \infty\} \quad (3.47)$$

and be furnished with the corresponding weak topology, i.e.,

$$\mu_n \Rightarrow \mu \quad (\text{weak convergence}) \quad \text{in} \quad \mathbf{M}_1^{(1)}(R)$$

if and only if

$$\int_R f(x) \mu_n(dx) \rightarrow \int_R f(x) \mu(dx)$$

for any continuous function f on R such that

$$\sup_{x \in R} \frac{|f(x)|}{1 + |x|} < \infty. \quad (3.48)$$

It is known that $\mathbf{M}_1^{(1)}(R)$ is a Polish space and $\mathbf{M}_1^{(1)}(R) \subset (F^*, \|\cdot\|_{F^*})$ [29].

Next we introduce the Wasserstein metric $W_T(\mu, \nu)$ on $M_1(C([0, T], R))$ as following:

$$W_T(\mu, \nu) = \inf \left\{ \int_{C([0, T], R) \times C([0, T], R)} \sup_{t \leq T} |X_t(\omega_1) - X_t(\omega_2)| m(d\omega_1, d\omega_2) \right\} \quad (3.49)$$

$$\mu, \nu \in \mathbf{M}_1(C([0, T], R))$$

where X_t is the canonical map from $C([0, T], R)$ to R , and the infimum is taken over the space $\mathbf{M}_1(C([0, T], R) \times C([0, T], R))$ of all probability measures m on $C([0, T], R) \times C([0, T], R)$ which have marginal measures μ and ν . It is known that under this metric $\mathbf{M}_1(C([0, T], R))$ becomes a complete metric space which generates the usual topology of weak convergence plus the convergence of the first moment [10].

Lemma 3.1 For any $T > 0$ and $\mu, \nu \in \mathbf{M}_1(C([0, T], R))$

$$\sup_{t \leq T} \|\mu(t) - \nu(t)\|_{F^*} \leq W_T(\mu, \nu) \quad (3.50)$$

Proof: Let ξ and η be two $C([0, T], R)$ -valued random variables with distributions μ and ν , respectively. Then for any $t \leq T$ and $f \in F$

$$\begin{aligned} | \langle f, \mu(t) - \nu(t) \rangle | &= | E[f(\xi(t)) - f(\eta(t))] | \\ &\leq \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|} E(|\xi(t) - \eta(t)|) \end{aligned} \quad (3.51)$$

This yields

$$\begin{aligned} \|\mu(t) - \nu(t)\|_{F^*} &\leq E(|\xi(t) - \eta(t)|) \\ &\leq E(\sup_{t \leq T} |\xi(t) - \eta(t)|) \end{aligned} \quad (3.52)$$

Thus

$$\sup_{t \leq T} \|\mu(t) - \nu(t)\|_{F^*} \leq \inf_{\xi, \eta} \{E(\sup_{t \leq T} |\xi(t) - \eta(t)|)\} = W_T(\mu, \nu) \quad (3.53)$$

where the infimum is taken over all $C([0, T], R)$ -valued random variables ξ and η whose distributions are μ and ν , respectively. $\square$

Finally, we define the metric $\mathbf{W}_T(\mu, \nu)$ on $M_1(C([0, T], R))^{\otimes p}$ as following

$$\mathbf{W}_T(\mu, \nu) = \max_{1 \leq k \leq p} W_T(\mu_k, \nu_k) \quad (3.54)$$

$$\mu = (\mu_1, \dots, \mu_p), \nu = (\nu_1, \dots, \nu_p) \in M_1(C([0, T], R))^{\otimes p}$$

We are now ready to prove the main theorem of this section.

Theorem 3.1 *Suppose that the functions $b(x)$, $\sigma(x)$ and $\Phi(x, y)$ satisfy the hypotheses H_1 and H_2 of section 3.1. Furthermore, assume $x_\xi(0) \in L^2(\Omega, \mathcal{F}_0, P)$ is independent of w . Then for each given $\xi \in \Theta$, the non-linear stochastic differential equation (3.43) has a unique solution $x_\xi = \{x_\xi(t) : t \geq 0\}$.*

Proof: It suffices to show that for any $T > 0$ and $\xi \in \Theta$, the equation (3.43) has a unique solution $x_\xi = \{x_\xi(t) : t \leq T\}$. We first prove that for each given $m = (m_1, \dots, m_p) \in M_1(C([0, \infty), R))^{\otimes p}$ such that the marginal measure $m(t)$ of m belongs to $M_1^{(1)}(R)^{\otimes p}$ for all $t \geq 0$ the following ordinary stochastic differential equation

$$dx_\xi(t) = b(x_\xi(t))dt + \sum_{l=1}^p \lambda_l c(\xi, \theta_l) \int_R \Phi(x_\xi(t), x) m_l(t, dx) dt + \sigma(x_\xi(t)) dw(t) \quad (3.55)$$

has a unique solution $x_\xi^m = \{x_\xi^m(t) : t \geq 0\}$. To show this we note that, for given m and ξ , equation (3.55) has the form of

$$dx_\xi^m(t) = \alpha(x_\xi^m(t), m(t), \xi) dt + \sigma(x_\xi^m(t)) dw(t). \quad (3.56)$$

where $\alpha : R \times M_1^{(1)}(R)^{\otimes p} \times R^q \rightarrow R$ is the continuous function defined by

$$\alpha(x, m, \xi) = b(x) + \sum_{l=1}^p \lambda_l c(\xi, \theta_l) \int_R \Phi(x, z) m_l(dz). \quad (3.57)$$

So according to theorem 1.1, it is enough to check that the condition 1 and 2 in that theorem are satisfied. In the following we always denote $c = \max_{1 \leq k, l \leq p} c(\theta_k, \theta_l)$.

For $x, y \in R$, by the hypotheses, we have

$$\begin{aligned} & |\alpha(x, m, \xi) - \alpha(y, m, \xi)|^2 + |\sigma(x) - \sigma(y)|^2 \\ &= |b(x) - b(y) + \sum_{l=1}^p \lambda_l c(\xi, \theta_l) \int_R [\Phi(x, z) - \Phi(y, z)] m_l(dz)|^2 \\ &+ |\sigma(x) - \sigma(y)|^2 \\ &\leq 2(|b(x) - b(y)|^2 + |\sigma(x) - \sigma(y)|^2) \\ &+ 2pc^2 \sum_{l=1}^p \int_R |\Phi(x, z) - \Phi(y, z)|^2 m_l(dz) \\ &\leq 2K|x - y|^2 + 2pc^2 K^2 \sum_{l=1}^p |x - y|^2 \\ &\leq k_1 |x - y|^2 \end{aligned} \quad (3.58)$$

for some constant k_1 .

Similarly, for $x \in R$, by the hypotheses

$$\begin{aligned} & x \cdot \alpha(x, m, \xi) + |\sigma(x)|^2 \\ &= x \cdot [b(x) + \sum_{l=1}^p \lambda_l c(\xi, \theta_l) \int_R \Phi(x, z) m_l(dz)] + |\sigma(x)|^2 \\ &\leq xb(x) + |\sigma(x)|^2 + pc \sum_{l=1}^p |x| \int_R |\Phi(x, z)| m_l(dz) \\ &\leq K(1 + |x|^2) + pc \sum_{l=1}^p |x| \int_R K(|x| + |z|) m_l(dz) \\ &\leq K(1 + |x|^2) + p^2 c K |x|^2 + pc K \sum_{l=1}^p |x| \int_R |z| m_l(dz) \\ &\leq k_2 + k_3 |x|^2 \\ &\leq k_4 (1 + |x|^2). \end{aligned} \quad (3.59)$$

for some constants k_2, k_3 and k_4 . Hence the equation (3.55) has a unique solution and

$$E(\sup_{t \leq T} |x'_\xi(t)|^2) < \infty \quad (3.60)$$

for each $\xi \in \Theta$ and any $T > 0$.

Next, for each $k = 1, \dots, p$, we denote by $\mathcal{L}(x_{\theta_k}^m)$ the probability law of $x_{\theta_k}^m$ on $C([0, T], R)$, and put them into a vector

$$\mathcal{L}(x_\xi^m) = (\mathcal{L}(x_{\theta_1}^m), \dots, \mathcal{L}(x_{\theta_p}^m)). \quad (3.61)$$

Then $\mathcal{L}(x_\xi^m) \in \mathbf{M}_1(C([0, T], R))^{\otimes p}$. Furthermore, (3.60) implies that the marginal measure $\mathcal{L}(x_\xi^m)(t)$ of $\mathcal{L}(x_\xi^m)$ at time t belongs to $\mathbf{M}_1^{(1)}(R)^{\otimes p}$ for all $t \leq T$. Thus we can define a map Π on $\mathbf{M}_1(C([0, T], R))^{\otimes p}$ by

$$\Pi : m \mapsto \mathcal{L}(x_\xi^m) \quad (3.62)$$

which maps m to $\mathcal{L}(x_\xi^m)$. Suppose that $x_\xi = \{x_\xi(t) : t \geq 0\}$ is a solution of the non-linear equation (3.43). Then the vector $\mathcal{L}(x_\xi) = (\mathcal{L}(x_{\theta_1}), \dots, \mathcal{L}(x_{\theta_p}))$, where $\mathcal{L}(x_{\theta_k})$ denotes the law of x_{θ_k} restricted on $C([0, T], R)$, is a fixed point of Π . Conversely, if m is a fixed point of Π , then the equation (3.55) defines a solution of (3.43) up to time $t \leq T$. Therefore there is a one to one correspondence between the solutions of (3.43) and the fixed points of the operator Π . We now show that the operator Π has a unique fixed point. To this end, let $\mu = (\mu_1, \dots, \mu_p)$ and $\nu = (\nu_1, \dots, \nu_p)$ be arbitrary two elements in $\mathbf{M}_1(C([0, T], R))^{\otimes p}$ such that the marginal probability measures $\mu(t)$ and $\nu(t)$ belong to $\mathbf{M}_1^{(1)}(R)^{\otimes p}$ for all $t \leq T$ and let x_ξ^μ and x_ξ^ν be the corresponding solutions of the ordinary stochastic differential equation (3.55), i.e.,

$$x_\xi^\mu(t) = x_\xi(0) + \int_0^t \alpha(x_\xi^\mu(s), \mu, \xi) ds + \int_0^t \sigma(x_\xi^\mu(s)) dw(s), \quad (3.63)$$

$$x_\xi^\nu(t) = x_\xi(0) + \int_0^t \alpha(x_\xi^\nu(s), \nu, \xi) ds + \int_0^t \sigma(x_\xi^\nu(s)) dw(s), \quad (3.64)$$

we show that

$$\mathbf{W}_T(\Pi(\mu), \Pi(\nu)) < c\mathbf{W}_T(\mu, \nu)$$

for some constant c strictly less than 1.

For any $t \leq T$ we have

$$\begin{aligned} & |x_\xi^\mu(t) - x_\xi^\nu(t)| \\ & \leq \int_0^t |\alpha(x_\xi^\mu(s), \mu, \xi) - \alpha(x_\xi^\nu(s), \nu, \xi)| ds \\ & + \left| \int_0^t [\sigma(x_\xi^\mu(s)) - \sigma(x_\xi^\nu(s))] dw(s) \right| \\ & \leq \int_0^t |b(x_\xi^\mu(s)) - b(x_\xi^\nu(s))| ds + \sum_{l=1}^p \lambda_l |c(\xi, \theta_l)| \\ & \quad \int_0^t \left| \int_R [\Phi(x_\xi^\mu(s), z) \mu_l(s, dz) - \Phi(x_\xi^\nu(s), z) \nu_l(s, dz)] dz \right| ds \\ & + \left| \int_0^t [\sigma(x_\xi^\mu(s)) - \sigma(x_\xi^\nu(s))] dw(s) \right| \\ & \leq \int_0^t \sqrt{K} |x_\xi^\mu(s) - x_\xi^\nu(s)| ds \\ & + c_1 \sum_{l=1}^p \int_0^t \left| \int_R [\Phi(x_\xi^\mu(s), z) \mu_l(s, dz) - \Phi(x_\xi^\nu(s), z) \nu_l(s, dz)] dz \right| ds \\ & + \left| \int_0^t [\sigma(x_\xi^\mu(s)) - \sigma(x_\xi^\nu(s))] dw(s) \right| \\ & \leq \sqrt{K} T \sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)| + c_1 I_1(t) + I_2(t) \end{aligned} \quad (3.65)$$

where

$$I_1(t) = \sum_{l=1}^p \int_0^t \left| \int_R [\Phi(x_\xi^\mu(s), z) \mu_l(s, dz) - \Phi(x_\xi^\nu(s), z) \nu_l(s, dz)] dz \right| ds \quad (3.66)$$

and

$$I_2(t) = \left| \int_0^t [\sigma(x_\xi^\mu(s)) - \sigma(x_\xi^\nu(s))] dw(s) \right| \quad (3.67)$$

Since for fixed $x \in R$, $\Phi(x, \cdot)$ is an element of F , we have

$$\left| \langle \Phi(x_\xi^\mu(s), \cdot), \mu_l(s) - \nu_l(s) \rangle \right| \leq \| \Phi(x_\xi^\mu(s), \cdot) \|_F \| \mu_l(s) - \nu_l(s) \|_{F^*} \quad (3.68)$$

Using (3.68) and lemma 3.1

$$\begin{aligned}
I_1(t) &= \sum_{l=1}^p \int_0^t \left| \int_R [\Phi(x_\xi^\mu(s), z) \mu_l(s, dz) - \Phi(x_\xi^\nu(s), z) \nu_l(s, dz)] dz \right| ds \\
&\leq \sum_{l=1}^p \int_0^t \left| \int_R [\Phi(x_\xi^\mu(s), z) - \Phi(x_\xi^\nu(s), z)] \nu_l(s, dz) \right| ds \\
&\quad + \sum_{l=1}^p \int_0^t \left| \int_R \langle \Phi(x_\xi^\mu(s), \cdot), \mu_l(s) - \nu_l(s) \rangle dz \right| ds \\
&\leq Kp \int_0^t |x_\xi^\mu(s) - x_\xi^\nu(s)| ds + \sum_{l=1}^p \int_0^t \|\Phi(x_\xi^\mu(s), \cdot)\|_F \|\mu_l(s) - \nu_l(s)\|_F ds \\
&\leq KpT \sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)| + \sum_{l=1}^p T \sup_{t \leq T} \|\Phi(x_\xi^\mu(t), \cdot)\|_F \sup_{t \leq T} \|\mu_l(t) - \nu_l(t)\|_F \\
&\leq KpT \sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)| + T \sup_{t \leq T} \|\Phi(x_\xi^\mu(t), \cdot)\|_F \sum_{l=1}^p W_T(\mu_l, \nu_l) \\
&\leq KpT \sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)| + Tp \sup_{t \leq T} \|\Phi(x_\xi^\mu(t), \cdot)\|_F \mathbf{W}_T(\mu, \nu) \tag{3.69}
\end{aligned}$$

But by the hypotheses on $\Phi(x, y)$

$$\begin{aligned}
&\|\Phi(x_\xi^\mu(t), \cdot)\|_F \\
&= \sup_{x \in R} \frac{|\Phi(x_\xi^\mu(t), x)|}{1 + |x|} + \sup_{x \neq y} \frac{|\Phi(x_\xi^\mu(t), x) - \Phi(x_\xi^\mu(t), y)|}{|x - y|} \\
&\leq \sup_{x \in R} \frac{K(|x_\xi^\mu(t)| + |x|)}{1 + |x|} + K \\
&\leq \sup_{x \in R} \frac{K(|x_\xi^\mu(t)| + 1 + |x|)}{1 + |x|} + K \\
&\leq K|x_\xi^\mu(t)| + 2K \tag{3.70}
\end{aligned}$$

so

$$\begin{aligned}
E(\sup_{t \leq T} I_1(t)) &\leq KpTE(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|) + TpE(\sup_{t \leq T} \|\Phi(x_\xi^\mu(t), \cdot)\|_F) \mathbf{W}_T(\mu, \nu) \\
&\leq KpTE(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|) + Tp(K E(\sup_{t \leq T} |x_\xi^\mu(t)|) + 2K) \mathbf{W}_T(\mu, \nu) \\
&\leq c_2TE(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|) + c_3T \mathbf{W}_T(\mu, \nu) \tag{3.71}
\end{aligned}$$

because $E(\sup_{t \leq T} |x_\xi^\mu(t)|)$ is bounded by (3.60).

Also, by Burkholder-Davis-Gundy martingale inequality [38]

$$\begin{aligned}
 E(\sup_{t \leq T} I_2(t)) &= E(\sup_{t \leq T} |\int_0^t [\sigma(x_\xi^\mu(s)) - \sigma(x_\xi^\nu(s))] dw(s)|) \\
 &\leq c_1 E[(\int_0^T |\sigma(x_\xi^\mu(s)) - \sigma(x_\xi^\nu(s))|^2 ds)^{\frac{1}{2}}] \\
 &\leq c_4 E[(\int_0^T K |x_\xi^\mu(s) - x_\xi^\nu(s)|^2 ds)^{\frac{1}{2}}] \\
 &\leq c_5 \sqrt{T} E(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|)
 \end{aligned} \tag{3.72}$$

Combine (3.65), (3.71) and (3.72), we obtain

$$\begin{aligned}
 &E(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|) \\
 &\leq (k_1 T + k_2 \sqrt{T}) E(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|) + k_3 T \mathbf{W}_T(\mu, \nu) \\
 &\leq [k_1 T + k_2 \sqrt{T} + k_3 T] \left(E(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|) + \mathbf{W}_T(\mu, \nu) \right)
 \end{aligned} \tag{3.73}$$

We can choose small enough T such that $[k_1 T + k_2 \sqrt{T} + k_3 T] < \frac{1}{3}$. Then

$$E(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|) < \frac{1}{3} (E(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|) + \mathbf{W}_T(\mu, \nu)) \tag{3.74}$$

This implies that

$$E(\sup_{t \leq T} |x_\xi^\mu(t) - x_\xi^\nu(t)|) < \frac{1}{2} \mathbf{W}_T(\mu, \nu), \tag{3.75}$$

so, by the definition of $\mathbf{W}_T$,

$$\begin{aligned}
 \mathbf{W}_T(\Pi(\mu), \Pi(\nu)) &= \max_{1 \leq k \leq p} W_T(\mu_k, \nu_k) \\
 &\leq \max_{1 \leq k \leq p} E(\sup_{t \leq T} |x_{\theta_k}^\mu(t) - x_{\theta_k}^\nu(t)|) \\
 &< \frac{1}{2} \mathbf{W}_T(\mu, \nu)
 \end{aligned} \tag{3.76}$$

Hence, for small enough T , Π is a contraction mapping and therefore produces a unique fixed point on $\mathbf{M}_1(C([0, T], R))^{\otimes p}$. This proves the existence and uniqueness for the solutions of the equation (3.43) up to some small time T . We can easily extend this solution to any finite interval by considering $[0, T], [T, 2T], [2T, 3T], \dots$, etc. □

Chapter 4

McKean-Vlasov Type Limit

For each realization $\xi \in \Xi$, by the result in chapter 3, the system of stochastic differential equations

$$dx_i^{(n)}(t) = b(x_i^{(n)}(t))dt + \frac{1}{n} \sum_{j=1}^n c(\xi_i, \xi_j) \Phi(x_i^{(n)}(t), x_j^{(n)}(t))dt + \sigma(x_i^{(n)}(t))dw_i(t), \quad (4.1)$$

$$i = 1, \dots, n.$$

determines a unique Markov diffusion process $x_\xi^{(n)} = \{x_\xi^{(n)}(t) : t \geq 0\}$ in R^n , where

$$x_\xi^{(n)}(t) = (x_{\xi,1}^{(n)}(t), \dots, x_{\xi,n}^{(n)}(t)).$$

To such $x_\xi^{(n)}$ there are associated p local empirical measures

$$X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)}$$

which are $\mathbf{M}_1(C([0, \infty), R))$ -valued random measures and the empirical measure $X_\xi^{(n)}$ for the system (4.1) can be written as a convex combination of $X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)}$:

$$X_\xi^{(n)} = \sum_{k=1}^p \frac{n_{\xi, \theta_k}}{n} X_{\xi, \theta_k}^{(n)}$$

where $\frac{n_{\xi, \theta_k}}{n}$ is a random variable from $(\Xi, \mathcal{B}(\Xi), \wp)$ to $(0, 1)$ defined by

$$\frac{n_{\xi, \theta_k}}{n} = \frac{1}{n} \sum_{i=1}^n 1_{\theta_k}(\xi_i).$$

In this chapter we study the behavior of the random measure $X_\xi^{(n)}$ in the limit when n goes to infinity.

4.1 Statement of the result

The main result of this chapter can be stated as following:

Theorem 4.1 *Suppose that the hypotheses H_1 to H_3 in chapter 3 are satisfied. Furthermore, let the following two initial conditions are satisfied :*

1. $\sup_{n \geq 1} E(|x_{\xi,i}^{(n)}(0)|^4) < \infty, \quad \mathbb{P}$ -a.s., *for each* $i \geq 1$,
2. $X_\xi^{(n)}(0) \Rightarrow u$ (weak convergence) in $\mathbf{M}_1(R)$ *as* $n \rightarrow \infty, \quad \mathbb{P}$ -a.s.

Then for almost all $\xi \in \Xi$, the probability laws $(\mathcal{L}(X_\xi^{(n)}))_{n=1}^\infty$ of $(X_\xi^{(n)})_{n=1}^\infty$ converge weakly in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R)))$ to a Dirac measure concentrating at the point $\sum_{k=1}^p \lambda_k m_{\theta_k}$ i.e.

$$\mathbb{P}\{\xi : \mathcal{L}(X_\xi^{(n)}) \Rightarrow \delta_{\sum_{k=1}^p \lambda_k m_{\theta_k}}\} = 1$$

where $m_{\theta_k}, 1 \leq k \leq p$, is the law on $C([0, \infty), R)$ of the unique solution x_ξ of the following stochastic differential equation under the condition that $\xi = \theta_k$:

$$dx_\xi(t) = b(x_\xi(t))dt + \sum_{i=1}^p \lambda_i c(\xi, \theta_i) \left[\int_R \Phi(x_\xi(t), y) m_{\theta_i}(t, dy) dt \right] + \sigma(x_\xi(t))dw(t), \quad (4.2)$$

$$m_{\theta_i}(t, dy) = P\{x_{\theta_i}(t) \in dy\}$$

$$x_\xi(0) \text{ has distribution } u, \quad \mathbb{P}\text{-a.s.}$$

The proof of this theorem consists of two steps. In section 1 we first show that the local empirical measures $(X_\xi^{(n)}|_{\theta_1}, \dots, X_\xi^{(n)}|_{\theta_p})$ converge in probability to $(m_{\theta_1}, \dots, m_{\theta_p})$ almost surely with respect to ξ . Then in section 2 we carry this weak convergence result to the empirical measure $X_\xi^{(n)}$ by using a result on propagation of chaos. Finally, in section 3, we consider a simple example of the model described by (4.1).

4.2 Weak convergence of $\mathcal{L}(X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)})$

As in chapter 3 we let $\mathbf{M}_1(C([0, \infty), R))$ be the collection of probability measures on the path space $C([0, \infty), R)$ and $\mathbf{M}_1(C([0, \infty), R))^{\otimes p}$ be the p -fold product space of $\mathbf{M}_1(C([0, \infty), R))$. For $\xi \in \Xi$, $(X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)})$ is an element of $\mathbf{M}_1(C([0, \infty), R))^{\otimes p}$. Throughout this section we denote

$$Q_{\xi}^{(n)} = \mathcal{L}(X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)})$$

Then $Q_{\xi}^{(n)} \in \mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R))^{\otimes p})$.

Proposition 4.1 *Under the hypotheses of theorem 4.1, for almost all $\xi \in \Xi$, the probability laws $(Q_{\xi}^{(n)})_{n=1}^{\infty}$ of the local empirical measures converges weakly in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R))^{\otimes p})$ to a Dirac measure concentrating on the point $(m_{\xi, 1}, \dots, m_{\xi, p})$,*

$$\text{i.e.,} \quad \psi\{\xi : \mathcal{L}(X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)}) \Rightarrow \delta_{(m_{\xi, 1}, \dots, m_{\xi, p})}\} = 1$$

where $m_{\xi, k}$, $1 \leq k \leq p$, is the law on $C([0, \infty), R)$ of the process x_{n_k} determined uniquely by the equation (4.2).

Proof: We start by proving the following lemma.

Lemma 4.1 *Suppose that the condition 1, in theorem 4.1 is satisfied, i.e.,*

$$\sup_{n \geq 1} E(|x_{\xi, i}^{(n)}(0)|^4) < \infty, \quad \psi\text{-a.s.,} \quad \text{for each } i \geq 1. \quad (4.3)$$

Then for any $i \geq 1$ and $T > 0$

$$\sup_{n \geq 1} E\left(\sup_{0 \leq t \leq T} |x_{\xi, i}^{(n)}(t)|^4\right) < \infty, \quad \psi\text{-a.s.} \quad (4.4)$$

Proof: Denote $c = \max_{1 \leq k, l \leq p} |c(\theta_k, \theta_l)|$. Let $1 \leq i \leq n$ be fixed. By Itô's formula,

$$\begin{aligned} |x_{\xi, i}^{(n)}(t)|^2 &= |x_{\xi, i}^{(n)}(0)|^2 + \int_0^t [\sigma^2(x_{\xi, i}^{(n)}(s)) + 2x_{\xi, i}^{(n)}(s)b(x_{\xi, i}^{(n)}(s))] \\ &\quad + 2\frac{1}{n} \sum_{j=1}^n c(\xi_i, \xi_j) x_{\xi, i}^{(n)}(s) \Phi(x_{\xi, i}^{(n)}(s), x_{\xi, j}^{(n)}(s))] ds + 2M_{\xi, i}^{(n)}(t) \end{aligned} \quad (4.5)$$

where

$$M_{\xi, i}^{(n)}(t) = \int_0^t x_{\xi, i}^{(n)}(s) \sigma(x_{\xi, i}^{(n)}(s)) dw_i(s) \quad (4.6)$$

is a $\mathcal{F}_t$ -martingale for each given ξ because of the assumptions on $\sigma(x)$. For any $t \leq T$, using the inequality

$$(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$$

and the hypotheses H_1 and H_2 ,

$$\begin{aligned} |x_{\xi, i}^{(n)}(t)|^4 &= \{|x_{\xi, i}^{(n)}(0)|^2 + \int_0^t [\sigma^2(x_{\xi, i}^{(n)}(s)) + 2x_{\xi, i}^{(n)}(s)b(x_{\xi, i}^{(n)}(s))] \\ &\quad + \frac{2}{n} \sum_{j=1}^n c(\xi_i, \xi_j) x_{\xi, i}^{(n)}(s) \Phi(x_{\xi, i}^{(n)}(s), x_{\xi, j}^{(n)}(s))] ds + 2M_{\xi, i}^{(n)}(t)\}^2 \\ &\leq 3\{|x_{\xi, i}^{(n)}(0)|^4 + (\int_0^t [\sigma^2(x_{\xi, i}^{(n)}(s)) + 2x_{\xi, i}^{(n)}(s)b(x_{\xi, i}^{(n)}(s)) \\ &\quad + \frac{2}{n} \sum_{j=1}^n c(\xi_i, \xi_j) x_{\xi, i}^{(n)}(s) \Phi(x_{\xi, i}^{(n)}(s), x_{\xi, j}^{(n)}(s))] ds)^2 + (2M_{\xi, i}^{(n)}(t))^2\} \\ &\leq 3\{|x_{\xi, i}^{(n)}(0)|^4 + 2T \int_0^t K^2(1 + |x_{\xi, i}^{(n)}(s)|^2) ds \\ &\quad + \frac{8}{n} T c^2 \sum_{j=1}^n \int_0^t |x_{\xi, i}^{(n)}(s)|^2 K^2(|x_{\xi, i}^{(n)}(s)|^2 + |x_{\xi, j}^{(n)}(s)|^2) ds \\ &\quad + 4(M_{\xi, i}^{(n)}(t))^2\} \\ &\leq c_1 + c_2 |x_{\xi, i}^{(n)}(0)|^4 + c_3 \int_0^t |x_{\xi, i}^{(n)}(s)|^4 ds \\ &\quad + c_4 (M_{\xi, i}^{(n)}(t))^2 + c_5 \frac{1}{n} \sum_{j=1}^n \int_0^t |x_{\xi, j}^{(n)}(s)|^4 ds \end{aligned} \quad (4.7)$$

for some positive constants c_1, c_2, c_3, c_4 , and c_5 which depend only on T . Let

$$y_{\xi}^{(n)}(t) = \frac{1}{n} \sum_{j=1}^n \int_0^t |x_{\xi, j}^{(n)}(s)|^4 ds. \quad (4.8)$$

Then we have

$$\begin{aligned} E(\sup_{t \leq T} |x_{\xi, i}^{(n)}(t)|^4) &\leq c_1 + c_2 E(|x_{\xi, i}^{(n)}(0)|^4) + c_3 \int_0^T E(\sup_{s \leq t} |x_{\xi, i}^{(n)}(s)|^4) dt \\ &\quad + c_4 E(\sup_{t \leq T} (M_{\xi, i}^{(n)}(t))^2) + c_5 E(\sup_{t \leq T} y_{\xi}^{(n)}(t)) \end{aligned} \quad (4.9)$$

Since according to the stochastic differential equations (4.1), $x_{\xi,1}^{(n)}, \dots, x_{\xi,n}^{(n)}$ are identically distributed, we can write

$$\begin{aligned} E(\sup_{t \leq T} |y_{\xi}^{(n)}(t)|) &\leq \frac{1}{n} \sum_{j=1}^n \int_0^T E(\sup_{s \leq t} |x_{\xi,j}^{(n)}(s)|^4) dt \\ &= \int_0^T E(\sup_{s \leq t} |x_{\xi,1}^{(n)}(s)|^4) dt \end{aligned} \quad (4.10)$$

Also, by Doob's martingale inequality

$$\begin{aligned} E(\sup_{t \leq T} |M_{\xi,1}^{(n)}(t)|^2) &\leq 4E(|M_{\xi,1}^{(n)}(T)|^2) \\ &= 4E \int_0^T |x_{\xi,1}^{(n)}(s)|^2 |\sigma(x_{\xi,1}^{(n)}(s))|^2 ds \\ &\leq 4E \int_0^T |x_{\xi,1}^{(n)}(s)|^2 K(1 + |x_{\xi,1}^{(n)}(s)|^2) ds \\ &\leq 4KT + 8K \int_0^T E(\sup_{s \leq t} |x_{\xi,1}^{(n)}(s)|^4) dt \end{aligned} \quad (4.11)$$

(4.9), (4.10) and (4.11) together yields

$$\begin{aligned} E(\sup_{t \leq T} |x_{\xi,1}^{(n)}(t)|^4) &\leq k_1 + k_2(|x_{\xi,1}^{(n)}(0)|^4) + k_3 \int_0^T E(\sup_{s \leq t} |x_{\xi,1}^{(n)}(s)|^4) dt \end{aligned} \quad (4.12)$$

for some new positive constants k_1 , k_2 and k_3 . So by the Gronwall's inequality

$$\begin{aligned} E(\sup_{t \leq T} |x_{\xi,1}^{(n)}(t)|^4) &\leq \{k_1 + k_2 E(|x_{\xi,1}^{(n)}(0)|^4)\} \exp\{k_3 T\} \end{aligned} \quad (4.13)$$

Since the constants k_1 , k_2 and k_3 do not depend on n , we have

$$\begin{aligned} \sup_{n \geq 1} E(\sup_{t \leq T} |x_{\xi,1}^{(n)}(t)|^4) &\leq \{k_1 + k_2 \sup_{n \geq 1} E(|x_{\xi,1}^{(n)}(0)|^4)\} \exp\{k_3 T\} \\ &< \infty, \quad \text{a.s.} \quad \square \end{aligned} \quad (4.14)$$

Tightness of $(Q_\xi^{(n)})_{n=1}^\infty$

In order to prove the proposition 4.1 it is necessary to show that the sequence of probability measures $(Q_\xi^{(n)})_{n=1}^\infty$ is relatively compact in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R))^{\otimes p})$. Since $(\mathbf{M}_1(C([0, \infty), R))^{\otimes p})$ is a complete and separable metric space, it is equivalent, by theorem 1.2, to show that $(Q_\xi^{(n)})_{n=1}^\infty$ is tight in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R))^{\otimes p})$.

Proposition 4.2 *Under the conditions of theorem 4.1 the sequence $(Q_\xi^{(n)})_{n=1}^\infty$ of probability measures is tight in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R))^{\otimes p})$ uniformly in $\xi \in \Xi$ in the sense that for each $\epsilon > 0$, there exists a compact subset K_ϵ in $\mathbf{M}_1(C([0, \infty), R))^{\otimes p}$ such that*

$$\inf_n Q_\xi^{(n)}(K_\epsilon) > 1 - \epsilon$$

for almost all $\xi \in \Xi$.

Proof: Let $\pi_k : \mathbf{M}_1(C([0, \infty), R))^{\otimes p} \rightarrow \mathbf{M}_1(C([0, \infty), R))$ be the k th projective map and $Q_{\xi, k}^{(n)}$ be the image measure of $Q_\xi^{(n)}$ under this map:

$$Q_{\xi, k}^{(n)} = Q_\xi^{(n)} \circ \pi_k^{-1} \quad (4.15)$$

By theorem 1.4 of chapter 1, it suffices to prove that for each $k = 1, \dots, p$, the sequence $(Q_{\xi, k}^{(n)})_{n=1}^\infty$ of probability measures is tight in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R)))$. To this end it suffices to show, according to theorem 1.5 of chapter 1, that the intensity measure $(I_{\xi, k}^{(n)})_{n=1}^\infty$ of $(Q_{\xi, k}^{(n)})_{n=1}^\infty$ is tight in $\mathbf{M}_1(C([0, \infty), R))$ for each $k = 1, \dots, p$.

For any $f \in C_b(C([0, \infty), R))$, using the fact that $x_{\xi, 1}^{(n)}, \dots, x_{\xi, p}^{(n)}$ have the same distribution, we can write

$$\begin{aligned} & \langle I_{\xi, k}^{(n)}, f \rangle \\ &= \int_{\mathbf{M}_1(C([0, \infty), R))} \langle m, f \rangle Q_{\xi, k}^{(n)}(dm) \\ &= E(\langle \frac{1}{n_{\xi, \theta_k}} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \delta_{x_{\xi, i}^{(n)}}, f \rangle) \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{n_{\xi, \theta_k}} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n E(f(x_{\xi, i}^{(n)})) \\
 &= E(f(x_{\xi, 1}^{(n)}))
 \end{aligned} \tag{4.16}$$

Thus it suffices to show that the laws of $(x_{\xi, 1}^{(n)})_{n=1}^{\infty}$ is tight in $\mathbf{M}_1(C([0, \infty), R))$.

Now $x_{\xi, 1}^{(n)}$ is a continuous semimartingale which has the representation

$$x_{\xi, 1}^{(n)}(t) = A_{\xi}^{(n)}(t) + M_{\xi}^{(n)}(t) \tag{4.17}$$

where

$$\begin{aligned}
 A_{\xi}^{(n)}(t) &= x_{\xi, 1}^{(n)}(0) + \int_0^t b(x_{\xi, 1}^{(n)}(s)) ds \\
 &+ \frac{1}{n} \sum_{l=1}^p c(\xi, \theta_l) \sum_{\substack{j=1 \\ \xi_j = \theta_l}}^n \int_0^t \Phi(x_{\xi, 1}^{(n)}(s), x_{\xi, j}^{(n)}(s)) ds
 \end{aligned} \tag{4.18}$$

and

$$M_{\xi}^{(n)}(t) = \int_0^t \sigma(x_{\xi, 1}^{(n)}(s)) dw_1(s) \tag{4.19}$$

Hence to prove the proposition it suffices to check that conditions in theorem 1.6 are satisfied.

For each fixed $t > 0$, by the lemma 4.1

$$\sup_{n \geq 1} E(|x_{\xi, 1}^{(n)}(t)|^4) < \infty, \quad \varphi\text{-a.s.} \tag{4.20}$$

which implies the tightness of $(\mathcal{L}(x_{\xi, 1}^{(n)}(t)))_{n=1}^{\infty}$ in $\mathbf{M}_1(R)$. Thus the condition 1 of theorem 1.6 is satisfied.

To check the condition 2 of theorem 1.6 we let τ_n be any stopping time bounded by T and let θ be any positive number bounded by $\delta > 0$. Then we have for all $n \geq 1$

$$\begin{aligned}
 &E(|A_{\xi}^{(n)}(\tau_n + \theta) - A_{\xi}^{(n)}(\tau_n)|^2) \\
 &\leq 2E(|\int_{\tau_n}^{\tau_n + \theta} b(x_{\xi, 1}^{(n)}(s)) ds|^2)
 \end{aligned}$$

$$\begin{aligned}
 & + 2E(|\frac{1}{n} \sum_{l=1}^p c(\xi, \theta_l) \sum_{\substack{j=1 \\ \xi_j = \theta_l}}^n \int_{\tau_n}^{\tau_n + \theta} \Phi(x_{\xi, 1}^{(n)}(s), x_{\xi, j}^{(n)}(s)) ds|^2) \\
 & \leq 2\delta E(\int_{\tau_n}^{\tau_n + \theta} |b(x_{\xi, 1}^{(n)}(s))|^2 ds) \\
 & + \frac{2p}{n^2} \sum_{l=1}^p |c(\xi, \theta_l)|^2 n_{\xi, \theta_l} \sum_{\substack{j=1 \\ \xi_j = \theta_l}}^n \delta E(\int_{\tau_n}^{\tau_n + \theta} |\Phi(x_{\xi, 1}^{(n)}(s), x_{\xi, j}^{(n)}(s))|^2 ds) \\
 & \leq 2\delta E(\int_{\tau_n}^{\tau_n + \theta} K(1 + \sup_{s \leq T+\delta} |x_{\xi, 1}^{(n)}(s)|^2) ds) \\
 & + \frac{4p}{n^2} \sum_{l=1}^p c^2 n_{\xi, \theta_l} \sum_{\substack{j=1 \\ \xi_j = \theta_l}}^n \delta E(\int_{\tau_n}^{\tau_n + \theta} K^2(\sup_{s \leq T+\delta} |x_{\xi, 1}^{(n)}(s)|^2 + \sup_{s \leq T+\delta} |x_{\xi, j}^{(n)}(s)|^2) ds) \\
 & \leq 2\delta^2 K[1 + E(\sup_{s \leq T+\delta} |x_{\xi, 1}^{(n)}(s)|^2)] \\
 & + \frac{4p}{n^2} \sum_{l=1}^p c^2 n_{\xi, \theta_l} \sum_{\substack{j=1 \\ \xi_j = \theta_l}}^n \delta^2 K^2(E \sup_{s \leq T+\delta} |x_{\xi, 1}^{(n)}(s)|^2 + E \sup_{s \leq T+\delta} |x_{\xi, j}^{(n)}(s)|^2) \\
 & \leq \delta^2 [c_1 + c_2 \sup_{n \geq 1} E(\sup_{s \leq T+\delta} |x_{\xi, 1}^{(n)}(s)|^2)] \quad \text{for some constants of } c_1 \text{ and } c_2 \\
 & \rightarrow 0 \text{ as } \delta \rightarrow 0, \quad \varphi\text{-a.s.}
 \end{aligned} \tag{4.21}$$

because $\sup_{n \geq 1} E(\sup_{s \leq T+\delta} |x_{\xi, 1}^{(n)}(s)|^2)$ is finite φ -a.s. by 3 of theorem 3.1 Then by Chebyshev's inequality, for any positive ϵ and η , there exists $\delta > 0$ such that

$$\limsup_{n \rightarrow \infty} \sup_{\theta \leq \delta} P(|A_{\xi}^{(n)}(\tau_n + \theta) - A_{\xi}^{(n)}(\tau_n)| > \epsilon) \leq \eta, \quad \varphi\text{-a.s.} \tag{4.22}$$

Similarly for any $n \geq 1$

$$\begin{aligned}
 & E(|\ll M_{\xi}^{(n)} \gg_{\tau_n + \theta} - \ll M_{\xi}^{(n)} \gg_{\tau_n} |^2) \\
 & = E(|\int_{\tau_n}^{\tau_n + \theta} \sigma^2(x_{\xi, 1}^{(n)}(s)) ds|^2) \\
 & \leq \delta E(\int_{\tau_n}^{\tau_n + \theta} \sigma^4(x_{\xi, 1}^{(n)}(s)) ds) \\
 & \leq \delta E(\int_{\tau_n}^{\tau_n + \theta} 2K^2(1 + \sup_{s \leq T+\delta} |x_{\xi, 1}^{(n)}(s)|^4) ds) \\
 & \leq \delta^2 c_3 [1 + \sup_{n \geq 1} E(\sup_{s \leq T+\delta} |x_{\xi, 1}^{(n)}(s)|^4)] \\
 & \rightarrow 0 \text{ as } \delta \rightarrow 0, \quad \varphi\text{-a.s.}
 \end{aligned} \tag{4.23}$$

because $\sup_{n \geq 1} E(\sup_{s \leq T+\delta} |x_{\xi, 1}^{(n)}(s)|^4)$ is finite $\mathcal{P}$ -a.s. by lemma 4.1. By Chebyshev's inequality, for any positive ϵ and η positive, there exists $\delta > 0$ such that

$$\limsup_{n \rightarrow \infty} \sup_{\theta \leq \delta} P(|\langle M_{\xi}^{(n)} \rangle_{\tau_n + \theta} - \langle M_{\xi}^{(n)} \rangle_{\tau_n}| > \epsilon) \leq \eta, \quad \mathcal{P}\text{-a.s.} \quad (4.24)$$

Thus the condition 2 of theorem 1.6 is satisfied. Therefore for each $\epsilon > 0$, there exists a compact subset K_{ϵ} in $\mathbf{M}_1(C([0, \infty), R))^{\otimes p}$ such that

$$\inf_n \mathcal{L}(x_{\xi, 1}^{(n)})(K_{\epsilon}) > 1 - \epsilon, \quad \mathcal{P}\text{-a.s.} \quad (4.25)$$

It follows that $(Q_{\xi}^{(n)})_{n=1}^{\infty}$ is tight in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R))^{\otimes p})$ uniformly in ξ . $\square$

Identification of the limits

Since $(Q_{\xi}^{(n)})_{n=1}^{\infty}$ is relatively compact in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R))^{\otimes p})$, it has convergent subsequences. The next step is to characterize all the possible limit points.

For $m = (m_1, \dots, m_p) \in \mathbf{M}_1(R)^{\otimes p}$ consider the operators $L_{\theta_1}(m), \dots, L_{\theta_p}(m)$ which act on the function space $C_K^2(R)$ according to the following formulas

$$L_{\theta_k}(m)f(x) = \frac{\sigma^2(x)}{2}f''(x) + b(x)f'(x) + \sum_{l=1}^p \lambda_l c(\theta_k, \theta_l) \int_R \Phi(x, y) m_l(dy) f'(x) \quad (4.26)$$

$$k = 1, \dots, p.$$

Let X be the usual projective map on $C([0, \infty), R)$ and $\{\mathcal{B}_t, t \geq 0\}$ be the family of sigma algebras on $C([0, \infty), R)$ generated by X . let $u \in \mathbf{M}_1(R)$. A vector of probability measures

$$m = (m_{\theta_1}, \dots, m_{\theta_p}) \in \mathbf{M}_1(C([0, \infty), R))^{\otimes p}$$

is said to be a solution to the martingale problem associated with the operators $L_{\theta_k}(m)$, $k = 1, \dots, p$, and the initial data u if the following conditions are satisfied:

1. for each $k = 1, \dots, p$ X has initial distribution u under m_{θ_k}

$$m_{\theta_k}(X(0) \in \cdot) = u(\cdot), \quad (4.27)$$

2. for each $k = 1, \dots, p$ and each $f \in C_K^2(R)$ the process

$$M_{\theta_k, t}^{f, m}(X) = f(X(t)) - f(X(0)) - \int_0^t L_{\theta_k}(m(s))f(X(s))ds \quad (4.28)$$

is a $(\mathcal{B}_t, m_{\theta_k})$ -martingale.

Proposition 4.3 *Let the conditions in theorem 4.1 be satisfied. Suppose that Q_ξ is a limit point of $(Q_\xi^{(n)})_{n=1}^\infty$. Then Q_ξ does not depend on ξ and concentrates on the measures $m = (m_{\theta_1}, \dots, m_{\theta_p})$ which are the solutions of the martingale problem associated with the operators $L_{\theta_k}(m)$, $k = 1, \dots, p$, and the initial data u .*

Proof: For each fixed $f \in C_K^2(R)$ and $0 \leq s \leq t$ we define the map

$$F : \mathbf{M}_1(C([0, \infty), R))^{\otimes p} \rightarrow [0, \infty)$$

by

$$F(m_1, \dots, m_p) = \sum_{k=1}^p (\langle m_k, [M_{\theta_k, t}^{f, m} - M_{\theta_k, s}^{f, m}]G_k \rangle)^2 \quad (4.29)$$

where $M_{\theta_k, t}^{f, m}$ is given by (4.28) and G_k , $k = 1, \dots, p$, are bounded continuous functions on $C([0, \infty), R)$ which are $\mathcal{B}_s$ -measurable. For each fixed positive integer K , $F(\cdot) \wedge K$ is a bounded continuous function on $\mathbf{M}_1(C([0, \infty), R))$. For simplicity, we still use $(Q_\xi^{(n)})_{n=1}^\infty$ to denote the subsequence which converges to Q_ξ . As in (3.30) we denote

$$Y_\xi^{(n)} = (X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)}).$$

Then for all $\xi \in \Xi$, we have

$$\langle F(\cdot) \wedge K, Q_\xi \rangle$$

$$\begin{aligned}
 &= \lim_{n \rightarrow \infty} \langle F(\cdot) \wedge K, Q_{\xi}^{(n)} \rangle \\
 &= \lim_{n \rightarrow \infty} \sum_{k=1}^p E \left\{ \left(\frac{1}{n_{\xi, \theta_k}} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n [M_{\theta_k, t}^{f, Y_{\xi}^{(n)}}(x_{\xi, i}^{(n)}) - M_{\theta_k, s}^{f, Y_{\xi}^{(n)}}(x_{\xi, i}^{(n)})] G_k(x_{\xi, i}^{(n)}) \right)^2 \wedge K \right\} \\
 &= \lim_{n \rightarrow \infty} [I_{1, K}^{(n)} + I_{2, K}^{(n)}] \tag{4.30}
 \end{aligned}$$

where

$$I_{1, K}^{(n)} = \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n E([M_{\theta_k, t}^{f, Y_{\xi}^{(n)}}(x_{\xi, i}^{(n)}) - M_{\theta_k, s}^{f, Y_{\xi}^{(n)}}(x_{\xi, i}^{(n)})]^2 G_k^2(x_{\xi, i}^{(n)}) \wedge K) \tag{4.31}$$

and

$$\begin{aligned}
 I_{2, K}^{(n)} &= \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i, j \\ \xi_i = \xi_j = \theta_k}} E([M_{\theta_k, t}^{f, Y_{\xi}^{(n)}}(x_{\xi, i}^{(n)}) - M_{\theta_k, s}^{f, Y_{\xi}^{(n)}}(x_{\xi, i}^{(n)})] \\
 &\quad [M_{\theta_k, t}^{f, Y_{\xi}^{(n)}}(x_{\xi, j}^{(n)}) - M_{\theta_k, s}^{f, Y_{\xi}^{(n)}}(x_{\xi, j}^{(n)})] G_k(x_{\xi, i}^{(n)}) G_k(x_{\xi, j}^{(n)}) \wedge K) \tag{4.32}
 \end{aligned}$$

By definition

$$\begin{aligned}
 &M_{\theta_k, t}^{f, Y_{\xi}^{(n)}}(x_{\xi, i}^{(n)}) \\
 &= f(x_{\xi, i}^{(n)}(t)) - f(x_{\xi, i}^{(n)}(0)) - \int_0^t L_{\theta_k}(Y_{\xi}^{(n)}(s)) f(x_{\xi, i}^{(n)}(s)) ds \\
 &= f(x_{\xi, i}^{(n)}(t)) - f(x_{\xi, i}^{(n)}(0)) - \int_0^t L_{\xi, \theta_k}^{(n)}(Y_{\xi}^{(n)}(s)) f(x_{\xi, i}^{(n)}(s)) ds \\
 &+ \sum_{l=1}^p \left(\frac{n_{\xi, \theta_l}}{n} - \lambda_l \right) c(\mathcal{U}_s, \mathcal{J}_l) \int_0^t \int_R \Phi(x_{\xi, i}^{(n)}(s), z) X_{\xi, l}^{(n)}(s, dz) f'(x_{\xi, i}^{(n)}(s)) ds \\
 &= \int_0^t \sigma(x_{\xi, i}^{(n)}(s)) dw_i(s) \\
 &+ \sum_{l=1}^p \left(\frac{n_{\xi, \theta_l}}{n} - \lambda_l \right) c(\theta_k, \theta_l) \int_0^t \int_R \Phi(x_{\xi, i}^{(n)}(s), z) X_{\xi, l}^{(n)}(s, dz) f'(x_{\xi, i}^{(n)}(s)) ds \tag{4.33}
 \end{aligned}$$

where $L_{\xi, \theta_k}^{(n)}(Y_{\xi}^{(n)}(s))$ is the operator defined by (3.32). We have

$$\begin{aligned}
 &\lim_{n \rightarrow \infty} I_{1, K}^{(n)} \\
 &= \lim_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n E \left\{ \left[\int_s^t \sigma(x_{\xi, i}^{(n)}(u)) dw_i(u) + \sum_{l=1}^p \left(\frac{n_{\xi, \theta_l}}{n} - \lambda_l \right) c(\theta_k, \theta_l) \right. \right.
 \end{aligned}$$

$$\begin{aligned}
 & \int_s^t \int_R \Phi(x_{\xi, i}^{(n)}(u), z) X_{\xi, i}^{(n)}(u, dz) f'(x_{\xi, i}^{(n)}(u)) du]^2 G_k^2(x_{\xi, i}^{(n)}) \wedge K \} \\
 \leq & c_1 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n E(|\int_s^t \sigma(x_{\xi, i}^{(n)}(u)) dw_i(u)|^2) \\
 + & c_2 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \sum_{l=1}^p (\frac{n_{\xi, \theta_l}}{n} - \lambda_l)^2 \\
 & (t-s) \int_s^t E(\int_R |\Phi(x_{\xi, i}^{(n)}(u), z)|^2 X_{\xi, i}^{(n)}(u, dz) |f'(x_{\xi, i}^{(n)}(u))|^2) du \\
 \leq & c_1 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \int_s^t K(1 + |x_{\xi, i}^{(n)}(u)|^2) du \\
 + & c_3 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \sum_{l=1}^p (\frac{n_{\xi, \theta_l}}{n} - \lambda_l)^2 \\
 & (t-s) \int_s^t \int_R 2K^2 E(|x_{\xi, i}^{(n)}(u)|^2 + |z|^2) X_{\xi, i}^{(n)}(u, dz) du \\
 \leq & c_1 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n K(t-s)(1 + E(\sup_{u \leq t} |x_{\xi, i}^{(n)}(u)|^2)) \\
 + & c_3 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \sum_{l=1}^p (\frac{n_{\xi, \theta_l}}{n} - \lambda_l)^2 (t-s)^2 c_4 E(\sup_{u \leq t} |x_{\xi, i}^{(n)}(u)|^2) \\
 \leq & c_5 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n [1 + E(\sup_{u \leq t} |x_{\xi, i}^{(n)}(u)|^2)] \\
 + & c_6 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \sum_{l=1}^p (\frac{n_{\xi, \theta_l}}{n} - \lambda_l)^2 E(\sup_{u \leq t} |x_{\xi, i}^{(n)}(u)|^2) \\
 \leq & c_5 \sum_{k=1}^p \limsup_{n \rightarrow \infty} \frac{1}{n_{\xi, \theta_k}} [1 + \sup_{n \geq 1} E(\sup_{u \leq t} |x_{\xi, i}^{(n)}(u)|^2)] \\
 + & c_6 \sum_{k=1}^p \limsup_{n \rightarrow \infty} \frac{1}{n_{\xi, \theta_k}} \sum_{l=1}^p (\frac{n_{\xi, \theta_l}}{n} - \lambda_l)^2 \sup_{n \geq 1} E(\sup_{u \leq t} |x_{\xi, i}^{(n)}(u)|^2) \\
 = & 0 \quad \text{p- a.s.} \tag{4.34}
 \end{aligned}$$

because of the lemma 4.1 and the strong law of large numbers for $\frac{n_{\xi, \theta_k}}{n}$.

Also, for $i \neq j$

$$\lim_{n \rightarrow \infty} I_{2, K}^{(n)}$$

$$\begin{aligned}
 &= \lim_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i \neq j \\ \xi_i = \xi_j = \theta_k}} E \left\{ \left(\int_s^t \sigma(x_{\xi, i}^{(n)}(u)) dw_i(u) + \sum_{l=1}^p \left(\frac{n_{\xi, \theta_l}}{n} - \lambda_l \right) c(\theta_k, \theta_l) \right. \right. \\
 &\quad \left. \left. \int_s^t \int_R \Phi(x_{\xi, i}^{(n)}(u), z) X_{\xi, l}^{(n)}(u, dz) f'(x_{\xi, i}^{(n)}(u)) du \right) \right. \\
 &\quad \left. \left(\int_s^t \sigma(x_{\xi, j}^{(n)}(u)) dw_j(u) + \sum_{l'=1}^p \left(\frac{n_{\xi, \theta_{l'}}}{n} - \lambda_{l'} \right) c(\theta_k, \theta_{l'}) \right) \right. \\
 &\quad \left. \left. \int_s^t \int_R \Phi(x_{\xi, j}^{(n)}(u), z) X_{\xi, l'}^{(n)}(u, dz) f'(x_{\xi, j}^{(n)}(u)) du \right) G_k(x_{\xi, i}^{(n)}) G_k(x_{\xi, j}^{(n)}) \wedge K \right\} \\
 &\leq c_7 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i \neq j \\ \xi_i = \xi_j = \theta_k}} E \left\{ \sum_{l=1}^p \sum_{l'=1}^p \left(\frac{n_{\xi, \theta_l}}{n} - \lambda_l \right) \left(\frac{n_{\xi, \theta_{l'}}}{n} - \lambda_{l'} \right) \right. \\
 &\quad \left. \int_s^t \left[\int_R \Phi(x_{\xi, i}^{(n)}(u), z) X_{\xi, l}^{(n)}(u, dz) \right] f'(x_{\xi, i}^{(n)}(u)) du \right. \\
 &\quad \left. \int_s^t \left[\int_R \Phi(x_{\xi, j}^{(n)}(u), z) X_{\xi, l'}^{(n)}(u, dz) \right] f'(x_{\xi, j}^{(n)}(u)) du \right\} \\
 &\leq c_7 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \sum_{l=1}^p \sum_{l'=1}^p \left(\frac{n_{\xi, \theta_l}}{n} - \lambda_l \right) \left(\frac{n_{\xi, \theta_{l'}}}{n} - \lambda_{l'} \right) \\
 &\quad E \left(\int_s^t \left[\int_R \Phi(x_{\xi, i}^{(n)}(u), z) X_{\xi, l}^{(n)}(u, dz) \right] f'(x_{\xi, i}^{(n)}(u)) du \right)^2 \\
 &\leq c_7 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n \sum_{l=1}^p \sum_{l'=1}^p \left(\frac{n_{\xi, \theta_l}}{n} - \lambda_l \right) \left(\frac{n_{\xi, \theta_{l'}}}{n} - \lambda_{l'} \right) \\
 &\quad (t-s) \int_s^t E \left[\int_R c_8 (|x_{\xi, i}^{(n)}(u)|^2 + |z|^2) X_{\xi, l}^{(n)}(u, dz) \right] du \\
 &\leq c_9 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^p \sum_{l=1}^p \sum_{l'=1}^p \left(\frac{n_{\xi, \theta_l}}{n} - \lambda_l \right) \left(\frac{n_{\xi, \theta_{l'}}}{n} - \lambda_{l'} \right) E \left(\sup_{u \leq t} |x_{\xi, i}^{(n)}(u)|^2 \right) \\
 &\leq c_9 \limsup_{n \rightarrow \infty} \sum_{k=1}^p \frac{1}{n_{\xi, \theta_k}^2} \sum_{l=1}^p \sum_{l'=1}^p \left(\frac{n_{\xi, \theta_l}}{n} - \lambda_l \right) \left(\frac{n_{\xi, \theta_{l'}}}{n} - \lambda_{l'} \right) \sup_{n \geq 1} E \left(\sup_{u \leq t} |x_{\xi, i}^{(n)}(u)|^2 \right) \\
 &= 0 \quad \text{p-a.s.} \tag{4.35}
 \end{aligned}$$

because of the lemma 4.1 and the strong law of large numbers for $\frac{n_{\xi, \theta_k}}{n}$, (4.34) and (4.35) together gives $\langle F(\cdot) \wedge K, Q_\xi \rangle = 0$ for each positive number K and almost all ξ . Let K goes to infinity yields

$$\langle F(\cdot), Q_\xi \rangle = 0 \quad \text{for almost all } \xi \in \Xi \tag{4.36}$$

Thus for almost all $\xi \in \Xi$

$$Q_\xi\{(m_1, \dots, m_p) : \langle m_k, [M_{\theta_k, t}^{f, m} - M_{\theta_k, s}^{f, m}]G_k \rangle = 0, 1 \leq k \leq p\} = 1. \quad (4.37)$$

This means that the limit measure Q_ξ concentrates on the probability measures $m = (m_1, \dots, m_p)$ in $\mathbf{M}_1(C([0, \infty), R))^{\otimes p}$ such that the process $M_{\theta_k, t}^{f, m}$ defined by (4.28) is a $(\mathcal{B}_t, m_{\theta_k})$ -martingale for each $k = 1, \dots, p$ and each $f \in C_K^2(R)$. Thus if $X_\xi^{(n)}(0) \Rightarrow u$ in $\mathbf{M}_1(R)$ as given by the condition 2 of theorem 4.1. Then m is the corresponding solution of the martingale problem associated with the operators $L_{\theta_k}(m)$, $k = 1, \dots, p$ and the initial data u . Since this martingale problem does not depend on the particular realization of the random variable ξ . The limit measure Q_ξ is in fact no longer random with respect to ξ . $\square$

So far we have proved that under the hypotheses of theorem 4.1, for almost all $\xi \in \Xi$, the probability law $\mathcal{L}(X_{\xi, \theta_1}^{(n)}, \dots, X_{\xi, \theta_p}^{(n)})$ of the local empirical measures is relatively compact in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R))^{\otimes p})$ and any limit points concentrate on the set of probability measures which are solutions to the martingale problem associated with the operators $L_{\theta_k}(m)$, $k = 1, \dots, p$, and the initial data u . It remains to show that the martingale problem has a unique solution.

Proposition 4.4 *The martingale problem associated with the operators $L_{\theta_k}(m)$, $k = 1, \dots, p$, and the initial data u has a unique solution $(m_{\theta_1}, \dots, m_{\theta_p})$, where m_{θ_k} is the probability distribution of x_{θ_k} of the unique solution x_ξ of the following stochastic differential equation*

$$dx_\xi(t) = b(x_\xi(t))dt + \sum_{i=1}^p \lambda_i c(\xi, \theta_i) \int_R \Phi(x_\xi(t), y) m_{\theta_i}(t, dy) dt + \sigma(x_\xi(t)) dw(t), \quad (4.38)$$

$$m_{\theta_i}(t, dy) = P\{x_{\theta_i}(t) \in dy\}$$

$$x_\xi(0) \text{ has distribution } u \quad \text{P-a.s.}$$

Proof: From theorem 3.2 of chapter 3, the equation (4.38) has a unique solution x_ξ . Let $\mu = (\mu_{\theta_1}, \dots, \mu_{\theta_p})$ be the vectors of the laws of this solution $(x_{\theta_1}, \dots, x_{\theta_p})$. Then, by Itô's formula, it is easy to check that μ is a solution of the martingale problem. The uniqueness of the martingale problem follows directly from the uniqueness of the equation (4.38). $\square$

This completes the proof of proposition 4.1. In the next section we use the fact that the empirical measure $X_\xi^{(n)}$ is a convex combination of $X_{\xi,1}^{(n)}, \dots, X_{\xi,p}^{(n)}$ to derive the almost sure weak convergence result for $(\mathcal{L}(X_\xi^{(n)}))_{n=1}^\infty$. $\square$

4.3 Weak convergence of $\mathcal{L}(X_\xi^{(n)})$

We first introduce the notion of propagation of chaos [46]. Let S be a complete and separable metric space with the Borel σ -algebra $\mathcal{B}(S)$, and $S^{\otimes n}$ be the n -fold product space of S with corresponding product σ -algebra $\mathcal{B}(S)^{\otimes n}$. Let $\mu \in \mathbf{M}_1(S)$.

Definition 4.1 A sequence $\{\mu_n \in \mathbf{M}_1(S^{\otimes n}), n \geq 1\}$ of symmetric probability measures is said to be μ -chaotic if for any $m \geq 1$ and any collection of functions $f_1, \dots, f_m \in C_b(S)$,

$$\lim_{n \rightarrow \infty} \langle \mu_n, f_1 \otimes \dots \otimes f_m \otimes 1 \otimes \dots \otimes 1 \rangle = \prod_{k=1}^m \langle \mu, f_k \rangle \quad (4.39)$$

Let $x_1, \dots, x_n$ be S -valued random variables with distribution μ_n and let X_n be the corresponding empirical measure defined by

$$X_n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i} \in \mathbf{M}_1(S)$$

Proposition 4.5 The following statements are equivalent:

1. The sequence $(\mu_n)_{n=1}^\infty$ is μ -chaotic;

2 OF 2

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

2. $\mathcal{L}(X_n) \Rightarrow \delta_\mu$ in $\mathbf{M}_1(\mathbf{M}_1(S))$ as $n \rightarrow \infty$;

3. $\mu_n\{(x_1, \dots, x_n) : |\frac{1}{n} \sum_{i=1}^n f(x_i) - \langle \mu, f \rangle| > \epsilon\} \rightarrow 0$ as $n \rightarrow \infty$ for every $f \in C_b(S)$.

Proof: The proof is mainly from [46]. We use E^{μ_n} to denote the expectation with respect to the measure μ_n

1 $\rightarrow$ 3. Suppose that (μ_n) is μ -chaotic and $f \in C_b(S)$. By the symmetric assumption on μ_n , we have

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} E^{\mu_n}(\langle X_n, f \rangle - \langle \mu, f \rangle)^2 \\
 &= \lim_{n \rightarrow \infty} E^{\mu_n} \left(\frac{1}{n^2} \sum_{i=1}^n f^2(x_i) + \frac{1}{n^2} \sum_{i \neq j} f(x_i) f(x_j) \right. \\
 &\quad \left. - \langle \mu, f \rangle^2 \right) \\
 &= \lim_{n \rightarrow \infty} \left[\frac{1}{n} E^{\mu_n}(f^2(x_1)) + \frac{n-1}{n} E^{\mu_n}(f(x_1) f(x_2)) \right. \\
 &\quad \left. - 2 \langle \mu, f \rangle E^{\mu_n}(f(x_1)) + \langle \mu, f \rangle^2 \right] \\
 &= \langle \mu, f \rangle^2 - 2 \langle \mu, f \rangle^2 + \langle \mu, f \rangle^2 = 0
 \end{aligned} \tag{4.40}$$

This yield 3 by Chebychev's inequality.

3 $\rightarrow$ 2. Let $(P_n)_{n=1}^\infty$ and P be random measures in $\mathbf{M}_1(S)$. Then it is well known that $P_n \Rightarrow P$ if and only if $\langle P_n, f \rangle \Rightarrow \langle P, f \rangle$ for all $f \in C_K(S)$. Now take $P_n = X_n$ and $P = \mu$, then it is clear that 3 implies 2.

2 $\rightarrow$ 1. suppose that 2 holds. For $f_1, \dots, f_m \in C_b(S)$ we have

$$\begin{aligned}
 & |E^{\mu_n}(f_1 \otimes \dots \otimes f_m \otimes 1 \otimes \dots \otimes 1) - \prod_{k=1}^m \langle \mu, f_k \rangle| \\
 &\leq |E^{\mu_n}(f_1 \otimes \dots \otimes f_m \otimes 1 \otimes \dots \otimes 1) - E^{\mu_n}(\prod_{k=1}^m \langle X_n, f_k \rangle)| \\
 &\quad + |E^{\mu_n}(\prod_{k=1}^m \langle X_n, f_k \rangle) - \prod_{k=1}^m \langle \mu, f_k \rangle|
 \end{aligned} \tag{4.41}$$

Since $\mathcal{L}(X_n) \Rightarrow \delta_\mu$ in $\mathbf{M}_1(\mathbf{M}_1(S))$ as $n \rightarrow \infty$ the second term in the right hand side of (4.41) certainly goes to 0 as $n \rightarrow \infty$. So we need to show that the first term in the right hand side of (4.41) converges to zero as $n \rightarrow \infty$. Since the measure μ_n are symmetric and

$$\prod_{k=1}^m \langle X_n, f_k \rangle = \frac{1}{n^m} \prod_{k=1}^m \sum_{l=1}^n f_k(x_l) \quad (4.42)$$

it suffices to verify that

$$\sup_{x \in S^n} \left| \frac{1}{n!} \sum_{\sigma \in \pi_n} \prod_{k=1}^m f_k(x_{\sigma(k)}) - \frac{1}{n^m} \prod_{k=1}^m \sum_{l=1}^n f_k(x_l) \right| \rightarrow 0 \text{ as } n \rightarrow \infty \quad (4.43)$$

where π_n is a set of all permutations of $\{1, 2, \dots, n\}$.

Since

$$\sum_{\sigma \in \pi_n} \prod_{k=1}^m f_k(x_{\sigma(k)}) = (n-m)! \sum_{l_1, \dots, l_m}^* f_1(x_{l_1}) \cdots f_m(x_{l_m}) \quad (4.44)$$

where $\sum^*$ stands for the summation on all different sets of indices $l_1, \dots, l_m$ such that $l_i \neq l_j$ if $i \neq j$ and

$$\prod_{k=1}^m \sum_{l=1}^n f_k(x_l) = \sum_{l_1, \dots, l_m}^* f_1(x_{l_1}) \cdots f_m(x_{l_m}) + A \quad (4.45)$$

where

$$|A| \leq M^m \left(n^m - \frac{n!}{(n-m)!} \right), \quad M = \max_{1 \leq k \leq m} \|f_k\|, \quad (4.46)$$

we have

$$\begin{aligned} & \left| \frac{1}{n!} \sum_{\sigma \in \pi_n} \prod_{k=1}^m f_k(x_{\sigma(k)}) - \frac{1}{n^m} \prod_{k=1}^m \sum_{l=1}^n f_k(x_l) \right| \\ & \leq \left| \frac{(n-m)!}{n!} \sum_{l_1, \dots, l_m}^* \prod_{k=1}^m f_k(x_{l_k}) - \frac{1}{n^m} \sum_{l_1, \dots, l_m}^* \prod_{k=1}^m f_k(x_{l_k}) \right| + |A| \frac{1}{n^m} \\ & \leq \left[\frac{(n-m)!}{n!} - \frac{1}{n^m} \right] \frac{n!}{(n-m)!} M^m + M^m \left[1 - \frac{n!}{n^m (n-m)!} \right] \\ & = 2M^m \left[1 - \frac{n!}{n^m (n-m)!} \right] \end{aligned} \quad (4.47)$$

Since for each fixed m and M

$$2M^m \left[1 - \frac{n!}{n^m (n-m)!} \right] \rightarrow 0 \text{ as } n \rightarrow \infty \quad (4.48)$$

the proof is completed. $\square$

We are now ready to complete the proof of the main theorem 4.1 of this chapter.

Let $S = C([0, \infty), \mathbb{R})$ and $\mu = \sum_{k=1}^p \lambda_k m_{\theta_k}$. Then for any $f \in C_b(C([0, \infty), \mathbb{R}))$

$$\begin{aligned}
 & E\left\{\left|\frac{1}{n} \sum_{i=1}^n f(x_{\xi,i}^{(n)}) - \langle \mu, f \rangle\right|^2\right\} \\
 &= E\left\{\left|\sum_{k=1}^p \frac{n_{\xi,\theta_k}}{n} \langle X_{\xi,\theta_k}^{(n)}, f \rangle - \sum_{k=1}^p \lambda_k \langle m_{\theta_k}, f \rangle\right|^2\right\} \\
 &\leq 2E\left\{\left|\sum_{k=1}^p \left[\frac{n_{\xi,\theta_k}}{n} - \lambda_k\right] \langle m_{\theta_k}, f \rangle\right|^2\right\} \\
 &+ 2E\left\{\left|\sum_{k=1}^p \frac{n_{\xi,\theta_k}}{n} (\langle X_{\xi,\theta_k}^{(n)}, f \rangle - \langle m_{\theta_k}, f \rangle)\right|^2\right\} \\
 &\leq 2p \sum_{k=1}^p \left[\frac{n_{\xi,\theta_k}}{n} - \lambda_k\right]^2 E(|\langle m_{\theta_k}, f \rangle|^2) \\
 &+ 2p \sum_{k=1}^p \left|\frac{n_{\xi,\theta_k}}{n}\right|^2 E(|\langle X_{\xi,\theta_k}^{(n)}, f \rangle - \langle m_{\theta_k}, f \rangle|^2) \\
 &\leq k_1 \sum_{k=1}^p \left[\frac{n_{\xi,\theta_k}}{n} - \lambda_k\right]^2 \\
 &+ k_2 \sum_{k=1}^p E(|\langle X_{\xi,\theta_k}^{(n)}, f \rangle - \langle m_{\theta_k}, f \rangle|^2) \tag{4.49}
 \end{aligned}$$

for some constants k_1 and k_2 . Since by the strong law of large numbers, $\frac{n_{\xi,\theta_k}}{n}$ tends to λ_k almost surely, the first term above clearly tends to zero as $n \rightarrow \infty$. The second term above also tends to zero because the random probability measures $(X_{\theta_1}^{(n)}, \dots, X_{\theta_p}^{(n)})$ converges in probability to $(m_{\theta_1}, \dots, m_{\theta_p})$ and the real valued random variables $\langle X_{\theta_1}^{(n)}, f \rangle, \dots, \langle X_{\theta_p}^{(n)}, f \rangle$ are bounded. It follows from the Chebychev's inequality and proposition 4.2 that $(\mathcal{L}(X_\xi^{(n)}))_{n=1}^\infty$ converges weakly to $\delta_{\sum_{k=1}^p \lambda_k m_{\theta_k}}$. This completes the proof of the theorem 4.1 $\square$

4.4 A nonlinear neural network model

In this section we apply the results developed in this chapter to analyze the behavior of a continuous nonlinear neural network model. This model is given by the following system of n Itô equations in R :

$$dx_i^{(n)}(t) = -x_i^{(n)}(t)dt + \frac{1}{n} \sum_{j=1}^n (\xi_i \cdot \xi_j) \varphi(x_j^{(n)}(t))dt + \sigma dw_i(t), \quad (4.50)$$

$$i = 1, \dots, n,$$

where $\xi_1, \dots, \xi_n$ are n i.i.d. Θ -valued random vectors with distribution $\lambda = \sum_{k=1}^p \lambda_k \delta_{\theta_k}$. $\xi_i \cdot \xi_j$ denotes the inner product in R^q . σ is a positive constant and

$$\varphi(x) = \tanh(\alpha x), \quad \alpha > 0 \quad (4.51)$$

Clearly the system of equations (4.50) satisfies the conditions in theorem 3.1. So it defines a Markov diffusion process $(x_{\xi,1}^{(n)}(\cdot), \dots, x_{\xi,n}^{(n)}(\cdot))$. Let $X_\xi^{(n)}$ be the associated empirical measure defined by

$$X_\xi^{(n)} = \frac{1}{n} \sum_{i=1}^n \delta_{x_{\xi,i}^{(n)}(\cdot)}. \quad (4.52)$$

According to the theorem 4.1, for almost all ξ , $(X_\xi^{(n)})_{n=1}^\infty$ converges as $n \rightarrow \infty$ in probability to a deterministic measure $m = \sum_{k=1}^p \lambda_k m_{\theta_k}$ in $\mathbf{M}_1(C([0, \infty), R))$, where m_{θ_k} is the law of the unique solution of the following stochastic differential equation

$$dx_\xi(t) = -x_\xi(t)dt + \sum_{l=1}^p \lambda_l (\xi \cdot \theta_l) \int_R \varphi(y) m_{\theta_l}(t, dy)dt + \sigma dw(t), \quad (4.53)$$

given $\xi = \theta_k$

$$m_{\theta_l}(t, dy) = P\{x_{\theta_l}(t) \in dy\}, \quad l = 1, \dots, p \quad (4.54)$$

The limit process (4.53) has a unique invariant probability measure which has the form of

$$P_\xi(x) = \frac{1}{Z_\xi} \exp\left\{-\frac{1}{\sigma^2}(x^2 - 2\Phi_\xi(a_1, \dots, a_p)x)\right\} \quad (4.55)$$

where

$$\Phi_\xi(a_1, \dots, a_p) = \sum_{l=1}^p \lambda_l (\xi \cdot \theta_l) a_l \quad (4.56)$$

and

$$a_l = \int_R \varphi(x) P_{\theta_l}(x) dx, \quad l = 1, \dots, p \quad (4.57)$$

Thus the invariant measure $P_\xi(x)$ of the limit process (4.52) is completely determined by the self-consistent equations (4.57). These equations can be written as

$$\begin{aligned} a_k &= \int_R \varphi(x) P_{\theta_k}(x) dx \\ &= \int_R \varphi(x) \frac{1}{Z_{\theta_k}} \exp\left\{-\frac{1}{\sigma^2}(x^2 - 2\Phi_{\theta_k}(a_1, \dots, a_p)x)\right\} dx \\ &= \int_R \varphi(x) \frac{1}{\sqrt{\pi}\sigma} \exp\left\{-\frac{1}{\sigma^2}(x - \Phi_{\theta_k}(a_1, \dots, a_p))^2\right\} dx \\ &= \int_R \varphi(x) \mu_{\sigma,k}(dx), \quad k = 1, \dots, p \end{aligned} \quad (4.58)$$

where $\mu_{\sigma,k}$ are Gaussian measures on R with mean $\Phi_{\theta_k}(a_1, \dots, a_p)$ and variance $\frac{\sigma^2}{2}$. Since $\varphi(x)$ is a bounded continuous function on R and $\mu_{\sigma,k}$ converges weakly to the Dirac measure concentrating on $\delta_{\Phi_{\theta_k}(a_1, \dots, a_p)}$ as σ goes to zero

$$\mu_{\sigma,k} \Rightarrow \delta_{\Phi_{\theta_k}(a_1, \dots, a_p)} \quad \text{as } \sigma \rightarrow 0, \quad (4.59)$$

we obtain, in the weak noise limit, the equations

$$\begin{aligned} a_k &= \varphi(\Phi_{\theta_k}(a_1, \dots, a_p)) \\ &= \tanh\left(\alpha \sum_{l=1}^p \lambda_l (\theta_k \cdot \theta_l) a_l\right) \\ k &= 1, \dots, p \end{aligned} \quad (4.60)$$

Suppose that the p patterns $\theta_1, \dots, \theta_p$ are orthonormal vectors in R^q , i.e.,

$$\theta_k \cdot \theta_l = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \neq l \end{cases} \quad (4.61)$$

Then the equations (4.60) can be reduced to

$$a_k = \tanh(\alpha \lambda_k a_k) \quad (4.62)$$

$$k = 1, \dots, p$$

Denote $a = (a_1, \dots, a_p)$. Then clearly $a = 0$ is always a solution of equation (4.62) for any $\alpha > 0$. We are more interested in the non-zero solutions.

Proposition 4.6 *There exists a critical value α_c such that for $\alpha \leq \alpha_c$ $a = 0$ is the only solution of (4.62), and for $\alpha > \alpha_c$ the equation (4.62) has nontrivial solutions.*

Proof: Let

$$\alpha_c = \min\left\{\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_p}\right\} \quad (4.63)$$

Then $\alpha \leq \alpha_c$ implies $\max_{1 \leq k \leq p} \alpha \lambda_k \leq 1$. This, in turn, implies that (4.62) has only a zero solution. On the other hand, if $\alpha > \alpha_c$ then $\alpha \lambda_k > 1$ for at least one k . This implies (4.62) has non-zero solution. $\square$

When $\alpha > \alpha_c$, the fixed point equations have non-zero solutions. In general there are many of them. For example, if $\lambda_1, \dots, \lambda_p$ and α satisfy the condition

$$\frac{1}{\lambda_k} < \alpha \leq \min_{l: l \neq k} \left(\frac{1}{\lambda_l}\right) \quad (4.64)$$

Then the non-zero solution will have the form of

$$(a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_p) = (0, \dots, 0, c, 0, \dots, 0)$$

where $c \in (-1, 1)$. Thus the non-zero term a_k occurs at the position k which corresponds to the pattern θ_k with highest probability λ_k . One of the functions of non-linear neural networks in modelling the associative memory model is the so called recall of stored patterns. One may interpret the non-zero solution as the recall the k th pattern. According to (4.64), for fixed p patterns, α_c is determined by

the largest λ_k . Thus the patterns with high probabilities will be recalled first. As the parameter α further increases above the critical value α_c there are other nonzero solutions. For example, if $\lambda_1, \dots, \lambda_p$ and α satisfy the condition

$$\max\left(\frac{1}{\lambda_k}, \frac{1}{\lambda_l}\right) < \alpha \leq \min_{m: m \neq k, m \neq l} \left(\frac{1}{\lambda_m}\right) \quad (4.65)$$

Then the non-zero solution will have the form of

$$(0, \dots, 0, c, 0, \dots, 0, \dots, 0, d, 0, \dots, 0)$$

where $c, d \in (-1, 1)$. This solution corresponds to the mixture of two patterns θ_k and θ_l which have higher probabilities than the rest of the patterns.

Chapter 5

Asymptotic Behavior of Random Gibbs Measures

In this chapter we consider the asymptotic behavior of the random Gibbs measures given by

$$P_{n,\xi}(dx_1, \dots, dx_n) = \frac{1}{Z_{n,\xi}} \exp\left\{\frac{\beta}{2n} \sum_{i,j=1}^n c(\xi_i, \xi_j) x_i x_j\right\} \prod_{i=1}^n \mu(dx_i), \quad (5.1)$$

where β is a positive parameter; $c(x, y)$ is a symmetric continuous function defined on $R^q \times R^q$. μ is a probability measure on R ; $\xi = (\xi_1, \xi_2, \dots)$ consists of a sequence of i.i.d. $\{\theta_1, \dots, \theta_p\} \subset R^q$ -valued random variables with distribution

$$\lambda = \sum_{k=1}^p \lambda_k \delta_{\theta_k}, \quad (5.2)$$

where $\lambda_1, \dots, \lambda_p$ are strictly positive constants satisfying $\sum_{k=1}^p \lambda_k = 1$; and $Z_{n,\xi}$ is the normalizing constant

$$Z_{n,\xi} = \int_{R^n} \exp\left\{\frac{\beta}{2n} \sum_{i,j=1}^n c(\xi_i, \xi_j) x_i x_j\right\} \prod_{i=1}^n \mu(dx_i). \quad (5.3)$$

The random probability measure defined by (5.1) is related to the equilibrium behavior of a special class of stochastic models introduced in chapter 3, i.e., the gradient type models given by the following systems of Itô equations in R :

$$dx_i^{(n)}(t) = -U'(x_i^{(n)}(t))dt + \frac{1}{n} \sum_{j=1}^n c(\xi_i, \xi_j) x_j^{(n)}(t)dt + \sigma dw_i(t), \quad (5.4)$$

$$i = 1, \dots, n,$$

where $w_1, \dots, w_n$ are n independent 1-dimensional Wiener processes, $U(x)$ is a differentiable function defined on R such that

$$\int_R \exp\left\{-\frac{2}{\sigma^2}U(x)\right\}dx < \infty \quad (5.5)$$

and σ is a positive constant. The equations (5.4) describe the motion of n particles $x_1^{(n)}(t), \dots, x_n^{(n)}(t)$ in R under the influence of the potential U , the mean field interaction with random coefficients $c(\xi_i, \xi_j)$, and the internal white noise fluctuations with intensity σ^2 . For each n and ξ , the Markov diffusion system (5.4) has a unique invariant probability measure $P_{n,\xi}(dx_1, \dots, dx_n)$ which has the form of (5.1) with

$$\beta = \frac{2}{\sigma^2} \quad (5.6)$$

$$\mu(dx) = \frac{1}{Z} \exp\left\{-\frac{2}{\sigma^2}U(x)\right\}dx \quad (5.7)$$

where Z denotes the normalizing constant.

$$Z = \int_R \exp\left\{-\frac{2}{\sigma^2}U(x)\right\}dx \quad (5.8)$$

Among the asymptotic behavior of the random Gibbs measures given by (5.1) we are mainly interested in the following two problems:

1. to compute the free energy per particle in the thermodynamic limit $n \rightarrow \infty$

$$f(\beta) = -\frac{1}{\beta} \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,\xi}$$

2. to determine whether $(P_{n,\xi})_{n=1}^\infty$ are convergent to a certain probability measure on $R^{\otimes \infty}$ as $n \rightarrow \infty$ and, if so, what is the limit probability measure.

In Section 5.1 we state the main results. In section 5.2 we introduce the notion of a.s. large deviations and provide an asymptotic method to study the a.s. large deviations. In section 5.3 we prove a Varadhan type variational formula for evaluating certain integrals. In Section 5.4 and Section 5.5 we prove the main results.

5.1 Statement of the results

As in chapter 3, $(\Xi, \mathcal{B}(\Xi), \varphi)$ still denotes the infinite product probability space of $(\Theta, \mathcal{B}(\Theta), \lambda)$. For $(m_1, \dots, m_p) \in R^p$ let

$$\Phi_\beta(m_1, \dots, m_p) = \frac{\beta}{2} \sum_{k,l=1}^p c(\theta_k, \theta_l) \lambda_k \lambda_l m_k m_l. \quad (5.9)$$

The first main result in this chapter can be stated as following.

Theorem 5.1 *Suppose that μ satisfies the condition*

$$\int_R \exp(kx^2) \mu(dx) < \infty, \quad \text{for all } k \in R. \quad (5.10)$$

Then for almost all $\xi \in \Xi$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,\xi} = \sup_{(m_1, \dots, m_p) \in R^p} [\Phi_\beta(m_1, \dots, m_p) - \sum_{k=1}^p \lambda_k L^*(m_k)], \quad (5.11)$$

where $L^(x)$ is the Legendre transformation of $L(t)$*

$$L^*(x) = \sup_{t \in R} \{tx - L(t)\}, \quad x \in R \quad (5.12)$$

and $L(t)$ is the logarithmic moment generating function of $\mu(dx)$

$$L(t) = \log \int_R \exp(tx) \mu(dx), \quad x \in R. \quad (5.13)$$

Next, for $\nu \in \mathbf{M}_1(R)$, we extend $P_{n,\xi}$ to the probability $\hat{P}_{n,\xi}$ on $R^{\otimes \infty}$ by defining

$$\hat{P}_{n,\xi}(dx_1, dx_2, \dots) = P_{n,\xi}(dx_1, \dots, dx_n) \prod_{i=n+1}^{\infty} \nu(dx_i). \quad (5.14)$$

Let $F(m)$ be the function on R^p defined by

$$F(m) = \frac{1}{2\beta}(m, Cm) - \sum_{k=1}^p \lambda_k \log \int_R \exp(x[Cm]_k) \mu(dx) \quad (5.15)$$

where $C = \{c(\theta_k, \theta_l)\}$ is the $p \times p$ matrix and the symbol $[Cm]_k$ denotes the k th component of the vector Cm . Let K_0 be the subset of R^p defined by

$$K_0 = \{m \in R^p : F(m) \text{ attains its minimum at } m\} \quad (5.16)$$

Then the second main result in this chapter can be stated as following.

Theorem 5.2 Suppose that $C = \{c(\theta_k, \theta_l)\}$ is a $p \times p$ strictly positive definite matrix and μ satisfies the condition (5.10). Then

1. for almost all $\xi \in \Xi$, $(\tilde{P}_{n,\xi})_{n=1}^\infty$ is relatively compact in $M_1(\mathbb{R}^{\otimes \infty})$, and
2. any limit point P_ξ of $(\tilde{P}_{n,\xi})_{n=1}^\infty$ has the representation

$$P_\xi(dx_1, dx_2, \dots) = \int_{R^p} Q(dm) \prod_{k=1}^p \prod_{\substack{i=1 \\ \xi_i = \theta_k}}^\infty \left(\frac{\exp\{x_i[Cm]_k\} \mu(dx_i)}{\int_R \exp\{x_i[Cm]_k\} \mu(dx_i)} \right) \quad (5.17)$$

where $Q(dm)$ is a probability measure in $M_1(R^p)$ which concentrates on the set K_0 .

5.2 φ -a.s. large deviations

Throughout this section let the following objects be given:

- $(\Xi, \mathcal{A}, \varphi)$ a probability space;
- (X, ρ) a complete separable metric space with Borel σ -algebra $\mathcal{B}(X)$;
- $(P_{n,\xi})$ a sequence of probability measures on X depending on $\xi \in \Xi$;
- (a_n) a sequences of positive numbers tending to infinity;
- I a function $X \rightarrow [0, \infty]$.

Definition 5.1 $(X, P_{n,\xi}, a_n)$ is said to form a φ -a.s. large deviation system with rate function I , if the following conditions are satisfied:

- (i) for each $n \geq 1$ and $A \in \mathcal{B}(X)$, $\xi \mapsto P_{n,\xi}(A)$ is $\mathcal{A}$ -measurable;
- (ii) the set $\Phi(s) = \{x \in X : I(x) \leq s\}$ is compact for each $s \geq 0$;
- (iii) there exists a φ -null set $\mathcal{N}$ such that for $\xi \notin \mathcal{N}$

$$\begin{aligned} - \inf_{x \in A^*} I(x) &\leq \liminf_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(A) \leq \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(A) \leq - \inf_{x \in A} I(x) \end{aligned} \quad (5.18)$$

for every Borel subset A of X , where A° and $\bar{A}$ denotes the interior and the closure of the set A , respectively.

Condition (iii) is equivalent to the following two conditions

(iii-a) there exists a φ -null set $\mathcal{N}_f$ such that for $\xi \notin \mathcal{N}_f$

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(F) \leq - \inf_{x \in F} I(x) \quad (5.19)$$

for every closed subset F of X ;

(iii-b) there exists a φ -null set $\mathcal{N}_g$ such that for $\xi \notin \mathcal{N}_g$

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(G) \geq - \inf_{x \in G} I(x) \quad (5.20)$$

for every open subset G of X .

The conditions (iii-a) and (iii-b) in the above definition are essentially the type of conditions used in Varadhan large deviation theory [54]. On the other hand, Freidlin and Wentzell [19] used different but equivalent conditions to describe (iii-a) and (iii-b) in their large deviation theory. We now give this equivalent description in our setting.

Proposition 5.1 *Under condition (i) and (ii), (iii-a) is equivalent to the condition: for any $\delta > 0$, any $r > 0$ and any $s > 0$ there exists an φ -null set $\mathcal{N}_f$ and a $n_0 > 0$ such that for $\omega \notin \mathcal{N}_f$*

$$P_{n,\xi}\{y : \rho(y, \Phi(s)) \geq \delta\} \leq \exp\{-a_n(s - r)\} \quad (5.21)$$

for $n \geq n_0$.

Proof: Suppose (iii-a) holds. For any $\delta > 0$, and $s > 0$ let

$$F = \{y : \rho(y, \Phi(s)) \geq \delta\} \quad (5.22)$$

Then F is a closed subset of X . For any $x \in F$, we have $x \notin \Phi(s)$, i.e. $I(x) > s$. Therefore

$$\inf_{x \in F} I(x) \geq s \quad (5.23)$$

By (5.19) there exists a φ -null set $\mathcal{N}_f$ such that for $\xi \notin \mathcal{N}_f$

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(F) \leq - \inf_{x \in F} I(x) \leq -s \quad (5.24)$$

which means that for every $r > 0$ and sufficiently large n

$$\frac{1}{a_n} \log P_{n,\xi}(F) \leq -s + r, \quad \varphi\text{-a.s.} \quad (5.25)$$

i.e. (5.21) is satisfied.

Conversely, suppose (5.21) holds. For any closed subset F of X , choose an arbitrary $r > 0$ and put

$$s = \inf_{y \in F} I(y) - r \quad (5.26)$$

Then $x \in \Phi(s)$ implies $I(x) \leq \inf_{y \in F} I(y) - r$ which implies $x \notin F$. So

$$F \cap \Phi(s) = \emptyset. \quad (5.27)$$

Since F is closed and $\Phi(s)$ is compact, we must have

$$\delta = \rho(F, \Phi(s)) > 0 \quad (5.28)$$

By (5.21) there exists a φ -null set $\mathcal{N}_f$ such that for $\xi \notin \mathcal{N}_f$

$$\begin{aligned} P_{n,\xi}(F) &\leq P_{n,\xi}\{y : \rho(y, \Phi(s)) \geq \delta\} \\ &\leq \exp\{-a_n(s - r)\} \\ &= \exp\{-a_n[\inf_{x \in F} I(x) - 2r]\} \end{aligned} \quad (5.29)$$

for sufficiently large n . Therefore

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(F) \leq - \inf_{y \in F} I(y) + 2r, \quad \varphi\text{-a.s.} \quad (5.30)$$

which implies (5.19), since $r > 0$ is arbitrary. $\square$

Proposition 5.2 Under condition (i) and (ii), (iii-b) is equivalent to the condition: for any $\delta > 0$, any $r > 0$ and any $x \in X$ there exists a φ -null set N_δ and a $n_0 > 0$ such that for $\xi \notin N_\delta$

$$P_{n,\xi}\{y : \rho(x, y) < \delta\} \geq \exp\{-a_n(I(x) + r)\} \quad (5.31)$$

for $n \geq n_0$.

Proof: Suppose (iii-b) holds. For any $\delta > 0$, and $x \in X$ let

$$G = \{y : \rho(y, x) < \delta\} \quad (5.32)$$

Then G is an open subset of X . Since $x \in G$, we have

$$\inf_{y \in G} I(y) \leq I(x) \quad (5.33)$$

By (5.20) there exists a φ -null set N_δ such that for $\xi \notin N_\delta$

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(G) \geq -\inf_{y \in G} I(y) \geq -I(x) \quad (5.34)$$

which means that for every $r > 0$ and sufficiently large n

$$\frac{1}{a_n} \log P_{n,\xi}(G) \geq -I(x) - r, \quad \varphi\text{-a.s.} \quad (5.35)$$

i.e. (5.31) is satisfied.

Conversely, suppose (5.31) holds. For any non-empty open subset G of X , any $r > 0$ choose an $x \in G$ such that

$$I(x) \leq \inf_{y \in G} I(y) + r \quad (5.36)$$

and choose a $\delta > 0$ such that

$$\{y, \rho(y, x) < \delta\} \subset G. \quad (5.37)$$

By (5.31) there exists a φ -null set $\mathcal{N}_\delta$ such that for $\xi \notin \mathcal{N}_\delta$

$$\begin{aligned} P_{n,\xi}(G) &\geq P_{n,\xi}\{y : \rho(y, x) < \delta\} \\ &\geq \exp\{-a_n(I(x) + r)\} \\ &\geq \exp\{-a_n(\inf_{y \in G} I(y) + 2r)\} \end{aligned} \quad (5.38)$$

for sufficiently large n . Therefore

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(G) \geq -\inf_{y \in F} I(y) - 2r, \quad \varphi - a.s. \quad (5.39)$$

which implies (5.20), since $r > 0$ is arbitrary. $\square$

We now present an asymptotic method to obtain the φ -a.s. large deviations on R^d . It essentially uses the same ideas as in [19].

Theorem 5.3 *Suppose that the following conditions are satisfied:*

1. *for each $n \geq 1$ and $A \in \mathcal{B}(X)$, $P_{n,\xi}(A)$ is $\mathcal{A}$ -measurable with respect to ξ ;*
2. *for each $x \in R^d$ the limit*

$$\begin{aligned} H(x) &= \lim_{n \rightarrow \infty} \frac{1}{a_n} H_{n,\xi}(x) \\ &= \lim_{n \rightarrow \infty} \frac{1}{a_n} \log \int_{R^d} \exp(a_n(y, x)) P_{n,\xi}(dy) \end{aligned} \quad (5.40)$$

exists φ -a.s. such that $H(x)$ does not depend on ξ and is finite.

3. *the function H is differentiable for every $x \in R^d$.*

Then $(R^d, P_{n,\xi}, a_n)$ forms a φ -a.s. large deviation system. The corresponding rate functional $L(x)$ is given by

$$L(x) = \sup_{y \in R^d} \{(x, y) - H(y)\}, \quad x \in R^d \quad (5.41)$$

Proof: let $P_{n,\xi}$ be a sequence of probability measures on R^d which depends on $\xi \in \Xi$. Define

$$H_{n,\xi}(\alpha) = \log \int_{R^d} \exp\{(\alpha, x)\} P_{n,\xi}(dx), \quad \alpha \in R^d. \quad (5.42)$$

Then, for each $\xi \in \Xi$, by means of the Hölder inequality and the Fatou's lemma $H_{n,\xi}(\alpha)$ is a convex and lower semicontinuous function with values in $(-\infty, +\infty]$. It follows that the function $H_\xi(\alpha)$ is also a convex and lower semicontinuous function with $H_\xi(0) = 0$. Since $H_\xi(\alpha)$ tends to infinity as α tends to infinity the set $\Phi(s) = \{\beta : L(\beta) \leq s\}$ is compact in X for each $s \geq 0$. Hence in order to prove the theorem we need to show that the conditions (iii-a) and (iii-b) in definition 5.1 are satisfied. We now divide the proof into the following two propositions.

Proposition 5.3 (*Upper bound*) Suppose that for each $\alpha \in R^d$, the limit

$$H(\alpha) = \lim_{n \rightarrow \infty} \frac{1}{a_n} H_{n,\xi}(a_n \alpha) \quad (5.43)$$

exists φ -a.s. such that $H(x)$ does not depend on ξ and is finite. Let $L(\beta)$ be the Legendre transformation of $H(\alpha)$

$$L(\beta) = \sup_{\alpha \in R^d} \{(\alpha, \beta) - H(\alpha)\} \quad (5.44)$$

and let $\Phi(s)$ be the level set

$$\Phi(s) = \{\beta : L(\beta) \leq s\} \quad (5.45)$$

Then the upper bound condition (iii-a) in definition 5.1 is satisfied.

Proof: According to proposition 5.1 it is equivalent to show that (5.21) is satisfied. By definition, the compact set $\Phi(s)$ can be written as

$$\Phi(s) = \bigcap_{\alpha \in R^d} \{\beta : (\alpha, \beta) - H(\alpha) \leq s\}. \quad (5.46)$$

Denote by $\Phi_\delta(s)$ the δ -neighbourhood of $\Phi(s)$ and let $\partial\Phi_\delta(s)$ be its boundary set:

$$\partial\Phi_\delta(s) = \{y : \rho(y, \Phi(s)) = \delta\}. \quad (5.47)$$

Let $G_\alpha, \alpha \in R^d$, be the open sets defined by

$$G_\alpha = \{y \in R^d : (\alpha, y) - H(\alpha) > s\}. \quad (5.48)$$

For any $y \in \partial\Phi_\delta(s)$, clearly $y \in \Phi^C(s)$. So there exists $\alpha \in R^d$ such that

$$(\alpha, y) - H(\alpha) > s. \quad (5.49)$$

Hence the open set $G = \bigcup_{\alpha \in R^d} G_\alpha$ covers the set $\partial\Phi_\delta(s)$. Since $\partial\Phi_\delta(s)$ is compact, we can choose finite points $\alpha_1, \dots, \alpha_m$ such that

$$\partial\Phi_\delta(s) \subset \bigcup_{i=1}^m G_{\alpha_i}, \quad (5.50)$$

and so

$$\bigcap_{i=1}^m G_{\alpha_i}^c \subset (\partial\Phi_\delta(s))^c \quad (5.51)$$

On the other hand

$$\bigcap_{i=1}^m G_{\alpha_i}^c \supset \Phi_\delta(s) \quad (5.52)$$

(5.51) and (5.52) together implies

$$\Phi(s) \subset \bigcap_{i=1}^m \{y : (\alpha_i, y) - H(\alpha_i) \leq s\} \subset \Phi_\delta(s). \quad (5.53)$$

For each $\xi \in \Xi$, by Chebyshev exponential inequality,

$$\begin{aligned} & P_{n,\xi}\{y : \rho(y, \Phi(s)) \geq \delta\} \\ & \leq P_{n,\xi}(\bigcup_{i=1}^m \{y : (\alpha_i, y) - H(\alpha_i) > s\}) \\ & \leq \sum_{i=1}^m P_{n,\xi}\{y : (\alpha_i, y) - H(\alpha_i) > s\} \\ & \leq \sum_{i=1}^m \int_{R^d} \exp\{a_n[(\alpha_i, y) - H(\alpha_i)]\} \exp\{-a_n s\} P_{n,\xi}(dy) \\ & = \sum_{i=1}^m \exp\{a_n[\frac{1}{a_n} H_{n,\xi}(a_n \alpha_i) - H(\alpha_i)]\} \exp\{-a_n s\} \end{aligned} \quad (5.54)$$

By assumption (5.43), for each $i = 1, \dots, m$, let $\mathcal{N}_i$ be the φ -null set such that for $\xi \notin \mathcal{N}_i$ the limit

$$H(\alpha_i) = \lim_{n \rightarrow \infty} \frac{1}{a_n} H_{n,\xi}(a_n \alpha_i), \quad \varphi\text{-a.s.} \quad (5.55)$$

exists and is finite. Denote $\mathcal{N} = \cup_{i=1}^m \mathcal{N}_i$. Then $\mathcal{N}$ is still a φ -null set and for $\xi \notin \mathcal{N}$ and sufficiently large n

$$P_{n,\xi}\{y : \rho(y, \Phi(s)) \geq \delta\} \leq \exp\{-a_n(s - r)\}. \quad (5.56)$$

Hence (5.21) is satisfied. $\square$

Proposition 5.4 (Lower bound) *Suppose that for each $\alpha \in R^d$ the limit*

$$H(\alpha) = \lim_{n \rightarrow \infty} \frac{1}{a_n} H_{n,\xi}(a_n \alpha) \quad (5.57)$$

exists φ -a.s. such that $H(x)$ does not depend on ξ and is finite. Furthermore, suppose that $H(\alpha)$ is differentiable at every point $\alpha \in R^d$. Let $L(\beta)$ be the Legendre transformation of $H(\alpha)$:

$$L(\beta) = \sup_{\alpha \in R^d} \{(\alpha, \beta) - H(\alpha)\} \quad (5.58)$$

and let $\Phi(s)$ be the level set

$$\Phi(s) = \{\beta : L(\beta) \leq s\} \quad (5.59)$$

Then the lower bound condition (iii-b) in definition 5.1 is satisfied.

Proof: According to proposition 5.2 it is equivalent to show that (5.31) is satisfied. For that it suffices to prove that (5.31) holds for $x \in \text{ri}(\text{dom}(L))$, where $\text{dom}(L)$ is the effective domain of the convex function L and $\text{ri}(\text{dom}(L))$ is the relative interior of the convex set $\text{dom}(L)$ (see [20] and [40]). So let $x \in \text{ri}(\text{dom}(L))$. Since H is differentiable, L is strictly convex in the sense that there exists $\alpha_0 \neq 0$ such that

$$L(\alpha) > L(x) + (\alpha_0, \alpha - x), \quad \text{for } \alpha \neq x, \quad (5.60)$$

i.e.,

$$(\alpha, \alpha_0) - L(\alpha) < (\alpha_0, x) - L(x), \text{ for } \alpha \neq x. \quad (5.61)$$

Hence we have

$$H(\alpha_0) = \sup_{\alpha \in R^d} \{(\alpha, \alpha_0) - L(\alpha)\} = (\alpha_0, x) - L(x) \quad (5.62)$$

Since $H(\alpha_0)$ is finite, there exists a φ -null set $\mathcal{N}$ such that for $\xi \notin \mathcal{N}$ $H_{n,\xi}(a_n \alpha_0)$ is finite for sufficiently large n . For such ξ and n we consider the probability measure P_{n,ξ,α_0} defined by the relation

$$P_{n,\xi,\alpha_0}(A) = \frac{1}{Z_{n,\xi,\alpha_0}} \int_A \exp\{a_n(\alpha_0, y)\} P_{n,\xi}(dy) \quad (5.63)$$

where Z_{n,ξ,α_0} is a normalizing factor

$$\begin{aligned} Z_{n,\xi,\alpha_0} &= \int_{R^d} \exp\{a_n(\alpha_0, y)\} P_{n,\xi}(dy) \\ &= \exp\{H_{n,\xi}(a_n \alpha_0)\}. \end{aligned} \quad (5.64)$$

Since $P_{n,\xi}$ and P_{n,ξ,α_0} are mutually absolutely continuous we have

$$P_{n,\xi}\{y : \rho(y, x) < \delta\} = \int_{\{y : \rho(y, x) < \delta\}} \exp\{-a_n(\alpha_0, y) + H_{n,\xi}(a_n \alpha_0)\} P_{n,\xi,\alpha_0}(dy) \quad (5.65)$$

For any $\delta > 0$ and $r > 0$, let

$$\delta' = \min(\delta, \frac{r}{|\alpha_0|}) \quad (5.66)$$

Using (5.65) we have

$$\begin{aligned} P_{n,\xi}\{y : \rho(y, x) < \delta\} &\geq P_{n,\xi}\{y : \rho(y, x) < \delta'\} \\ &\geq P_{n,\xi,\alpha_0}\{y : \rho(y, x) < \delta'\} \\ &\quad \exp\{-a_n[(\alpha_0, x) - \frac{1}{a_n} H_{n,\xi}(a_n \alpha_0)]\} \\ &\quad \exp\{-a_n r\} \end{aligned} \quad (5.67)$$

By the assumption (5.57) there exists a φ -null set $\mathcal{N}$ such that for $\xi \notin \mathcal{N}$ and sufficiently large n the exponential term above tends to

$$\exp\{-a_n(L(x) + \frac{r}{3})\}.$$

Hence to finish the proof it remains to show that

$$P_{n,\xi,\alpha_0}\{y : \rho(y, x) < \delta'\} \rightarrow 1 \quad \varphi\text{-a.s.} \quad (5.68)$$

as $n \rightarrow \infty$.

For this we apply proposition 5.3 to the family of measures P_{n,ξ,α_0} . Since

$$\begin{aligned} H_{n,\xi,\alpha_0}(\alpha) &= \log \int_{R^d} \exp\{(\alpha, y)\} P_{n,\xi,\alpha_0}(dy) \\ &= \log \int_{R^d} \exp\{(\alpha, y)\} \exp\{a_n(\alpha_0, y) - H_{n,\xi}(a_n\alpha_0)\} P_{n,\xi}(dy) \\ &= \log \int_{R^d} \exp\{(\alpha + a_n\alpha_0, y) - H_{n,\xi}(a_n\alpha_0)\} P_{n,\xi}(dy) \\ &= H_{n,\xi}(\alpha + a_n\alpha_0) - H_{n,\xi}(a_n\alpha_0); \end{aligned} \quad (5.69)$$

the limit

$$\begin{aligned} H_{\alpha_0}(\alpha) &= \lim_{n \rightarrow \infty} \frac{1}{a_n} H_{n,\xi,\alpha_0}(a_n\alpha) \\ &= \lim_{n \rightarrow \infty} \frac{1}{a_n} [H_{n,\xi}(a_n(\alpha + \alpha_0)) - H_{n,\xi}(a_n\alpha_0)] \\ &= H(\alpha + \alpha_0) - H(\alpha_0) \end{aligned} \quad (5.70)$$

exists φ -a.s. and is finite. Let $L_{\alpha_0}(\beta)$ be the Legendre transformation of $H_{\alpha_0}(\alpha)$ defined by

$$\begin{aligned} L_{\alpha_0}(\beta) &= \sup_{\alpha \in R^d} \{(\alpha, \beta) - H_{\alpha_0}(\alpha)\} \\ &= \sup_{\alpha \in R^d} \{(\alpha, \beta) - H(\alpha + \alpha_0) + H(\alpha_0)\} \\ &= \sup_{\alpha \in R^d} \{(\alpha + \alpha_0, \beta) - H(\alpha + \alpha_0)\} - [(\alpha_0, \beta) - H(\alpha_0)] \\ &= L(\beta) - [(\alpha_0, \beta) - H(\alpha_0)] \end{aligned} \quad (5.71)$$

By (5.58) and (5.60) the function $L_{\alpha_0}(\beta)$ vanishes at $\beta = x$ and is strictly positive for all $\beta \neq x$. Hence

$$r_1 = \min\{L_{\alpha_0}(\beta) : \rho(\beta, x) \geq \frac{\delta'}{2}\} > 0 \quad (5.72)$$

Choose positive $r_0 < r_1$ and $s \in (r_0, r_1)$. Let $y \in R^d$ be such that

$$\rho(y, x) \geq \delta', \quad (5.73)$$

then by triangle inequality we may have $\rho(y, z) \geq \frac{\delta'}{2}$ or $\rho(z, x) \geq \frac{\delta'}{2}$ for all $z \in \Phi_{\alpha_0}(s)$.

But for any $z \in \Phi_{\alpha_0}(s)$

$$L_{\alpha_0}(z) \leq s < r_1. \quad (5.74)$$

This implies that

$$\rho(z, x) < \frac{\delta'}{2} \quad (5.75)$$

so we must have

$$\rho(y, z) \geq \frac{\delta'}{2} \text{ for all } z \in \Phi_{\alpha_0}(s). \quad (5.76)$$

By proposition 5.3 with $\frac{\delta'}{2}$ instead of δ and sufficiently large n

$$\begin{aligned} P_{n,\xi,\alpha_0}\{y : \rho(y, x) \geq \delta'\} &\leq P_{n,\xi,\alpha_0}\{y : \rho(y, \Phi_{\alpha_0}(s)) \geq \frac{\delta'}{2}\} \\ &\leq \exp\{-a_n(s - r_0)\}, \end{aligned} \quad (5.77)$$

which converges to zero as $n \rightarrow \infty$. □

5.3 Varadhan type theorem

In the standard large deviation theory one has the useful Varadhan theorem which is often used to evaluate certain integrations over function spaces. It is natural that similar theorem should also hold in the p -a.s. large deviation setting. The purpose of this section is to prove such theorem.

Let $(P_{n,\xi})_{n=1}^{\infty}$ be a sequence of probability measures defined on the metric space (X, ρ) with $\xi \in \Xi$ and let $\Phi(x)$ be a continuous function defined on X .

Theorem 5.4 Suppose that $(X, P_{n,\xi}, a_n)$ forms a φ -a.s. large deviation system with rate function $I(x)$, and Φ satisfies the condition

$$\lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \int_{\{\Phi(x) \geq L\}} \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) = -\infty, \quad \varphi\text{-a.s.} \quad (5.78)$$

Then we have

$$\varphi\{\xi : \lim_{n \rightarrow \infty} \frac{1}{a_n} \log \int_X \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) = \sup_{x \in X} [\Phi(x) - I(x)]\} = 1. \quad (5.79)$$

Proof: (Upper bound) We show that

$$\varphi\{\xi : \lim_{n \rightarrow \infty} \frac{1}{a_n} \log \int_X \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) \leq \sup_{x \in X} [\Phi(x) - I(x)]\} = 1. \quad (5.80)$$

First let us assume that Φ is bounded. Denote

$$l = \sup_{x \in X} [\Phi(x) - I(x)] \quad (5.81)$$

Let $x_0 \in X$ be an arbitrary point such that $I(x_0) < \infty$ (such x_0 exists because $I(x)$ is not always $+\infty$). Since $\Phi(x)$ is continuous and $I(x)$ is lower semicontinuous, for $\epsilon > 0$, there exists an open ball B_{x_0} of x_0 such that

$$\Phi(y) \leq \Phi(x_0) + \epsilon \quad (5.82)$$

and

$$I(y) \geq I(x_0) - \epsilon \quad (5.83)$$

for all $y \in B_{x_0}$. Let G_{x_0} be another open ball of x_0 such that $\bar{G}_{x_0} = F_{x_0} \subset B_{x_0}$, and let $\mathcal{N}$ be the φ -null set such that for all $\xi \notin \mathcal{N}$ we have

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(F_{x_0}) \leq - \inf_{x \in F_{x_0}} I(x) \quad (5.84)$$

Then for $\xi \notin \mathcal{N}$

$$\begin{aligned} \int_{G_{x_0}} \exp\{a_n \Phi(y)\} P_{n,\xi}(dy) &\leq \exp\{a_n \sup_{y \in G_{x_0}} \Phi(y)\} P_{n,\xi}(G_{x_0}) \\ &\leq \exp\{a_n [\Phi(x_0) + \epsilon]\} P_{n,\xi}(F_{x_0}) \end{aligned} \quad (5.85)$$

Therefore, φ -a.s.

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \int_{G_{x_0}} \exp\{a_n \Phi(y)\} P_{n,\xi}(dy) &\leq \Phi(x_0) + \epsilon - \inf_{x \in F_{x_0}} I(x) \\ &\leq \Phi(x_0) - I(x_0) + 2\epsilon \\ &\leq l + 2\epsilon. \end{aligned} \quad (5.86)$$

Now let

$$K = \{x : I(x) \leq \|\Phi\| + |l|\}. \quad (5.87)$$

Then for a given $\epsilon > 0$, any $x_0 \in K$, there exists an open ball G_{x_0} of x_0 such that (5.86) holds. But K is compact in X , there exists a finite union of $G = \cup_{x \in K} G_x$, which covers K . Then φ -a.s. we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \int_G \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) \\ \leq \max_{x=x_1, \dots, x_m} \limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \int_{G_x} \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) \\ \leq l + 2\epsilon. \end{aligned} \quad (5.88)$$

Let $F = X - G$. Then for any $x \in F$,

$$I(x) > \|\Phi\| + |l| \quad (5.89)$$

and

$$\int_F \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) \leq \exp\{a_n \|\Phi\|\} P_{n,\xi}(F). \quad (5.90)$$

It follows that, φ -a.s.,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \int_F \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) &\leq \|\Phi\| - (\|\Phi\| + |l|) \\ &= -|l| \end{aligned} \quad (5.91)$$

(5.88) and (5.91) together yields

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \int_X \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) &\leq l + 2\epsilon - |l| \\ &\leq l + 2\epsilon, \quad \text{p-a.s.} \end{aligned} \quad (5.92)$$

Since $\epsilon > 0$ is arbitrary, we obtain

$$\text{p}\{\xi : \limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \int_X \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) \leq l\} = 1 \quad (5.93)$$

when Φ is bounded and continuous.

For general Φ , let $\Phi_L = \Phi \wedge L$ for $L > 0$. Denote, for Borel subset A of X ,

$$J_{n,\xi}(A) = \frac{1}{a_n} \log \int_A \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) \quad (5.94)$$

Then

$$\begin{aligned} &\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \int_X \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) \\ &= \max\{\limsup_{n \rightarrow \infty} J_{n,\xi}(\{\Phi(x) < L\}), \limsup_{n \rightarrow \infty} J_{n,\xi}(\{\Phi(x) \geq L\})\} \\ &\leq \max\{\sup_{x \in X} [\Phi_L(x) - I(x)], \limsup_{n \rightarrow \infty} J_{n,\xi}(\{\Phi(x) \geq L\})\} \end{aligned} \quad (5.95)$$

Now let $L \rightarrow \infty$ and using the assumption (5.78) we obtain the upper bound (5.80).

(Lower bound) We show that

$$\text{p}\{\xi : \lim_{n \rightarrow \infty} \frac{1}{a_n} \log \int_X \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) \geq \sup_{x \in X} [\Phi(x) - I(x)]\} = 1. \quad (5.96)$$

Let $x_0 \in X$ be such that at least one of $\Phi(x)$ and $I(x)$ is finite. For such $x_0 \in X$ and for any $\epsilon > 0$ there exists a neighbourhood N_{x_0} of x_0 in X such that for $y \in N_{x_0}$,

$$\Phi(y) \geq \Phi(x_0) - \epsilon. \quad (5.97)$$

Then

$$\begin{aligned} \int_X \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) &\geq \int_{N_{x_0}} \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) \\ &\geq \exp\{a_n [\Phi(x_0) - \epsilon]\} P_{n,\xi}(N_{x_0}). \end{aligned} \quad (5.98)$$

Let $\mathcal{N}$ be the ρ -null set such that for $\xi \notin \mathcal{N}$

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \log P_{n,\xi}(N_{x_0}) \geq - \inf_{x \in N_{x_0}} I(x).$$

Then for $\xi \notin \mathcal{N}$

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{a_n} \int_X \exp\{a_n \Phi(x)\} P_{n,\xi}(dx) &\geq \Phi(x_0) - \epsilon - \inf_{x \in N_{x_0}} I(x) \\ &\geq \Phi(x_0) - I(x_0) - \epsilon. \end{aligned} \quad (5.99)$$

Since $\epsilon > 0$ is arbitrary, we obtain the lower bound (5.96). $\square$

5.4 Proof of theorem 5.1

We are now ready to prove the theorem 5.1. For each $n \geq 1$ and $\xi \in \Xi$, define the map $m_\xi^{(n)} : R^n \rightarrow R^p$ as following

$$m_\xi^{(n)}(x) = \left(\frac{1}{n} \sum_{\substack{i=1 \\ \xi_i = \theta_1}}^n x_i, \dots, \frac{1}{n} \sum_{\substack{i=1 \\ \xi_i = \theta_p}}^n x_i \right)$$

and let $Q_{n,\xi}$ be the image measure of $\prod_{i=1}^n \mu(dx_i)$ under this map. Then (5.1) can be written as

$$\begin{aligned} Z_{n,\xi} &= \int_{R^n} \exp\left\{ \frac{\beta}{2n} \sum_{i,j=1}^n c(\xi_i, \xi_j) x_i x_j \right\} \prod_{i=1}^n \mu(dx_i) \\ &= \int_{R^n} \exp\left\{ \frac{\beta}{2} n \sum_{k,l=1}^p c(\theta_k, \theta_l) \left(\frac{1}{n} \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n x_i \right) \left(\frac{1}{n} \sum_{\substack{j=1 \\ \xi_j = \theta_l}}^n x_j \right) \right\} \prod_{i=1}^n \mu(dx_i) \\ &= \int_{R^p} \exp\left\{ \frac{\beta}{2} n \sum_{k,l=1}^p c(\theta_k, \theta_l) m_k m_l \right\} Q_{n,\xi}(dm) \\ &= \int_{R^p} \exp\{n F_\beta(m)\} Q_{n,\xi}(dm) \end{aligned} \quad (5.100)$$

where

$$F_\beta(m) = \frac{\beta}{2} \sum_{k,l=1}^p c(\theta_k, \theta_l) m_k m_l \quad (5.101)$$

We first show that $(R^p, Q_{n,\xi}, n)$ forms a φ -a.s. large deviation system with rate function $I(m) = \sum_{k=1}^p \lambda_k L^*(\frac{m_k}{\lambda_k})$. We then use the Varadhan type theorem 5.4 to derive the final result.

Proposition 5.5 *Let the condition (5.10) hold. Then $(R^p, Q_{n,\xi}, n)$ forms a φ -a.s. large deviation system with rate function*

$$I(m) = \sum_{k=1}^p \lambda_k L^*\left(\frac{m_k}{\lambda_k}\right), \quad m = (m_1, \dots, m_p) \in R^p \quad (5.102)$$

where $L^*(x)$ is the Legendre transformation of $L(t)$

$$L^*(x) = \sup_{t \in R} \{tx - L(t)\}, \quad x \in R \quad (5.103)$$

and

$$L(t) = \log \int_R \exp(tx) \mu(dx), \quad t \in R \quad (5.104)$$

Proof: We use theorem 5.3.

1. For each $n \geq 1$ and $A \in \mathcal{B}(X)$, $P_{n,\xi}(A)$ is certainly $\mathcal{A}$ -measurable with respect to the random variable ξ .

2. For $y \in R^p$ we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} H_{n,\xi}(ny) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{R^p} \exp\{n(x, y)\} Q_{n,\xi}(dx) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{R^n} \exp\left\{\sum_{k=1}^p y_k \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n x_i\right\} \prod_{i=1}^n \mu(dx_i) \\ &= \sum_{k=1}^p \lim_{n \rightarrow \infty} \frac{n_{\xi, \theta_k}}{n} \log \int_R \exp(y_k x) \mu(dx) \end{aligned} \quad (5.105)$$

Since by strong law of large numbers,

$$\frac{n_{\xi, \theta_k}}{n} \rightarrow \lambda_k, \quad \varphi\text{-a.s.},$$

it follows that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} H_{n,\xi}(ny) &= \sum_{k=1}^p \lambda_k \log \int_R \exp(y_k x) \mu(dx) \\ &= H(y_1, \dots, y_p), \quad \varphi\text{-a.s.} \end{aligned} \quad (5.106)$$

We now show that $H(y_1, \dots, y_p)$ is finite for all $(y_1, \dots, y_p) \in R^p$. For any $y_k \in R$

$$\begin{aligned} \int_R \exp(y_k x) \mu(dx) &\leq \int_R \exp\left(\frac{y_k^2}{2} + \frac{x^2}{2}\right) \mu(dx) \\ &= \exp\left(\frac{y_k^2}{2}\right) \int_R \exp\left(\frac{x^2}{2}\right) \mu(dx) \end{aligned} \quad (5.107)$$

so by the assumption (5.10) in theorem 5.1

$$\begin{aligned} H(y_1, \dots, y_p) &= \sum_{k=1}^p \lambda_k \log \int_R \exp(y_k x) \mu(dx) \\ &\leq \sum_{k=1}^p \lambda_k \left[\frac{y_k^2}{2} + \log \int_R \exp\left(\frac{x^2}{2}\right) \mu(dx) \right] \\ &< \infty \end{aligned} \quad (5.108)$$

Hence condition 2 in theorem 5.3 is satisfied.

3. Finally we show that $H(y_1, \dots, y_p)$ is differentiable for all $(y_1, \dots, y_p) \in R^p$. It suffices to show that the real valued function $f(y) = \int_R \exp(yx) \mu(dx)$ has continuous derivative. For any $x \in R$ and $\delta > 0$ we have

$$\begin{aligned} \left| \frac{\exp(hx) - 1}{h} \right| &\leq \sum_{k=1}^{\infty} \delta^{k-1} \frac{|x|^k}{k!} \\ &\leq \frac{1}{\delta} \exp(\delta|x|) \\ &\leq \frac{1}{\delta} [\exp(\delta x) + \exp(-\delta x)], \quad |h| \leq \delta. \end{aligned} \quad (5.109)$$

By the Lebesgue dominated convergence theorem $f(y)$ is differentiable at each $y \in R$ with

$$\begin{aligned} f'(y) &= \lim_{h \rightarrow 0} \frac{f(y+h) - f(y)}{h} \\ &= \lim_{h \rightarrow 0} \int_R \left[\frac{\exp(hx) - 1}{h} \right] \exp(yx) \mu(dx) \\ &= \int_R \lim_{h \rightarrow 0} \left[\frac{\exp(hx) - 1}{h} \right] \exp(yx) \mu(dx) \\ &= \int_R x \exp(yx) \mu(dx) \end{aligned} \quad (5.110)$$

Similarly we can show that $f'(y)$ is differentiable at each $y \in R$ with

$$f''(y) = \int_R x^2 \exp(yx) \mu(dx) \quad (5.111)$$

This shows $H(y_1, \dots, y_p)$ is differentiable and the condition 3 of theorem 5.3 is satisfied. Therefore, by theorem 5.3, $(R^p, Q_{n,\xi}, n)$ forms a φ -a.s. large deviation system with rate function

$$\begin{aligned}
 I(m_1, \dots, m_p) &= \sup_{(y_1, \dots, y_p) \in R^p} \left\{ \sum_{k=1}^p [(m_k, y_k) - \lambda_k \log \int_R \exp(y_k x) \mu(dx)] \right\} \\
 &= \sum_{k=1}^p \sup_{y_k \in R} \{ (m_k, y_k) - \lambda_k \log \int_R \exp(y_k x) \mu(dx) \} \\
 &= \sum_{k=1}^p \lambda_k \sup_{y_k \in R} \left\{ \left(\frac{m_k}{\lambda_k}, y_k \right) - \log \int_R \exp(y_k x) \mu(dx) \right\} \\
 &= \sum_{k=1}^p \lambda_k L^* \left(\frac{m_k}{\lambda_k} \right). \tag{5.112}
 \end{aligned}$$

where $(m_1, \dots, m_p) \in R^p$. □

Now the theorem 5.1 will be proved if we can show that the condition (5.78) in Varadhan type theorem 5.4 is satisfied. Denote $c = \max_{1 \leq k, l \leq p} c(\theta_k, \theta_l)$. Then by the Chebyshev exponential inequality

$$\begin{aligned}
 &\int_{\{F_3(m) \geq L\}} \exp[n F_3(m)] Q_{n,\xi}(dm) \\
 &\leq \sum_{\alpha=1}^{\infty} \exp[n(\alpha+1)L] Q_{n,\xi}(F_3(m) \geq \alpha L) \\
 &\leq \sum_{\alpha=1}^{\infty} \exp[n(\alpha+1)L] \exp[-3n\alpha L] \\
 &\quad \int_{R^n} \exp\left[3n \frac{3}{2} \sum_{k,l=1}^p c(\theta_k, \theta_l) \left(\frac{1}{n} \sum_{\xi_i=\theta_k}^n x_i \right) \left(\frac{1}{n} \sum_{\xi_j=\theta_l}^n x_j \right) \right] \prod_{i=1}^n \mu(dx_i) \\
 &\leq \sum_{\alpha=1}^{\infty} \exp[-n(2\alpha-1)L] \int_{R^n} \exp\left[c \sum_{k=1}^p \sum_{\xi_i=\theta_k}^n x_i^2\right] \prod_{i=1}^n \mu(dx_i) \\
 &\leq \sum_{\alpha=1}^{\infty} \exp[-n(2\alpha-1)L] \prod_{k=1}^p \left(\int_R \exp[cx^2] \mu(dx) \right)^{n\epsilon_k} \tag{5.113}
 \end{aligned}$$

Using the assumption that $\int_R \exp[cr^2] \mu(dr)$ is finite and the strong law of large numbers we obtain

$$\begin{aligned}
 & \lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \int_{\{F_\beta(m) \geq L\}} \exp\{n F_\beta(m)\} Q_{n,\xi}(dm) \\
 &= \lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(\sum_{\alpha=1}^{\infty} \exp[-n(2\alpha - 1)L] \right) \\
 &+ \sum_{k=1}^p \limsup_{n \rightarrow \infty} \frac{n_{\xi, \theta_k}}{n} \log \left(\int_R \exp[cr^2] \mu(dr) \right) \\
 &= -\infty, \quad \varphi\text{-a.s.}
 \end{aligned} \tag{5.114}$$

Therefore for almost all $\xi \in \Xi$

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_{n,\xi} &= \sup_{(m_1, \dots, m_p) \in R^p} [F_\beta(m_1, \dots, m_p) - \sum_{k=1}^p \lambda_k L^*\left(\frac{m_k}{\lambda_k}\right)] \\
 &= \sup_{(m_1, \dots, m_p) \in R^p} [\Phi_\beta(m_1, \dots, m_p) - \sum_{k=1}^p \lambda_k L^*(m_k)] \tag{5.115}
 \end{aligned}$$

where $\Phi_\beta(m_1, \dots, m_p)$ is defined by (5.9). This completes the proof of the theorem 5.1. $\square$

5.5 Proof of theorem 5.2

Lemma 5.1 *Let C be a $p \times p$ strictly positive definite symmetric matrix and X be a $p \times 1$ vector. Then*

$$\exp\left\{\frac{1}{2}(X, CX)\right\} = \frac{1}{(2\pi)^{\frac{p}{2}} [\det(C^{-1})]^{\frac{1}{2}}} \int_{R^p} \exp\left\{-\frac{1}{2}(m, Cm) + (X, Cm)\right\} dm \tag{5.116}$$

where C^{-1} denotes the inverse matrix and $\det(C^{-1})$ denotes its determinant.

Proof: Since the matrix C is symmetric, for any $m, X \in R^p$, we have

$$(m, CX) = (X, Cm). \tag{5.117}$$

Thus

$$-\frac{1}{2}(m - X, C(m - X)) = -\frac{1}{2}(m, Cm) + (X, Cm) - \frac{1}{2}(X, CX). \quad (5.118)$$

The equality (5.116) then follows from the Gaussian integral

$$\frac{1}{(2\pi)^{\frac{p}{2}} \det(C^{-1})^{\frac{1}{2}}} \int_{R^p} \exp\left\{-\frac{1}{2}(m - X, C(m - X))\right\} dm = 1 \quad (5.119)$$

□

Lemma 5.2 Let $P_{n,\xi}(dx_1, \dots, dx_n)$ be the random Gibbs measure defined by (5.1) and $P_{n,l,\xi}(dx_1, \dots, dx_l)$ be the l -dimensional marginal distribution of $P_{n,\xi}(dx_1, \dots, dx_n)$.

Then

$$\frac{dP_{n,l,\xi}}{d\mu^{\otimes l}}(x_1, \dots, x_l) = \int_{R^p} f_m^{l,\xi}(x_1, \dots, x_l) Q_{n,\xi}^l(dm) \quad (5.120)$$

where for each $m \in R^p$ and $\xi \in \Xi$, $f_m^{l,\xi}(x_1, \dots, x_l)$ is the density function on R^l with respect to the measure $\mu^{\otimes l}$

$$f_m^{l,\xi}(x_1, \dots, x_l) = \prod_{k=1}^p \prod_{\substack{i=1 \\ \xi_i = \theta_k}}^l \frac{\exp\{x_i[Cm]_k\}}{\int_R \exp\{x_i[Cm]_k\} \mu(dx_i)} \quad (5.121)$$

and $Q_{n,\xi}^l(dm)$ is the probability measure on R^p given by

$$Q_{n,\xi}^l(dm) = \frac{\exp\{-nF_{n,\xi}^l(m)\} dm}{\int_{R^p} \exp\{-nF_{n,\xi}^l(m)\} dm} \quad (5.122)$$

where

$$F_{n,\xi}^l(m) = \frac{1}{23}(m, Cm) - \sum_{k=1}^p \frac{(n-l)_{\xi, \theta_k}}{n} \log \int_R \exp(x[Cm]_k) \mu(dx) \quad (5.123)$$

and $(n-l)_{\xi, \theta_k}$ is defined by

$$(n-l)_{\xi, \theta_k} = \sum_{i=l+1}^n 1_{\theta_k}(\xi_i) \quad (5.124)$$

Proof: For each $\xi \in \Xi$, $k = 1, \dots, p$, denote

$$x_{n,k,\xi} = \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n x_i \quad (5.125)$$

Then

$$\begin{aligned} \exp\left\{\frac{\beta}{2n} \sum_{i,j=1}^n c(\xi_i, \xi_j) x_i x_j\right\} &= \exp\left\{\frac{\beta}{2} \sum_{k,l=1}^p c(\theta_k, \theta_l) \frac{x_{n,k,\xi}}{\sqrt{n}} \frac{x_{n,l,\xi}}{\sqrt{n}}\right\} \\ &= \exp\left\{\frac{\beta}{2} \left(\frac{x_{n,\xi}}{\sqrt{n}}, C \frac{x_{n,\xi}}{\sqrt{n}}\right)\right\} \end{aligned} \quad (5.126)$$

where $x_{n,\xi} = (x_{n,1,\xi}, \dots, x_{n,p,\xi})^t$ is the $p \times 1$ vector and $C = (c(\theta_k, \theta_l))$ is the $p \times p$ strictly positive definite symmetric matrix. By lemma 5.1

$$\begin{aligned} \exp\left\{\frac{\beta}{2n} \sum_{i,j=1}^n c(\xi_i, \xi_j) x_i x_j\right\} &= \frac{1}{(2\pi)^{\frac{p}{2}} \det(C^{-1})^{\frac{1}{2}}} \int_{R^p} \exp\left\{-\frac{1}{2}(m, Cm) + \left(\frac{\sqrt{\beta} x_{n,\xi}}{\sqrt{n}}, Cm\right)\right\} dm \\ &= \frac{\sqrt{n}}{\sqrt{\beta} (2\pi)^{\frac{p}{2}} \det(C^{-1})^{\frac{1}{2}}} \int_{R^p} \exp\left\{-n \frac{1}{2\beta} (m, Cm) + (x_{n,\xi}, Cm)\right\} dm \end{aligned} \quad (5.127)$$

It follows that

$$\begin{aligned} \frac{dP_{n,l,\xi}}{d\mu^{\otimes l}}(x_1, \dots, x_l) &= \frac{1}{Z_{n,\xi}} \int_{R^{n-l}} \exp\left\{\frac{\beta}{2n} \sum_{i,j=1}^n c(\xi_i, \xi_j) x_i x_j\right\} \prod_{i=l+1}^n \mu(dx_i) \\ &= \frac{\sqrt{n}}{\sqrt{\beta} (2\pi)^{\frac{p}{2}} [\det(C^{-1})]^{\frac{1}{2}} Z_{n,\xi}} \int_{R^p} \exp\left\{-n \frac{1}{2\beta} (m, Cm)\right\} \\ &\quad \left(\int_{R^{n-l}} \exp\{(x_{n,\xi}, Cm)\} \prod_{i=l+1}^n \mu(dx_i) \right) dm \\ &= \frac{\sqrt{n}}{\sqrt{\beta} (2\pi)^{\frac{p}{2}} [\det(C^{-1})]^{\frac{1}{2}} Z_{n,\xi}} \int_{R^p} \exp\left\{-n \frac{1}{2\beta} (m, Cm)\right\} \\ &\quad \left(\int_{R^{n-l}} \exp\left\{\sum_{k=1}^p \left(\sum_{\substack{i=1 \\ \xi_i = \theta_k}}^n x_i\right) [Cm]_k\right\} \prod_{i=l+1}^n \mu(dx_i) \right) dm \\ &= \frac{\sqrt{n}}{\sqrt{\beta} (2\pi)^{\frac{p}{2}} [\det(C^{-1})]^{\frac{1}{2}} Z_{n,\xi}} \int_{R^p} \exp\left\{-n \frac{1}{2\beta} (m, Cm)\right\} \end{aligned}$$

$$\begin{aligned}
& \left(\prod_{k=1}^p \prod_{\substack{i=1 \\ \xi_i = \theta_k}}^n \int_R \exp\{x_i [Cm]_k\} \mu(dx_i) \right) \exp\left\{ \sum_{k=1}^p \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^l x_i [Cm]_k \right\} dm \\
&= \frac{\sqrt{n}}{\sqrt{\beta}(2\pi)^{\frac{p}{2}} [\det(C^{-1})]^{\frac{1}{2}} Z_{n,\xi}} \int_{R^p} \exp\left\{ -n \left[\frac{1}{2\beta} (m, Cm) \right. \right. \\
&\quad \left. \left. - \sum_{k=1}^p \frac{(n-l)_{\xi, \theta_k}}{n} \log \int_R \exp(x [Cm]_k) \mu(dx) \right] \right\} \exp\left\{ \sum_{k=1}^p \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^l x_i [Cm]_k \right\} dm \\
&= \int_{R^p} \frac{\exp\left\{ \sum_{k=1}^p \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^l x_i [Cm]_k \right\}}{\int_{R^l} \exp\left\{ \sum_{k=1}^p \sum_{\substack{i=1 \\ \xi_i = \theta_k}}^l x_i [Cm]_k \right\} \prod_{i=1}^l \mu(dx_i) \right\}} \\
&\quad \frac{\exp\left\{ -n \left[\frac{1}{2\beta} (m, Cm) - \sum_{k=1}^p \frac{(n-l)_{\xi, \theta_k}}{n} \log \int_R \exp(x [Cm]_k) \mu(dx) \right] \right\} dm}{\int_{R^p} \exp\left\{ -n \left[\frac{1}{2\beta} (m, Cm) - \sum_{k=1}^p \frac{(n-l)_{\xi, \theta_k}}{n} \log \int_R \exp(x [Cm]_k) \mu(dx) \right] \right\} dm} \\
&= \int_{R^p} f_m^{l,\xi}(x_1, \dots, x_l) Q_{n,\xi}^l(dm) \tag{5.128}
\end{aligned}$$

where $f_m^{l,\xi}(x_1, \dots, x_l)$ and $Q_{n,\xi}^l(dm)$ are given by (5.121) and (5.122) respectively. $\square$

We now consider the probability measures $Q_{n,\xi}^l(dm)$ on R^p defined by (5.122). Let $F(m)$ be the function defined on R^p by

$$F(m) = \frac{1}{2\beta} (m, Cm) - \sum_{k=1}^p \lambda_k \log \int_R \exp(x [Cm]_k) \mu(dx) \tag{5.129}$$

Then $F(m)$ is a lower semicontinuous convex function. Let K_0 be the set of points m of R^p such that $F(m)$ attains its minimum at m

$$K_0 = \{m : F(m) = \inf_{m \in R^p} F(m)\} \tag{5.130}$$

Lemma 5.3 *We have*

1. $(Q_{n,\xi}^l(dm))_{n=1}^\infty$ is tight in $\mathbf{M}_1(R^p)$ for each ξ and l and the limit points do not depend on ξ and l .
2. Any limit point Q of $(Q_{n,\xi}^l(dm))_{n=1}^\infty$ is concentrated on the set K_0 .

Proof: Let $a = \inf_{m \in R^p} F(m)$. For $\delta > 0$ define

$$K_\delta = \{m : F(m) \leq a + \delta\} \quad (5.131)$$

Then K_δ is a compact set because $F(m)$ is lower semicontinuous convex function.

Denote by $G(dm)$ the Gaussian measure on R^p with mean zero and covariance matrix C , i.e.,

$$G(A) = \frac{1}{(2\pi)^{\frac{p}{2}} [\det(C^{-1})]^{\frac{1}{2}}} \int_A \exp\left\{-\frac{1}{2}(m, Cm)\right\} dm, \quad A \in \mathcal{B}(R^p) \quad (5.132)$$

Let $V_{n,\xi}$ be defined by

$$V_{n,\xi} = \frac{n-1}{n} \frac{1}{2\beta} (m, Cm) - \sum_{k=1}^p \frac{(n-l)_{\xi, \theta_k}}{n} \log \int_R \exp\{x[Cm]_k\} \mu(dx) \quad (5.133)$$

Then by the strong law of large numbers, for any $m \in R^p$, we have

$$\lim_{n \rightarrow \infty} V_{n,\xi}(m) = F(m) \quad \varphi\text{-a.s.} \quad (5.134)$$

Now let $m \in K_\delta^c$, then (5.134) implies that there exists a φ -null set $\mathcal{N}$ and an integer n_0 such that for $\xi \notin \mathcal{N}$

$$V_{n,\xi}(m) > a + \delta \quad \text{for all } n \geq n_0 \quad (5.135)$$

Thus for sufficiently large n

$$\begin{aligned} Q'_{n,\xi}\{K_\delta^c\} &= \frac{\int_{K_\delta^c} \exp\{-nF_{n,\xi}^l(m)\} dm}{\int_{R^p} \exp\{-nF_{n,\xi}^l(m)\} dm} \\ &= \frac{\int_{K_\delta^c} \exp\{-nV_{n,\xi}(m)\} G(dm)}{\int_{R^p} \exp\{-nV_{n,\xi}(m)\} G(dm)} \\ &< \frac{\int_{K_\delta^c} \exp\{-n[a + \delta]\} G(dm)}{\int_{R^p} \exp\{-nV_{n,\xi}(m)\} G(dm)} \\ &< \frac{1}{\int_{R^p} \exp\{-n(V_{n,\xi}(m) - [a + \delta])\} G(dm)} \\ &< \frac{1}{\int_{\{\cap_{n=1}^\infty V_{n,\xi}(m) < [a + \delta]\}} \exp\{-n(V_{n,\xi}(m) - [a + \delta])\} G(dm)} \\ &\rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned} \quad (5.136)$$

Hence for any $\epsilon > 0$ there exists $n_0 > 0$ such that

$$Q_{n,\xi}^l\{K_\delta\} \geq 1 - \epsilon \quad \text{for } n \geq n_0 \quad (5.137)$$

Since K_δ is a compact subset of R^p which does not depend on ξ and l . $(Q_{n,\xi}^l)_{n=1}^\infty$ is tight and the limit points are independent of ξ and l . This proves 1.

To prove 2 we simply note that K_δ is a set of monotone decreasing sets as $\delta \downarrow 0$ such that

$$K_0 = \bigcap_\delta K_\delta \quad (5.138)$$

Let Q be any limit point of $Q_{n,\xi}^l$. Then we have

$$\begin{aligned} Q(K_0) &\geq \limsup_{n \rightarrow \infty} Q_{n,\xi}^l(K_0) \\ &= \limsup_{n \rightarrow \infty} \lim_{\delta \downarrow 0} Q_{n,\xi}^l(K_\delta) \\ &= \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} Q_{n,\xi}^l(K_\delta) \\ &= 1 \end{aligned} \quad (5.139)$$

This completes the proof of the lemma 5.3. $\square$

Now we are ready to prove theorem 5.2. Since $\tilde{P}_{n,\xi}(dx_1, dx_2, \dots)$ is a probability measure in $M_1(R^{\otimes \infty})$, it suffices to show that its finite dimensional distribution $\tilde{P}_{n,l,\xi}(dx_1, \dots, dx_l)$ converges to the required one. For any $g \in C_b(R^l)$, by Fubini theorem and the definition of weak convergence

$$\begin{aligned} &\lim_{n \rightarrow \infty} \int_{R^l} g(x_1, \dots, x_l) \tilde{P}_{n,l,\xi}(dx_1, \dots, dx_l) \\ &= \lim_{n \rightarrow \infty} \int_{R^l} \int_{R^p} g(x_1, \dots, x_l) f_m^{l,\xi}(x_1, \dots, x_l) Q_{n,\xi}^l(dm) \mu^{\otimes l}(dx_1, \dots, dx_l) \\ &= \lim_{n \rightarrow \infty} \int_{R^p} \int_{R^l} g(x_1, \dots, x_l) f_m^{l,\xi}(x_1, \dots, x_l) \mu^{\otimes l}(dx_1, \dots, dx_l) Q_{n,\xi}^l(dm) \\ &= \int_{R^p} \int_{R^l} g(x_1, \dots, x_l) f_m^{l,\xi}(x_1, \dots, x_l) \mu^{\otimes l}(dx_1, \dots, dx_l) Q(dm) \\ &= \int_{R^l} g(x_1, \dots, x_l) \int_{R^p} f_m^{l,\xi}(x_1, \dots, x_l) \mu^{\otimes l}(dx_1, \dots, dx_l) Q(dm) \end{aligned} \quad (5.140)$$

This means that $\tilde{P}_{n,l,\xi}(dx_1, \dots, dx_l) \Rightarrow \int_{R^p} f_m^{l,\xi}(x_1, \dots, x_l) \mu^{\otimes l}(dx_1, \dots, dx_l) Q(dm)$.

Thus the theorem 5.2 is completed. $\square$

Chapter 6

Gaussian Limit

Many stochastic systems in applications can be described by the following Itô stochastic differential equations:

$$dx_i^{(n)}(t) = -x_i^{(n)}(t)dt + \frac{1}{\sqrt{n}} \sum_{j=1}^n J_{i,j} \varphi(x_j^{(n)}(t))dt + \sigma dw_i(t) \quad (6.1)$$
$$i = 1, \dots, n$$

where $w_1, \dots, w_n$ are n independent 1-dimensional Wiener processes; $\{J_{i,j}\}_{i,j=1}^n$ is a set of i.i.d. R -valued random variables with zero mean and finite variance; σ is a positive constant; and φ is a bounded continuous function defined on R . One is interested in the behavior of this system as $n \rightarrow \infty$, and hopes that the limiting process will turn out to be a Gaussian one with the covariance satisfying some self-consistent equation. To explain the phenomenon behind these equations let us consider the following two examples.

The first example is a continuous non-linear neural network model studied by Sompolinsky and Crisanti [42]. The model consists of n continuous neurons denoted by $S_1(t), \dots, S_n(t)$, where $t \geq 0$ and $-1 \leq S_i(t) \leq +1$. Associated with each neuron, there is a local field $h_i \in R$. The relationship between $S_i(t)$ and $h_i(t)$ is defined by

$$S_i(t) = \varphi(h_i(t))$$

where φ is a non-linear function. In the biological context, h_i may be related to the membrane potential of the nerve cell and S_i to its electrical activity (e.g., its firing rate). The dynamics of the network is given by n coupled first order differential equations

$$\frac{d}{dt}h_i(t) = -h_i(t) + \frac{1}{\sqrt{n}} \sum_{j=1}^n J_{i,j} S_j(t) = -h_i(t) + \frac{1}{\sqrt{n}} \sum_{j=1}^n J_{i,j} \varphi(h_j(t)) \quad (6.2)$$

$$i = 1, \dots, n$$

where $\{J_{i,j}\}_{i,j=1}^n$ is a set of i.i.d. random variables with normal $N(0, J^2)$ distribution, which couples the output of the j th (presynaptic) neuron to the input of the i th (postsynaptic) neuron. $\varphi(x)$ is defined by

$$\varphi(x) = \tanh(\beta x) \quad (6.3)$$

where β is a positive constant representing the gain parameter. Based on the dynamical mean field theory on spin glass Sompolinsky and Crisanti [42] proposed that as $n \rightarrow \infty$ the behavior of this system can be described by a self-consistent single equation

$$\frac{d}{dt}h(t) = -h(t) + \xi(t) \quad (6.4)$$

where ξ is a Gaussian process with mean zero and variance satisfying the equation

$$E[\xi(t)\xi(s)] = E[\varphi(h(t))\varphi(h(s))] \quad (6.5)$$

Then they studied this limit process and showed that there is a transition from a stationary phase to a chaotic phase occurring at a critical value of the parameter βJ . The second example is a complex ecosystem model studied by Rieger [39]. The model is described by a generalized Lotka-Volterra equation

$$\frac{d}{dt}n_i(t) = n_i(t) \left(f_i(n_i(t)) + \frac{1}{\sqrt{n}} \sum_{\substack{j=1 \\ j \neq i}}^n J_{i,j} n_j(t) + \xi_i(t) \right) \quad (6.6)$$

$$i = 1, \dots, n$$

where $n_i(t)$ is the population density of the i th species at time t . The function $f_i(x)$ describes the development of the i th species without interaction with other species. The interactions among species are described by $\{J_{i,j}\}_{i,j=1}^n$ which is a set of i.i.d. normal $N(0, J^2)$ distributed random variables and ξ_i is a white noise that takes into account the influence of the environment, e.g., climate, disease, etc. If one considers $f_i(x) = -\log(x)$, the so called Gompertz self-regulation model and let $S_i(t) = \log n_i(t)$ then the equation (6.6) can be written as

$$\frac{d}{dt} S_i(t) = -S_i(t) + \frac{1}{\sqrt{n}} \sum_{\substack{j=1 \\ j \neq i}}^n J_{i,j} \exp(S_j(t)) + \xi_i(t) \quad (6.7)$$

$$i = 1, \dots, n$$

Again, by using the dynamical mean field theory on spin glass Rieger proposed that as $n \rightarrow \infty$ the behavior of this system can be described by a self-consistent single equation

$$\frac{dS(t)}{dt} = -S(t) + \zeta(t) + \xi(t) \quad (6.8)$$

where ζ is a Gaussian process with mean zero and variance satisfying the equation

$$E[\zeta(t)\zeta(s)] = E[\exp(S_i(t) + S_i(s))] \quad (6.9)$$

Based on this limit process Rieger then showed a stability-instability transition depending on the interaction strength J and the noise.

These two examples and many other models in applications suggest one to consider the equations of type (6.1). However, mathematically, it is very hard to deal with the equations (6.1) because of the random interactions $\{J_{i,j}\}_{i,j=1}^n$ (See [22][23] for similar difficulties). On the other hand, if we replace the i.i.d. random variables $\{J_{i,j}\}_{i,j=1}^n$ by the white noises $\{\dot{w}_{i,j}\}_{i,j=1}^n$, then one can use many powerful tools

from stochastic analysis to study this kind of equations. Hence in this chapter, we consider the following similar but mathematically more tractable Itô equations

$$dx_i^{(n)}(t) = -x_i^{(n)}(t)dt + \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi(x_j^{(n)}(t))dw_{ij}(t) \quad (6.10)$$

$$i = 1, \dots, n$$

where $\{w_{ij}\}_{i,j=1}^n$ is a set of independent 1-dimensional Brownian motions and $\varphi(x)$ is a bounded continuous function on R assuming the following form

$$\varphi(x) = \sqrt{2} \tanh(\beta x) \quad (6.11)$$

where β is a positive constant measuring the non-linearity of the function.

We show that, as $n \rightarrow \infty$, the empirical measure associated with $x_1^{(n)}, \dots, x_n^{(n)}$ converges in probability to a deterministic measure which is the distribution of the process $\{x(t) : t \geq 0\}$ determined uniquely by the equation

$$dx(t) = -x(t)dt + \sqrt{\langle \mu(t), \varphi^2 \rangle} dw(t) \quad (6.12)$$

$$\mu(t) = \text{distribution of } x(t)$$

where w is a 1-dimensional Wiener process, the notation $\langle \mu(t), \varphi^2 \rangle$ denotes the integral

$$\langle \mu(t), \varphi^2 \rangle = \int_R \varphi^2(x) \mu(t)(dx)$$

For sufficiently large β , this process has a unique invariant measure which turns out to a Gaussian measure with zero mean such that its variance is determined by the SK-spin glass fixed point equation (2.20). Therefore, if we start the process with this invariant measure, then equation (6.12) determines a unique stationary Gaussian process.

Since there is no major difficulty to consider the more general drift term instead of a linear one as in (6.10), we will prove the limiting theorem for general function $b(x)$.

The chapter consists of 4 sections. In Section 1 we state the main results. In section 2 we prove the existence and uniqueness of the equations (6.10) and (6.12). In Section 3 we prove the main limit theorem. Finally, in section 4 we study the stationary solution of the limiting process (6.12).

6.1 Statement of the results

As in previous chapters we let $(\Omega, \mathcal{F}, P)$ be a given complete probability space and $(\mathcal{F}_t)_{t \geq 0}$ be a family of sub σ -algebras of $\mathcal{F}$ which is assumed to satisfy the standard conditions (refer back to chapter 1). Let $\{w_{ij}\}_{i,j=1}^n$ be n^2 independent 1-dimensional $\mathcal{F}_t$ -Brownian motions defined on this probability space. Consider the Itô equations in R

$$dx_i^{(n)}(t) = b(x_i^{(n)}(t))dt + \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi(x_j^{(n)}(t))dw_{ij}(t) \quad (6.13)$$

$$i = 1, \dots, n$$

where $b(x)$ and $\varphi(x)$ are continuous functions defined on R which are assumed to satisfy the conditions

H_1 . There exists a constant K such that for all $x, y \in R$

$$|b(x) - b(y)| \leq K|x - y|.$$

$$|b(x)|^2 \leq K(1 + |x|^2).$$

H_2 . $\varphi(x)$ is bounded and Lipschitz with $\varphi(0) = 0$.

H_3 . $x^{(n)}(0) = (x_i^{(n)}(0))_{i=1}^n \in L^2(\Omega, \mathcal{F}_0, P)$ is independent of $\{w_{ij}\}_{i,j=1}^n$.

Under the above conditions it is easy to show (Proposition 6.1) that the system of stochastic differential equations (6.13) has a unique solution $(x_1^{(n)}(\cdot), \dots, x_n^{(n)}(\cdot))$

for each $n \geq 1$. Let X_n be the associated empirical measure defined by

$$X_n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i^{(n)}(\cdot)} \quad (6.14)$$

which is viewed as a random probability measure on the path space $C([0, \infty), R)$. For $t \geq 0$, we denote by $X_n(t)$ the marginal measure of X_n . Then the first result of this chapter can be stated as following:

Theorem 6.1 *Suppose that the conditions H_1 and H_3 are satisfied. Furthermore, let the following two initial conditions are satisfied :*

1. $\sup_{n \geq 1} E(|x_i^{(n)}(0)|^2) < \infty$ for each $i \geq 1$
2. $X_n(0) \Rightarrow u$ (weak convergence) in $M_1(R)$ as $n \rightarrow \infty$.

Then the laws $(\mathcal{L}(X_n))_{n=1}^\infty$ of $(X_n)_{n=1}^\infty$ converge weakly in $M_1(M_1(C([0, \infty), R)))$ to a Dirac measure δ_μ , written as $\mathcal{L}(X_n) \Rightarrow \delta_\mu$, where μ is the law on $C([0, \infty), R)$ of the unique solution $\{x(t) : t \geq 0\}$ of the following equation

$$dx(t) = b(x(t))dt + \sqrt{\langle \mu(t), \varphi^2 \rangle} dw(t) \quad (6.15)$$

$\mu(t)$ is the distribution of $x(t)$

$$\mu(0) = u$$

where w is the 1-dimensional Brownian motion.

Now let $b(x) = -x$ and $\varphi(x) = \sqrt{2} \tanh(\beta x)$ with $\beta > 0$ then we have the following result:

Theorem 6.2 *There exists a critical value $\beta_c = 1$ such that for $\beta \leq \beta_c$, the stochastic process $\{x(t) : t \geq 0\}$ determined uniquely by the equation*

$$dx(t) = -x(t)dt + \sqrt{\langle \mu(t), \varphi^2 \rangle} dw(t) \quad (6.16)$$

$\mu(t)$ is the distribution of $x(t)$

$$\varphi(x) = \sqrt{2} \tanh(\beta x)$$

has no invariant probability measure. For $\beta > \beta_c$, the equation (6.16) has a unique Gaussian invariant probability measure

$$P_\sigma(dx) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} dx \quad (6.17)$$

where σ is the unique positive solution of the Sherrington-Kirkpatrick spin glass fixed point equation

$$\sigma = \int_{-\infty}^{\infty} \tanh^2(\beta \sigma^{\frac{1}{2}} x) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx. \quad (6.18)$$

6.2 Existence and uniqueness

We first prove the existence and uniqueness of the finite system defined by the equation (6.13).

Proposition 6.1 Suppose that the hypotheses H_1 to H_3 are satisfied. Then for each $n \geq 1$, the equation (6.13) defines a unique Markov diffusion process

$$x^{(n)} = \{(x_1^{(n)}(t), \dots, x_n^{(n)}(t)); t \geq 0\}$$

on R^n . The domain $D(A^{(n)})$ of the infinitesimal generator $A^{(n)}$ contains $C_K^2(R^n)$. For $f \in C_K^2(R^n)$, $A^{(n)}$ coincides with the second order differential operator

$$L^{(n)}f(x_1, \dots, x_n) = \sum_{i=1}^n L_i(X_n)f(x_1, \dots, x_n) \quad (6.19)$$

where for $(x_1, \dots, x_n) \in R^n$, X_n is the associated empirical measure

$$X_n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}. \quad (6.20)$$

and $L(m)$, $m \in \mathbf{M}_1(R)$, is the operator defined by

$$L(m)f(x) = \frac{1}{2} \langle m, \varphi^2 \rangle f''(x) + b(x)f'(x). \quad (6.21)$$

In (6.19), the index i in $L_i(X_n)$ indicates that the operator $L(X_n)$ acts upon the i th variable x_i .

Proof: Denote

$$\begin{aligned} x(t) &= (x_1(t), \dots, x_n(t)) \\ U(x(t)) &= (b(x_1(t)), \dots, b(x_n(t))) \\ \Phi(x(t)) &= \frac{1}{\sqrt{n}}(\varphi(x_1(t)), \dots, \varphi(x_n(t))) \\ w(t) &= \begin{pmatrix} w_{11}(t) & \cdots & w_{n,1}(t) \\ w_{12}(t) & \cdots & w_{n,2}(t) \\ \vdots & & \vdots \\ w_{1n}(t) & \cdots & w_{nn}(t) \end{pmatrix} \end{aligned} \quad (6.22)$$

Then equation (6.13) can be written as an Itô equation in R^n

$$dx(t) = U(x(t))dt + \Phi(x(t))dw(t) \quad (6.23)$$

From the hypotheses on $b(x)$ and $\varphi(x)$ one can easily check that $U(x)$ is a Lipschitz continuous function on R^n which satisfying the linear growth condition. Also, $\Phi(x)$ is a bounded continuous function on R^n . The existence and uniqueness for the equation (6.13) then follow from the standard Picard iteration on stochastic differential equations (theorem 1.1). It remains to show that, for each $f \in C_K^2(R^n)$, the process

$$M_n^f(t) = f(x(t)) - f(x(0)) - \int_0^t L^{(n)}f(x(s))ds \quad (6.24)$$

is a martingale. But this follows from Itô's lemma. Hence the theorem is true. $\square$

We now prove the existence and uniqueness of the limiting process determined by the equation (6.15). Let $(BL(R), \|\cdot\|_{BL})$ be the Banach space of real valued

bounded Lipschitz continuous functions f on R , where the norm $\|\cdot\|_{BL}$ is defined by

$$\|f\|_{BL} = \sup_{x \in R} |f(x)| + \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}. \quad (6.25)$$

We denote by $(BL^*(R), \|\cdot\|_{BL}^*)$ the corresponding dual space, where the dual norm $\|\cdot\|_{BL}^*$ is defined by

$$\|\mu\|_{BL}^* = \sup_{f \in BL(R), \|f\|_{BL} \leq 1} \{|\langle f, \mu \rangle|\}, \quad \mu, \nu \in BL^*(R). \quad (6.26)$$

It is known that $\|\mu - \nu\|_{BL}^*$ is a metric on $M_1(R)$ which generates the usual topology of weak convergence [13]. As in chapter 3 we introduce the Wasserstein metric W_T on $M_1(C([0, T], R))$ as

$$W_T(\mu, \nu) = \inf_{(\xi, \eta)} E(\sup_{t \leq T} |\xi(t) - \eta(t)|) \quad (6.27)$$

where the infimum is taken over all the pair (ξ, η) which are continuous processes such that ξ has the distribution μ and η has the distribution ν .

Proposition 6.2 *Suppose that hypothesis H_1 and H_2 are satisfied. Furthermore, let $x(0) \in L^2(\Omega, \mathcal{F}_0, P)$ be independent of $w(t)$. Then the stochastic differential equation (6.15) has a unique solution $x = \{x(t) : t \geq 0\}$.*

Proof: Let $T > 0$ be arbitrary. We first note that under the given conditions, for each $m \in M_1(C([0, T], R))$, the following ordinary stochastic differential equation

$$dx(t) = b(x(t))dt + \sqrt{\langle m(t), \varphi^2 \rangle} dw(t) \quad (6.28)$$

has a unique solution $x_m = \{x_m(t) : t \leq T\}$. The law of x_m on $C([0, T], R)$ is denoted by $\mathcal{L}(x_m)$.

Let Π be the operator on $M_1(C([0, T], R))$ which maps m to $\mathcal{L}(x_m)$. Then as in the proof of theorem 3.2, it suffices to show that the map Π has a unique fixed point.

To this end let μ and ν be two arbitrary elements in $\mathbf{M}_1(C([0, T], R))$ and let x_μ and x_ν be the corresponding solutions of the equation (6.28), i.e.

$$x_\mu(t) = x(0) + \int_0^t b(x_\mu(s))ds + \int_0^t \sqrt{\langle \mu(s), \varphi^2 \rangle} dw(s), \quad (6.29)$$

$$x_\nu(t) = x(0) + \int_0^t b(x_\nu(s))ds + \int_0^t \sqrt{\langle \nu(s), \varphi^2 \rangle} dw(s) \quad (6.30)$$

Then by the Burkholder-Davis-Gundy martingale inequality and the hypothesis

$$\begin{aligned} & E(\sup_{t \leq T} |x_\mu(t) - x_\nu(t)|) \\ & \leq E(\sup_{t \leq T} |\int_0^t [b(x_\mu(s)) - b(x_\nu(s))]ds|) \\ & + E(\sup_{t \leq T} |\int_0^t [\sqrt{\langle \mu(s), \varphi^2 \rangle} - \sqrt{\langle \nu(s), \varphi^2 \rangle}]dw(s)|) \\ & \leq KTE(\sup_{t \leq T} |x_\mu(t) - x_\nu(t)|) \\ & + c_1 E\{(\int_0^T |\sqrt{\langle \mu(s), \varphi^2 \rangle} - \sqrt{\langle \nu(s), \varphi^2 \rangle}|^2 ds)^{\frac{1}{2}}\} \end{aligned} \quad (6.31)$$

Using the inequality

$$(\sqrt{x} - \sqrt{y})^2 \leq |x - y|, \quad x, y \geq 0, \quad (6.32)$$

and the fact that $\varphi^2(x)$ is a bounded Lipschitz continuous function and the lemma 3.1 we can write the last term in (6.31) as

$$\begin{aligned} & c_1 E[(\int_0^T |\sqrt{\langle \mu(s), \varphi^2 \rangle} - \sqrt{\langle \nu(s), \varphi^2 \rangle}|^2 ds)^{\frac{1}{2}}] \\ & \leq c_1 E[(\int_0^T |\langle \mu(s) - \nu(s), \varphi^2 \rangle| ds)^{\frac{1}{2}}] \\ & \leq c_1 \sqrt{T} E(\sup_{t \leq T} |\langle \mu(t) - \nu(t), \varphi^2 \rangle|^{\frac{1}{2}}) \\ & \leq c_1 \sqrt{T} E(\sup_{t \leq T} |\langle \mu(t) - \nu(t), 1 + \varphi^2 \rangle|) \\ & \leq c_1 \sqrt{T} \|1 + \varphi^2\|_{BL} \sup_{t \leq T} \|\mu(t) - \nu(t)\|_{BL}^* \\ & \leq c_2 \sqrt{T} W_T(\mu, \nu) \end{aligned} \quad (6.33)$$

Therefore

$$\begin{aligned}
 E(\sup_{t \leq T} |x_\mu(t) - x_\nu(t)|) & \\
 & \leq KT E(\sup_{t \leq T} |x_\mu(t) - x_\nu(t)|) + c_2 \sqrt{T} W_T(\mu, \nu) \\
 & \leq (KT + c_2 \sqrt{T}) [E(\sup_{t \leq T} |x_\mu(t) - x_\nu(t)|) + W_T(\mu, \nu)]
 \end{aligned} \tag{6.34}$$

For small enough T , we can have $(KT + c_2 \sqrt{T}) < \frac{1}{3}$. Then

$$E(\sup_{t \leq T} |x_\mu(t) - x_\nu(t)|) < \frac{1}{3} [E(\sup_{t \leq T} |x_\mu(t) - x_\nu(t)|) + W_T(\mu, \nu)] \tag{6.35}$$

Thus

$$E(\sup_{t \leq T} |x_\mu(t) - x_\nu(t)|) < \frac{1}{2} W_T(\mu, \nu). \tag{6.36}$$

This then yields

$$\begin{aligned}
 W_T(\Pi(\mu), \Pi(\nu)) &= W_T(\mathcal{L}(x_\mu), \mathcal{L}(x_\nu)) \\
 &\leq E(\sup_{t \leq T} |x_\mu(t) - x_\nu(t)|) \\
 &< \frac{1}{2} W_T(\mu, \nu)
 \end{aligned} \tag{6.37}$$

Hence, for small enough T , Π is a contraction map and therefore produces a unique fixed point on $\mathbf{M}_1(C([0, T], R))$. This proves the existence and uniqueness of the solution for the equation (6.15) up to some small time T . We can easily extend this solution to any finite interval by considering $[0, T], [T, 2T], [2T, 3T], \dots$, etc. $\square$

6.3 McKean-Vlasov limit

We start by proving the following lemma.

Lemma 6.1 *Suppose that for each $i \geq 1$*

$$\sup_{n \geq 1} E(|x_i^{(n)}(0)|^2) < \infty. \tag{6.38}$$

Then for any $T > 0$

$$\sup_{n \geq 1} E(\sup_{0 \leq t \leq T} |x_i^{(n)}(t)|^2) < \infty \quad (6.39)$$

Proof: Let i be fixed. Then for $t \leq T$

$$|x_i^{(n)}(t)|^2 \leq c_1(|x_i^{(n)}(0)|^2 + |\int_0^t b(x_i^{(n)}(s))ds|^2 + |\int_0^t \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi(x_j^{(n)}(s))dw_{1j}(s)|^2) \quad (6.40)$$

Since

$$\begin{aligned} E(\sup_{t \leq T} |\int_0^t b(x_i^{(n)}(s))ds|^2) &\leq TE(\int_0^T |b(x_i^{(n)}(s))|^2 ds) \\ &\leq K(T^2 + T \int_0^T E|x_i^{(n)}(s)|^2 ds) \\ &\leq K(T^2 + T \int_0^T E(\sup_{s \leq t} |x_i^{(n)}(s)|^2) dt) \end{aligned} \quad (6.41)$$

and according to Doob's martingale inequality

$$\begin{aligned} E(\sup_{t \leq T} |\int_0^t \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi(x_j^{(n)}(s))dw_{1j}(s)|^2) &\leq c_2 E(|\int_0^T \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi(x_j^{(n)}(s))dw_{1j}(s)|^2) \\ &\leq c_3 E(\int_0^T \frac{1}{n} \sum_{j=1}^n \varphi^2(x_j^{(n)}(s))ds) \\ &\leq c_4 \quad \text{because } \varphi(x) \text{ is bounded.} \end{aligned} \quad (6.42)$$

It follows that

$$\sup_{n \geq 1} E(\sup_{t \leq T} |x_i^{(n)}(t)|^2) \leq k_1 + k_2 \sup_{n \geq 1} E(|x_i^{(n)}(0)|^2) + k_3 \int_0^T \sup_{n \geq 1} E(\sup_{s \leq t} |x_i^{(n)}(s)|^2) dt \quad (6.43)$$

for some constants k_1, k_2 and k_3 . The result then follows from the Gronwall's inequality. $\square$

In order to prove theorem 6.1 we need to show that the sequence $(\mathcal{L}(X_n))_{n=1}^{\infty}$ of probability measures is tightness. Denote $Q_n = \mathcal{L}(X_n)$.

Proposition 6.3 $(Q_n)_{n=1}^{\infty}$ is tight in $\mathbf{M}_1(\mathbf{M}_1(C([0, \infty), R)))$

Proof: According to theorem 1.5 of chapter 1, it suffices to prove that the intensity measure $(I_n)_{n=1}^{\infty}$ of $(Q_n)_{n=1}^{\infty}$ is tight in $\mathbf{M}_1(C([0, \infty), R))$.

Since the processes $x_1^{(n)}, \dots, x_n^{(n)}$ have the same distribution, for $f \in C_b(R)$, we have

$$\begin{aligned}
 \langle I_n, f \rangle &= \int_{\mathbf{M}_1(C([0, \infty), R))} \langle m, f \rangle Q_n(dm) \\
 &= E\left(\left\langle \frac{1}{n} \sum_{i=1}^n \delta_{x_i^{(n)}}, f \right\rangle\right) \\
 &= \frac{1}{n} \sum_{i=1}^n E(f(x_i^{(n)})) \\
 &= E(f(x_1^{(n)}))
 \end{aligned} \tag{6.44}$$

Thus it suffices to show that the laws $(\mathcal{L}(x_1^{(n)}))_{n=1}^{\infty}$ of the processes $(x_1^{(n)})_{n=1}^{\infty}$ is tight in $\mathbf{M}_1(C([0, \infty), R))$.

Now $x_1^{(n)}$ is a continuous semimartingale which has the representation

$$x_1^{(n)}(t) = A_n(t) + M_n(t) \tag{6.45}$$

where

$$A_n(t) = x_1^{(n)}(0) + \int_0^t b(x_1^{(n)}(s)) ds \tag{6.46}$$

and

$$M_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \int_0^t \varphi(x_j^{(n)}(s)) dw_{1j}(s) \tag{6.47}$$

We check the two conditions in theorem 1.6 are satisfied.

1. For each fixed $t > 0$, the sequence $(\mathcal{L}(x_1^{(n)}(t)))_{n=1}^{\infty}$ of probability measures is tight in $M_1(R)$ because by lemma 6.1 we have

$$\sup_{n \geq 1} E(|x_1^{(n)}(t)|^2) < \infty. \quad (6.48)$$

2. Let τ_n be any stopping time bounded by T and let θ be any positive number bounded by δ . We have for all $n \geq 1$

$$\begin{aligned} & E(|A_n(\tau_n + \theta) - A_n(\tau_n)|^2) \\ &= E(|\int_{\tau_n}^{\tau_n + \theta} b(x_1^{(n)}(s)) ds|^2) \\ &\leq \delta E(\int_{\tau_n}^{\tau_n + \theta} |b(x_1^{(n)}(s))|^2 ds) \\ &\leq \delta E(\int_{\tau_n}^{\tau_n + \theta} K(1 + |x_1^{(n)}(s)|^2) ds) \\ &\leq \delta^2 K[1 + \sup_{n \geq 1} E(\sup_{s \leq T + \delta} |x_1^{(n)}(s)|^2)] \\ &\rightarrow 0 \text{ as } \delta \rightarrow 0. \end{aligned} \quad (6.49)$$

because $\sup_{n \geq 1} E(\sup_{s \leq T + \delta} |x_1^{(n)}(s)|^2)$ is finite by lemma 6.1. By Chebyshev's inequality, for any positive ϵ and η , there exists $\delta > 0$ such that

$$\limsup_{n \rightarrow \infty} \sup_{\theta \leq \delta} P(|A_n(\tau_n + \theta) - A_n(\tau_n)| > \epsilon) \leq \eta \quad (6.50)$$

Similarly, as $\varphi(x)$ is bounded, we have for all $n \geq 1$

$$\begin{aligned} & E(|\ll M_n \gg_{\tau_n + \theta} - \ll M_n \gg_{\tau_n}|^2) \\ &= E(|\frac{1}{n} \sum_{j=1}^n \int_{\tau_n}^{\tau_n + \theta} \varphi^2(x_j^{(n)}(s)) ds|^2) \\ &\leq \frac{1}{n} \sum_{j=1}^n \delta E \int_{\tau_n}^{\tau_n + \theta} \varphi^4(x_j^{(n)}(s)) ds \\ &\leq \delta^2 c_1 \\ &\rightarrow 0 \text{ as } \delta \rightarrow 0. \end{aligned} \quad (6.51)$$

This proves the tightness of $(\mathcal{L}(x_1^{(n)}))_{n=1}^{\infty}$, and hence the tightness of $(Q_n)_{n=1}^{\infty}$. $\square$

Next we show that any limit point of $(Q_n)_{n=1}^\infty$ concentrates on the subset of $\mathbf{M}_1(C([0, \infty), R))$ consisting of solutions to the following martingale problem.

For $m \in \mathbf{M}_1(R)$ and $f \in C_K^2(R)$ consider the following operator

$$L(m)f(x) = \frac{1}{2} \langle m, \varphi^2 \rangle f''(x) + b(x)f'(x) \quad (6.52)$$

Let $X : C([0, \infty), R) \rightarrow R$ be the usual projective map and $\mathcal{B}_t, t \geq 0$, be the family of sigma algebras on $C([0, \infty), R)$ generated by X . Let $u \in \mathbf{M}_1(R)$. A probability measure m in $\mathbf{M}_1(C([0, \infty), R))$ is said to be a solution to the martingale problem $(L(m), u)$ if the following conditions are satisfied:

1. X has the initial distribution u under the measure m

$$m(X(0) \in \cdot) = u(\cdot), \quad (6.53)$$

2. for $f \in C_K^2(R)$ the process

$$M^{f,m}(t) = f(x(t)) - f(x(0)) - \int_0^t L(m(s))f(x(s))ds \quad (6.54)$$

is a $(\mathcal{B}_t, m)$ -martingale.

Let S_u be the set of all solutions to the martingale problem $(L(m), u)$.

Proposition 6.4 Suppose that Q is a limit point of $(Q_n)_{n=1}^\infty$. If $X_n(0) \Rightarrow u$ in $\mathbf{M}_1(R)$ as $n \rightarrow \infty$, then Q concentrates on the set S_u .

Proof: For each fixed $f \in C_K^2(R)$ and $0 \leq s \leq t$ we define the map

$$F : \mathbf{M}_1(C([0, \infty), R)) \rightarrow [0, \infty)$$

such that

$$F(m) = \langle m, [M^{f,m}(t) - M^{f,m}(s)]G \rangle^2 \quad (6.55)$$

where G is a bounded continuous function on $C([0, \infty), R)$ which is $\mathcal{B}_s$ -measurable.

For $K \geq 1$

$$\begin{aligned}
 & \langle F(\cdot) \wedge K, Q \rangle \\
 &= \lim_{n \rightarrow \infty} \langle F(\cdot) \wedge K, Q_n \rangle \\
 &= \lim_{n \rightarrow \infty} E \left\{ \left[\frac{1}{n} \sum_{i=1}^n [M^{f, X_n}(t)(x_i^{(n)}) - M^{f, X_n}(s)(x_i^{(n)})] G(x_i^{(n)}) \right]^2 \wedge K \right\} \\
 &= \lim_{n \rightarrow \infty} [I_{1,K}^{(n)} + I_{2,K}^{(n)}] \tag{6.56}
 \end{aligned}$$

where

$$I_{1,K}^{(n)} = \frac{1}{n^2} \sum_{i=1}^n E \{ ([M^{f, X_n}(t)(x_i^{(n)}) - M^{f, X_n}(s)(x_i^{(n)})] G(x_i^{(n)}))^2 \wedge K \} \tag{6.57}$$

and

$$\begin{aligned}
 I_{2,K}^{(n)} &= \frac{1}{n^2} \sum_{i \neq j} E \{ [M^{f, X_n}(t)(x_i^{(n)}) - M^{f, X_n}(s)(x_i^{(n)})] \\
 &\quad [M^{f, X_n}(t)(x_j^{(n)}) - M^{f, X_n}(s)(x_j^{(n)})] G(x_i^{(n)}) G(x_j^{(n)}) \wedge K \} \tag{6.58}
 \end{aligned}$$

Since f and φ are bounded

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} I_{1,K}^{(n)} \\
 &\leq c_1 \limsup_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n E(|\int_s^t \frac{1}{\sqrt{n}} \sum_{j=1}^n f'(x_i^{(n)}(u)) \varphi(x_j^{(n)}(u)) dw_{1j}(u)|^2) \\
 &= c_2 \limsup_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n E(\int_s^t \frac{1}{n} \sum_{j=1}^n |f'(x_i^{(n)}(u)) \varphi(x_j^{(n)}(u))|^2 du) \\
 &= 0 \tag{6.59}
 \end{aligned}$$

Also for $i \neq j$

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} I_{2,K}^{(n)} \\
 &\leq c_3 \limsup_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i \neq j} E([M^{f, X_n}(t)(x_i^{(n)}) - M^{f, X_n}(s)(x_i^{(n)})] \\
 &\quad [M^{f, X_n}(t)(x_j^{(n)}) - M^{f, X_n}(s)(x_j^{(n)})]) \\
 &= c_4 \limsup_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i \neq j} E(\int_s^t \frac{1}{\sqrt{n}} \sum_{k=1}^n f'(x_i^{(n)}(u)) \varphi(x_k^{(n)}(u)) dw_{ik}(u) \\
 &\quad \int_s^t \frac{1}{\sqrt{n}} \sum_{k=1}^n f'(x_j^{(n)}(u)) \varphi(x_k^{(n)}(u)) dw_{jk}(u))
 \end{aligned}$$

$$\begin{aligned} & \int_s^t \frac{1}{\sqrt{n}} \sum_{l=1}^n f'(x_j^{(n)}(u)) \varphi(x_l^{(n)}(u)) dw_{jl}(u) \\ &= 0 \end{aligned} \quad (6.60)$$

Therefore

$$\langle F(\cdot) \wedge K, Q \rangle = 0 \quad (6.61)$$

Let K tend to infinity we have

$$F(m) = 0 \quad Q\text{-a.s.} \quad (6.62)$$

Hence Q concentrates on the measures $m \in \mathbf{M}_1(C([0, \infty), R))$ such that the process $M^{f,m}$ defined by (6.54) is a $(\mathcal{B}_t, m)$ -martingale. If $X_n(0) \Rightarrow u$ in $\mathbf{M}_1(R)$, then Q concentrates on the measures $m \in \mathbf{M}_1(C([0, \infty), R))$ such that m is a solution of the martingale problem $(L(m), u)$. $\square$

Finally, we show that the martingale problem has a unique solution.

Proposition 6.5 *The martingale problem $(L(m), u)$ has a unique solution μ which is the law of the unique solution $\{x(t) : t \geq 0\}$ of the equation:*

$$dx(t) = b(x(t))dt + \sqrt{\langle \mu(t), \varphi^2 \rangle} dw(t) \quad (6.63)$$

$\mu(t)$ is the distribution of $x(t)$

$x(0)$ has distribution u

Proof: Let μ be the law of the unique solution of (6.63). Then Itô's formula implies that μ is a solution of the martingale problem. Suppose that μ and ν are two solutions of the martingale problem. Then it is easy to see that the laws of the corresponding solutions of the equation (6.63) with μ as a parameter are necessarily the same as μ and ν , respectively. Since the non-linear equation (6.63) has a unique solution, we must have $\mu = \nu$. Therefore the martingale problem has a unique solution. $\square$

This completes the proof of the theorem 6.1. In the next section we prove the theorem 6.2.

6.4 Stationary Gaussian solutions

In this section we consider the equilibrium behavior of the following system of stochastic differential equations in R :

$$dx_i^{(n)}(t) = -x_i^{(n)}(t)dt + \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi(x_j^{(n)}(t))dw_{ij}(t), \quad (6.64)$$

$$i = 1, \dots, n,$$

where

$$\varphi(x) = \sqrt{2} \tanh(\beta x), \quad \beta > 0 \quad (6.65)$$

Let

$$X_n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i^{(n)}} \quad (6.66)$$

be the empirical measure associated with the unique solution $(x_1^{(n)}(\cdot), \dots, x_n^{(n)}(\cdot))$ of the equation (6.64). In the limit $n \rightarrow \infty$, by theorem 6.1, $(X_n)_{n=1}^{\infty}$ converges in probability to the deterministic measure μ which is the law of the unique solution of the following stochastic differential equation:

$$dx(t) = -x(t)dt + \sqrt{\langle \mu(t), 2 \tanh^2(\beta x) \rangle} dw(t) \quad (6.67)$$

$$\mu(t) = \text{the probability distribution of } x(t) \quad (6.68)$$

The equilibrium measure of the limit process (6.67) is given by

$$P_\sigma(x) = \frac{1}{Z_\sigma} \exp\left\{-\frac{1}{2\sigma} x^2\right\} \quad (6.69)$$

where Z_σ is the normalization factor

$$Z_\sigma(x) = \int_R \exp\left\{-\frac{1}{2\sigma} x^2\right\} dx \quad (6.70)$$

and σ satisfies the following self-consistent equation

$$\sigma = \frac{1}{Z_\sigma} \int_R \tanh^2(\beta x) \exp\left\{-\frac{1}{2\sigma} x^2\right\} dx \quad (6.71)$$

Using the fact that $Z_\sigma = \sqrt{2\pi\sigma}$ and a change of integral variable, the self-consistent equation (6.71) can be easily reduced to the following one

$$\begin{aligned} \sigma &= \int_R \tanh^2(\beta x) \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{1}{2\sigma} x^2\right\} dx \\ &= \int_R \tanh^2(\beta\sqrt{\sigma}x) \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} x^2\right\} dx \end{aligned} \quad (6.72)$$

We are interested in the positive solutions of this equation.

Proposition 6.6 *There exists a critical value $\beta_c = 1$ such that*

1. *for $\beta \leq \beta_c$, the fixed point equation (6.72) has only a trivial solution $\sigma = 0$. It has no positive solution.*
2. *for $\beta > \beta_c$, the fixed point equation (6.72) has a unique positive solution $\sigma > 0$.*

Proof: Let

$$F(\sigma) = \int_R \tanh^2(\beta\sqrt{\sigma}x) \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} x^2\right\} dx \quad (6.73)$$

Then clearly

$$|F(\sigma)| \leq \int_R |\tanh^2(\beta\sqrt{\sigma}x)| \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} x^2\right\} dx < 1 \quad (6.74)$$

because $|\tanh(z)| < 1$. Thus the solutions to (6.72) can only be in $[0, 1)$.

Since the Gaussian measure

$$\frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{1}{2\sigma} x^2\right\} dx \Rightarrow \delta_0 \quad (6.75)$$

as $\sigma \rightarrow 0$, a standard weak convergence argument gives

$$\lim_{\sigma \rightarrow 0} F(\sigma) = 0 \quad (6.76)$$

Thus $\sigma = 0$ is always a solution of the equation (6.72) for any $\beta > 0$.

The first derivative of $F(\sigma)$ is

$$\begin{aligned} F'(\sigma) &= \int_R \tanh(\beta\sqrt{\sigma}x) \operatorname{sech}^2(\beta\sqrt{\sigma}x) \beta\sigma^{-\frac{1}{2}} x \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx \\ &= \int_R \frac{\tanh(\beta\sqrt{\sigma}x)}{\beta\sqrt{\sigma}x} \operatorname{sech}^2(\beta\sqrt{\sigma}x) \beta^2 x^2 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx \end{aligned} \quad (6.77)$$

which is clearly positive for $\sigma > 0$. Since

$$\lim_{\sigma \rightarrow 0} \frac{\tanh(\beta\sqrt{\sigma}x)}{\beta\sqrt{\sigma}x} = 1 \quad (6.78)$$

and

$$\lim_{\sigma \rightarrow 0} \operatorname{sech}^2(\beta\sqrt{\sigma}x) = 1 \quad (6.79)$$

we obtain from (6.77) that

$$\lim_{\sigma \rightarrow 0} F'(\sigma) = \beta^2 \quad (6.80)$$

To compute the second derivative of $F(\sigma)$ we write $F'(\sigma)$ as follows

$$F'(\sigma) = \int_R f(\sigma, x) g(\sigma, x) dx \quad (6.81)$$

where

$$f(\sigma, x) = \frac{\tanh(\beta\sqrt{\sigma}x)}{\beta\sqrt{\sigma}x} \quad (6.82)$$

and

$$g(\sigma, x) = \operatorname{sech}^2(\beta\sqrt{\sigma}x) \beta^2 x^2 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} \quad (6.83)$$

Then

$$\frac{\partial f(\sigma, x)}{\partial \sigma} = \frac{1}{2} \operatorname{sech}^2(\beta\sqrt{\sigma}x) \sigma^{-1} - \frac{1}{2} \tanh(\beta\sqrt{\sigma}x) \beta^{-1} \sigma^{-\frac{3}{2}} x^{-1} \quad (6.84)$$

and

$$\frac{\partial g(\sigma, x)}{\partial \sigma} = -\operatorname{sech}^2(\beta\sqrt{\sigma}x) \tanh(\beta\sqrt{\sigma}x) \beta^3 \sigma^{-\frac{1}{2}} x^3 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} \quad (6.85)$$

This gives

$$\begin{aligned}
 F''(\sigma) &= \int_R \left[\frac{\partial f(\sigma, x)}{\partial \sigma} g(\sigma, x) + f(\sigma, x) \frac{\partial g(\sigma, x)}{\partial \sigma} \right] dx \\
 &= \int_R \operatorname{sech}^4(\beta \sqrt{\sigma} x) \frac{1}{2} \beta^2 \sigma^{-1} x^2 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx \\
 &\quad - \int_R \tanh(\beta \sqrt{\sigma} x) \operatorname{sech}^2(\beta \sqrt{\sigma} x) \frac{1}{2} \beta \sigma^{-\frac{1}{2}} x \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx \\
 &\quad - \int_R \tanh^2(\beta \sqrt{\sigma} x) \operatorname{sech}^2(\beta \sqrt{\sigma} x) \beta^2 \sigma^{-1} x^2 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx \\
 &\leq -\frac{1}{2} \int_R \operatorname{sech}^2(\beta \sqrt{\sigma} x) \beta \sigma^{-\frac{1}{2}} x \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} \\
 &\quad [\tanh(\beta \sqrt{\sigma} x) \sigma^{-\frac{1}{2}} - \operatorname{sech}^2(\beta \sqrt{\sigma} x) \beta x] dx \\
 &= -\int_0^\infty \operatorname{sech}^2(\beta \sqrt{\sigma} x) \beta \sigma^{-\frac{1}{2}} x \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} \\
 &\quad [\tanh(\beta \sqrt{\sigma} x) \sigma^{-\frac{1}{2}} - \operatorname{sech}^2(\beta \sqrt{\sigma} x) \beta x] dx \\
 &\leq 0
 \end{aligned} \tag{6.86}$$

In the last inequality above we used the fact that

$$\frac{\tanh(\beta \sqrt{\sigma} x) \sigma^{-\frac{1}{2}}}{\operatorname{sech}^2(\beta \sqrt{\sigma} x) \beta x} = \frac{\sinh(\beta \sqrt{\sigma} x)}{\beta \sqrt{\sigma} x} \cosh(\beta \sqrt{\sigma} x) \geq 1 \tag{6.87}$$

i.e.,

$$\tanh(\beta \sqrt{\sigma} x) \sigma^{-\frac{1}{2}} - \operatorname{sech}^2(\beta \sqrt{\sigma} x) \beta x \geq 0 \tag{6.88}$$

The equations (6.76), (6.77), (6.80) and (6.86) together shows that for $\sigma \in [0, \infty)$ the function $F(\sigma)$ is a monotone increasing concave function vanishing at 0 such that its tangent equals β^2 at $\sigma = 0$. Thus for $\beta \leq \beta_c = 1$, $\sigma > F(\sigma)$ for all $\sigma > 0$, i.e., the fixed point equation has no positive solution. But for $\beta > \beta_c = 1$, the two curves σ and $F(\sigma)$ intersect at some point $\sigma > 0$, so fixed point equation has a unique positive solution. $\square$

Bibliography

- [1] D.J.Amit, H.Gutfreund and H.Sompolinsky, *Spin-glass models of neural networks*, Phys. Rev., A32, 1985, pp 1007-1018.
- [2] L.L.Bonilla, *Stable nonequilibrium probability densities and phase transitions for mean-field models in the thermodynamic limit*, J. Stat. Phys., 46, 1987, pp 659-678.
- [3] L.L.Bonilla and J.M.Casado, *Dynamics of a soft-spin van Hemmen model. I. phase and bifurcation diagrams for stationary distributions*, J. Stat. Phys., 56, 1989, pp 113-125.
- [4] F.Comets, *Large deviation estimates for a conditional probability distribution, applications to random interaction Gibbs measures*, Probability Theory and Related Fields, 80, 1989, pp 407-432.
- [5] D.A.Dawson, *Critical dynamics and fluctuations for a mean-field model of cooperative behavior*, J. Stat. Phys., 31, 1983, pp 29-85.
- [6] D.A.Dawson and J.Gärtner, *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, Stochastics, 20, 1987, pp 247-308.
- [7] D.A.Dawson and J.Gärtner, *Large deviations, free energy functional and quasi-potential for a mean field model of interacting diffusions*, Memoirs of the American Mathematical Society, 398, 1989.
- [8] B.Derrida, *Random energy model: an exactly solvable model of disordered systems*, Phys. Rev., B24, 1981, pp 2613-2626.
- [9] B.Derrida, *A generalization of the random energy model which includes correlations between energies*, J. Phys. Lett., 46, 1985, pp 2401-2407.
- [10] R.L.Dobrushin, *Prescribing a system of random variables by conditional distributions*, Theory Prob. Appl., 15, 1970, pp 458-486.

- [11] C. De Dominicis and I.Kondor, *Eigenvalues of the stability matrix for Parisi solution of the long-range spin-glass*. Phys. Rev., B27, 1983, pp 606-608.
- [12] C. De Dominicis, *Dynamics as a substitute for replicas in systems with quenched random impurities*. Phys. Rev., B18, 1978, pp 4913-4919.
- [13] R.M.Dudley, *Distance of probability measures and random variables*. Ann. Math. Statist., 39, No.5, 1968, pp 1563-1572.
- [14] E.B.Dynkin, *Markov processes*, I,II, 1965, Springer-Verlag.
- [15] S.F.Edwards and P.W.Anderson, *Theory of spin glasses* , J. Phys., F5, 1975, pp 965-974.
- [16] T.Eisele, *On a third-order phase transition*. Commun. Math. Phys., 90, 1983, pp 125-159.
- [17] R.S.Ellis, *Entropy, large deviations, and statistical Mechanics*, 1985, Springer-Verlag.
- [18] S.N.Ethier and T.G.Kurtz, *Markov process: characterization and convergence*, 1986, Wiley.
- [19] M.I.Freidlin and A.D.Wentzell, *Random perturbations of dynamical systems*, 1984, Springer-Verlag.
- [20] J.Gärtner, *On large deviations from the invariant measure*. Theory of probability and its applications, XXII, No.1, 1977, pp 24-39.
- [21] J.Gärtner, *On the McKean-Vlasov limit for interacting diffusions I, II*. Akademie Der Wissenschaften Der DDR, Karl-Weierstrass-Institute Für Mathematik, Berlin, 1986.
- [22] S.Geman and C.R.Hwang, *A chaos hypothesis for some large systems of random equations*, Zeit. Wahr. Verw. Geb., 60, 1982, pp 291-314.
- [23] S.Geman, *Almost sure stable oscillation in a large system of randomly coupled equations*, SIAM J. Appl. Math., 42, 1984, pp 695-703.

- [24] J.J.Hopfield. *Neural networks and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. USA, **79**, 1982, pp 2554-2558.
- [25] C.R.Hwang, *Laplace's method revisited : weak convergence of probability measures*, The Annals of probability, **8**, No.6, 1980, pp 1177-1182.
- [26] N.Ikeda and S.Watanabe. *Stochastic differential equations and diffusion processes*, North-Holland/Kodansha, 1989, Second Edition.
- [27] A.Joffe and M.Metivier, *Weak convergence of sequences of semimartingales with applications to multitype branching processes*, Adv. in Appl. Prob., **18**, 1986, pp 20-65.
- [28] S.Kusuoka and Y.Tamura, *Gibbs measures for mean field Potentials*, J. Fac. Sci. Univ. Tokyo, Sect. 1A, math., **34**, 1984, pp 223-245.
- [29] C.Léonard, *Une loi des grands nombres pour des systèmes de diffusions avec interaction à coefficients nonbornés*, Ann. Inst. Henri Poincaré, **22**, 1984, pp 237-262.
- [30] C.Léonard, *Large deviations and law of large numbers for a mean field type interacting particle system*, Stochastic Processes and their Applications, **25**, 1987, pp 215-235.
- [31] H.P.McKean, *A class of Markov processes associated with nonlinear parabolic equations*, Proc. N.A.S., **56**, 1966, pp 1907-1911.
- [32] H.P.McKean, *Propagation of chaos for a class of nonlinear parabolic equations*, Lecture Series in Differential equation, **7**, Catholic University, 1969, pp 41-57, Van Nostrand Reinhold.
- [33] M.Mézard, G. Parisi and M.A.Virasoro, *Spin glass theory and beyond*, World Scientific Lecture Notes in Physics, **9**, 1987, World Scientific Publishing Co.Pte.Ltd.
- [34] K.Oeschläger, *A martingale approach to the law of large numbers for weakly interacting stochastic processes*, Ann. Prob., **12**, No.2, 1984, pp 458-479.

- [35] G.Parisi, *Toward a mean field theory for spin glasses*, Phys. Lett., A73, 1979, pp 203-205.
- [36] G.Parisi, *Order parameter for spin glasses*, Phys. Rev. Lett., A50, 1983, pp 1946-1948.
- [37] P.Picco, *On the R.E.M. and the G.R.E.M. statistical physics, automata networks and dynamical systems*, Edited by E.Goles and S. Martinez. Kluwer Academic Publishers, 1992, pp 173-207.
- [38] D.Revuz and M.Yor, *Continuous martingales and Brownian motion*, 1991, Springer-Verlag.
- [39] H.Rieger, *Solvable model of a complex ecosystem with randomly interacting species*, J. Phys., A22, 1989, pp 3447-3460.
- [40] R.T.Rockafellar, *Convex Analysis*, 1972, Princeton University Press.
- [41] D.Sherrington and S.Kirkpatrick, *Solvable model of a spin-glass*, Phys. Rev. Lett., 35, 1975, pp 1792-4403.
- [42] H.Sompolinsky and A.Crisanti, *Chaos in random neural networks*, Phy. Rev. Letters, 61, No.3, 1988, pp 259-262.
- [43] H.Sompolinsky and A.Zippelius, *Relaxation dynamics of the Edwards-Anderson model and the mean-field theory of spin-glasses*, Phy. Rev., B25, 1982, pp 6860-6875.
- [44] R.L.Stratonovich, *On a method of calculating quantum distribution functions*, Sov. Phys. Dokl., 2, 1958, pp 416.
- [45] D.W.Stroock, *An introduction to the theory of large deviations*, 1984, Springer-Verlag.
- [46] A.Sznitman, *Topics in propagation of chaos*, Ecole d'Eté de Probabilités de Saint-Flour XIX-1989, Lecture Notes in Mathematics 1464.

- [47] **H.Takayama, T.Shirakura and K.Nemoto**, *Dynamical mean field theory of spin glasses and their phase transitions in ac external fields*, Progress of Theoretical Physics Supplement., **87**, 1986, pp 191-213.
- [48] **Y.Tamura**, *On asymptotic behaviors of the solution of a non-linear diffusion equation*, J. Fac. Sci. Univ. Tokyo, Sect. 1A, math., **31**, 1984, pp 195-221.
- [49] **Y.Tamura**, *Free energy and the convergence of distributions of diffusion processes of McKean type*, J. Fac. Sci. Univ. Tokyo, Sect. 1A, Math., **34**, 1987, pp 443-484.
- [50] **D.J.Thouless, P.W.Anderson and R.G.Palmer**, *Solutions of a Solvable model of a spin glass*, Phil. Mag., **35**, 1977, pp 593-601.
- [51] **J.Vaillancourt**, *Interactive Fleming-Viot processes and related measure-valued processes*, Ph.D. Thesis, Depart. of Math. and Stat., Carleton University, 1987.
- [52] **J.L.van Hemmen, D.Greising, A.Huber and R.Kuhn**, *Nonlinear neural networks I, II*, J. of Stat., **50**, No.1,2, 1989, pp 231-293.
- [53] **J.L.van Hemmen and R.G.Palmer**, *The replica method and a solvable spin glass model*, J. Phys. A: Math. Gen., **12**, No.4, 1979, pp 563-580.
- [54] **S.R.S.Varadhan**, *Asymptotic probabilities and differential equations*, Communications on pure and applied mathematics, **XIX**, 1966, pp 261-286.

END

14-12-93

FIN