Bayesian Statistics, Statistical Physics, and Monte Carlo Methods: An interdisciplinary study

Yukito Iba

Contents

Chapter 1

Preamble

1.1 Preface

This thesis consists of a set of studies in statistical physics and statistical sciences. Formally, they are classified into three categories, studies on an analogy between Bayesian statistics and statistical mechanics (Chaps. 3,4,5,6 and 8), computer design of microscopic objects (Chap. 7), and design of Monte Carlo algorithms with Extended Ensembles (Chaps. 9 and 10). There are close relationships among these subjects. Statistical inference and computer design are special examples of inverse problems, to which we coherently apply probabilistic frameworks. Monte Carlo algorithms play essential roles in many of these chapters. Chaps 3,4 and 5 deal with the application of Monte Carlo algorithms in statistical inference. In Chaps 9 and 10, I discuss our contributions to a new trend of Dynamical Monte Carlo algorithms, i.e., Extended Ensemble Monte Carlo.

Keywords of my study are *analogy* and *algorithm*. The analogies between different fields are in themselves a core of my interest. I am also interested in the way they contribute each of the fields. Algorithm is also a key in cross-disciplinary studies, because studies on very different subjects can have common aspects in the algorithm level. My study is also regarded as an expression of my love for equilibrium statistical physics. I want to show how the methods of statistical physics works in other fields and how we can extend these methods.

This thesis contains works in the period 1987-1999. My interest in statistical sciences began around 1987, when I came to know the activity at the Institute of Statistical Mathematics and read a paper [1] by Geman and Geman on Bayesian image restoration. For 1987-1989, I gave a series of talks on the analogy between Bayesian statistics and statistical physics at the meetings of the physical society of Japan and Yukawa International Symposium (YKIS88') [2]. Chaps. 3 and 4 are based on the works published in 1991 in Japanese. Chap. 5 deals with the subject presented in the third pacific area statistical conference (1991) and appeared as an unpublished technical report (1992). The present version is based on a manuscript rewritten in 1999. Chap. 6, 7 and 8 are based on recent works that appeared in journals or proceedings in 1998-1999.

My interest on artificial ensembles designed for computation also goes back to the age when I was a graduate student. But I began with serious studies in this field after I read the seminal papers on simulated tempering [3] and multicanonical algorithm [4, 5]. My first contribution in this field is re-discovery of Metropolis-coupled chain (Parallel Tempering, Exchange Monte Carlo) algorithm at 1993 [6], but I decided not to publish it because I was not sure of my priority on this topic [1]. In this thesis, two topics from my recent study in the field are included – Chaps. 9 and 10. They are based on the papers published in 1998 and 1999.

In these twelve years, the situation around the topics of my interest drastically changed. Even at the end of 1980s, there were researchers who recognize central concepts in this field – the analogy between Bayesian statistics and statistical physics, the utility of Dynamical Monte Carlo methods in statistical sciences, and the possibility of artificial ensembles as a tool for computation. However, in the following ten years, a large number of researchers in various fields recognized the importance of these concepts and joined the study of this area. This process is still continued – for example, the concept of finite temperature information processing have been drawing attention of statistical physicists in a neural network community for the past five years.

Materials that I presented here as a thesis are pieces of a dream of the days when I was young. I feel that my contributions through this period of innovation are not so much as I expected. I hope that I can do better in the next decade, which will also be the days of revolution.

1.2 Acknowledgements

In the first place, I would acknowledge Prof. Akutsu, who gave me kind encouragement and useful advices for preparing the thesis. The studies in Chaps. 7 and 9 are the results of joint research with Mr. Chikenji (Osaka Univ.), Dr. Tokita (Osaka Univ.) and Prof. Kikuchi (Osaka Univ.). I would acknowledge their contributions and kind treatment that allows the inclusion of these works to this thesis. I am also grateful to Dr. Hukushima (Univ. Tokyo, ISSP) and Dr. Kabashima (Tittech.) for kind advices and critical reading of the manuscript. Dr. Hukushima is also kind enough to allow me the reproduction of an unpublished result of the joint research in Chap. 11. I would greatly appreciate kind encouragement and advices from Prof. Amari (RIKEN), Prof. Tanabe (ISM), and Dr. Kawato (ATR), which gave me great help in the earlier period of the research. Specifically, Prof. Tanabe gave me many important advices for the preparation of an earlier version of Chap. 5.

[1] An English version of [6] is included in Chap. 11 as an appendix. It is earlier than Hukushima and Nemoto (1994, published in 1996 [7]), but later than Geyer(1991) [8], and Kimura and Taki(1990) [9]. And I did not give a full-scale application in statistical physics (Hukushima and Nemoto *did* it.).

1.3 Organization of the Paper

This thesis consists of three parts — this preamble (Chap. 1), a survey on backgrounds and motivations of our studies (Chap. 2), and chapters on our original contributions (Chap. 3-Chap. 10). There are also some appendices (Chap. 11). Outlines of Chap. 2 and Chap. 3-Chap. 10 are shown in the following.

1.3.1 Outline of the Review Part (Chap.2)

In Chap.2, we give (1) a survey on backgrounds on statistical sciences (Sec. 2.1), and (2) a mini-review of Extended Ensemble Monte Carlo algorithms (Sec. 2.2), both of which are useful for the understanding of original contributions in later chapters. In a survey in Sec. 2.1, we systematically introduce the basic notions of modern statistical sciences with special emphasis on recent revival of Bayesian frameworks and an analogy to statistical physics. It gives a basis for the understanding of Chaps. 3–6 and Chap. 8. The survey also contains an introduction to Chap. 7, where we discuss computer design of lattice proteins. On the other hand, we give an introduction to Extended Ensemble Monte Carlo algorithms in Sec. 2.2, which is a subject of the studies in Chaps. 9 and 10.

An extraordinary length of the review part, Chap.2, is partly due to the theme of this thesis, which is not familiar to most physicists. Another aim of the long introduction is to show the hidden relations among seemingly divergent subjects in Chap. 3 –10, which are the outcome of the study in these twelve years. In the surveys, we make efforts to keep close contact to the contributions in later chapters. On the other hand, these surveys are also designed as a concise review of these fields of current interest. I would be happy if experts of these subjects find something fresh in the surveys.

In this thesis, we expect that the readers have some basic knowledge on concepts, methods, and models (e.g., spin glass models, lattice heteropolymers) in statistical physics. We make efforts to include the derivations of algorithms and mathematical definitions of models, but it is difficult to include comprehensive background on physics. A concise review on concepts and models in statistical physics are found in an unpublished technical report by the author [10]. The foundation of Dynamical Monte Carlo algorithms are also discussed in a survey paper by the author [11]. They might be useful to fill the gaps.

1.3.2 Summaries of the Contributions in Chap.3-Chap.10

Here we give summaries of chapters that deal with our original contributions. Each of these chapters is based on an independent paper published in (or will be submitted to) a journal or proceedings. The abstract of the original paper, which gives more specific details on the study, is reproduced at the beginning of each chapter.

§ **Chap. 3**
Estimation of the Coupling Constant of an Ising Model
This is the first of the three chapters that deal with application of Dynamical Monte Carlo algorithms to problems in statistical sciences. In this chapter, we study an inverse problem for Ising model, i.e., the estimation of the coupling constant of an Ising model from a pattern (snapshot) corrupted by noise. It is shown that the coupling constant and the strength of the noise are simultaneously estimated by a maximum marginal likelihood procedure with a Dynamical Monte Carlo algorithm for calculating posterior and prior averages. Relevance to image restoration problems is also discussed.

Publication:
Iba,Y. 1991 (in Japanese)
Macroscopic Parameter Estimation from Incomplete Data with Metropolis-type Monte Carlo Algorithm,
Proceedings of the Institute of Statistical Mathematics, **39**, 1-21, 1991.

Background:

Sec. 2.1.3.1:	Image Restoration, Ising Prior
Sec. 2.1.2.4:	Boltzmann Machine Learning Equation
Sec. 2.1.4.3:	Dynamical Monte Carlo Algorithms (with a historical note on application in statistical sciences)
Sec. 2.1.4:	Statistical Physics and Bayesian Statistics (in general)
Sec. 2.1.2:	Bayesian Framework (in general)

§ **Chap. 4**
Estimation of Change Points
In this chapter, we discuss the application of Dynamical Monte Carlo algorithms to the estimation of the positions and number of change points in linearly ordered data. We use model mixing approach base on the quasi-Bayesian method of Akaike and show that Dynamical Monte Carlo algorithms are useful for the calculation of posterior probabilities of the location of change points. Both of real-world data and artificial data are treated by the proposed method.

Publication:
Iba,Y. 1991 (in Japanese)
Metropolis-type Monte Carlo Algorithm and Quasi-Bayesian Estimation Procedure: An Application to a Change Point Problem,
Proceedings of the Institute of Statistical Mathematics, **39**, 225-244, 1991.

Background:

Sec. 2.1.3.2:	Change Point Detection
Sec. 2.1.1.3:	AIC, Model Selection
Sec. 2.1.4.3:	Dynamical Monte Carlo Algorithms (with a historical note on application in statistical sciences)
Sec. 2.1.1.1:	Maximum Likelihood Estimate (in general)
Sec. 2.1.4:	Statistical Physics and Bayesian Statistics (in general)
Sec. 2.1.2:	Bayesian Framework (in general)

§ **Chap. 5**
Bayesian Classification with Relational Data
This chapter also treats an analogy between statistical physics and Bayesian statistics and an application of Dynamical Monte Carlo algorithm to statistical sciences. Here we introduce a set of models for data analysis closely related to Ising and/or Potts spin models. These models are an extension of finite mixture models and designed for "classification with relational data", where data generation process is assumed to depend upon a pair of the objects to be classified. It is shown that a Dynamical Monte Carlo algorithm is successfully used for Bayesian inference with these models, e.g., assessment of the validity of estimates, calculation of an optimal estimator, and estimation of a hyperparameter.

Publication:
To be submitted. Corresponding Technical Reports are :
ISM Research Memo No.731 (1999)
ISM Research Memo No.440 (1992)

Background:

Sec. 2.1.3.3:	Finite Mixture Model
Sec. 2.1.4.3:	Dynamical Monte Carlo Algorithms (with a historical note on application in statistical sciences)
Sec. 2.1.4:	Statistical Physics and Bayesian Statistics (in general)
Sec. 2.1.2:	Bayesian Framework (in general)

§ **Chap. 6**
Mean Field Approximation in Variable Selection
In this chapter, we discuss an application of another family of methods in statistical physics, Mean Field Approximation, to statistical sciences. Here we treat a model mixing (model averaging) problem by a mean field approximation. That is, multiple regression models with different number of explanatory variables are expressed by a set of binary indicators ("Ising spins") and posterior averages over the set of models are calculated by a mean field approximation. Application to real world data, Boston housing data, is shown.

Publication:
Iba,Y. 1998
Mean Field Approximation in Bayesian Variable Selection,

Proceedings of ICONIP'98, **1**, 530-533, 1998 (ed. S.Usui and T.Omori, Ohmsha, Tokyo and IOS Press, Burke VA USA).

Background:

Sec. 2.1.1.3:	AIC, Model Selection
Sec. 2.1.3.5:	Model Mixing
Sec. 2.1.3.4:	Multiple Regression
Sec. 2.1.4.4:	Mean Field Approximation (with a note on application in statistical sciences)
Sec. 2.1.4:	Statistical Physics and Bayesian Statistics (in general)
Sec. 2.1.2:	Bayesian Framework (in general)

§ **Chap. 7**
Design of Lattice Proteins
In this chapter, we study a problem of computer biology, computer design or "inverse folding" of protein. While three-dimensional conformations is asked for a given sequence of amino acids in the *folding* problem, a sequence that folds into a given conformation is requested in the *inverse folding* problem. Here, starting from an analogy to statistical science, we develop a computational method for this problem. We test the proposed method for a simplified model known as HP model of lattice protein.

The results in this chapter are the outcome of a joint research with Kei Tokita (Osaka Univ.) and Macoto Kikuchi (Osaka. Univ.).

Publication:
Iba, Y., Tokita, K. and Kikuchi, M. 1998
Design Equation: A Novel Approach to Heteropolymer Design,
Journal of Physical Society of Japan, **67**, 3985-3990, 1998.

Background:

Sec. 2.1.1.2:	Design of Microscopic Objects
Sec. 2.1.1.1:	Maximum Likelihood Estimate (and Learning Equation)
Chap. 9:	Monte Carlo Methods for Lattice Polymers

§ **Chap. 8**
An Interpretation of the Nishimori line
In Chaps. 3 - Chap. 6, we discuss the applications of computational algorithms of statistical physics to the problems of statistical sciences. Here, we discuss an example of the use of the correspondence between statistical physics and Bayesian statistics in the reverse direction. That is, we show that a well-known rigorous result in statistical physics of disordered systems, theorems of the *Nishimori line*, has an interesting interpretation from the viewpoint of Bayesian statistics. We give a comprehensive survey on the issues on "finite temperature decoding" of error-correcting code, where a relation of gauge invariance of error-correcting code and the "equivariance" property in the statistical decision theory is shown.

Publication:
Iba, Y. 1999
The Nishimori line and Bayesian Statistics,
Journal of Physics A, Mathematical and General, **32**, 3875-3888, 1999.
Background:

Sec. 2.1.4:	Statistical Physics and Bayesian Statistics (in general)
Sec. 2.1.3.1:	Image Restoration
Chap. 3:	Image Restoration
Sec. 2.1.3.6:	Error Correcting Codes
Sec. 2.1.2.3:	Marginal Likelihood
Sec. 2.1.2.3:	Marginal Likelihood (and Learning Equation)

§ **Chap. 9**
Multi-Self-Overlap Ensemble for Lattice Heteropolymers
Contributions to Dynamical Monte Carlo algorithms are the subjects of Chaps. 9 and 10. In these chapters, we discuss Extended Ensemble Monte Carlo algorithms. In Chap. 9, we develop algorithms for the study of protein models on a lattice. The essence of our approach is the use of an artificial ensemble that contains an adequate amount of self-overlapping conformations, while such conformations are not allowed in the original models *. The inclusion of self-overlapping conformations accelerate the mixing of the Markov chain drastically at least at higher temperatures, because the path becomes able to penetrate the barriers introduced by the self-avoiding condition. If we use dynamics that satisfy the detailed balance condition and discard the self-overlapping conformations from samples for the calculation of averages, we still obtain correct canonical averages. A problem is that it is not easy to produce an adequate amount of self-overlapping conformations. Another problem is that, if we relax the self-avoiding condition, the polymers with attractive interactions between monomers collapsed at low temperatures. Here, we show that these problems are systematically treated by the methods of extended ensemble and give a family of efficient algorithms. We successfully apply an algorithm with this idea (*Multi-Self-Overlap Ensemble Monte Carlo*) to HP model of lattice protein and prove that the proposed algorithm outperforms a conventional multicanonical algorithm which uses an extended ensemble that does not contain self-overlapping conformations.

The results in this chapter are the outcome of a joint research with George Chikenji (Osaka Univ.) and Macoto Kikuchi (Osaka. Univ).

Publication:
Iba, Y., Chikenji, G. and Kikuchi, M. 1998
Simulation of Lattice Polymers with Multi-Self-Overlap Ensemble,
Journal of Physical Society of Japan, **67**, 3327-3330, 1998.

(*) After we published the above-mentioned paper (and [13]), we realize that a similar idea is discussed by Vorontsov-Velyaminov *et al.*(1996) [12].

They, however, did not apply the method to lattice heteropolymers/proteins and two-dimensional extension is not used. A Survey on algorithms with the relaxation of self-avoidingness is seen in the section "Realated Studies" of Chap. 9.

Background:

Sec. 2.2.4:	Multicanonical Algorithm, Entropic Sampling
Sec. 2.2.5 :	Special Purpose Ensembles
Sec. 2.2.6:	From Statistical Science to Computation
Sec. 2.2:	Extended Ensemble (in general)
Sec. 2.1.4.3:	Dynamical Monte Carlo Algorithm (Basics and Applications to Statistical Sciences)
Sec. 2.1.1.2:	HP model and Computer Design of Lattice Protein
Chap. 7:	HP model and Computer Design of Lattice Protein

§ **Chap. 10**
Multi-System-Size Ensemble for Spin Glass
In Chap. 10, we discuss another type of extended ensemble, *Multi-System-Size Ensemble*. We apply an algorithm with this ensemble to the SK model of spin glass and shows that it shows a better performance compared with a standard heat bath algorithm. In a simulation with Multi-System-Size Ensemble, the system size (in the present case, the number of spins) grows and decreases in a probabilistic manner within the range between zero and a given size. The detailed balance condition is satisfied and measurements at largest size give correct canonical averages. The adequate values of the penalty to the size are learned in preliminary runs in a way similar to that of multicanonical algorithm.

Publication:
Iba, Y. 1999
Simulation of Spin Glass with Multi-System-Size Ensemble,
Journal of Physical Society of Japan, **68**,1057-1058, 1999.

Background:

Sec. 2.2.4:	Multicanonical Algorithm, Entropic Sampling
Sec. 2.2.5 :	Special Purpose Ensembles
Sec. 2.2.6:	From Statistical Science to Computation
Sec. 2.2:	Extended Ensemble (in general)
Sec. 2.1.4.3:	Dynamical Monte Carlo Algorithm (Basics and Applications to Statistical Sciences)

1.4 Publications

1.4.1 List of Papers for the Requirement of the Degree

1. Iba,Y. 1989
   Bayesian Statistics and Statistical Mechanics,
   Cooperative Dynamics in Complex Physical Systems,
   235-236, ed. H. Takayama, Springer-Verlag, Berlin, 1989.
   `[Proceedings]`

2. Iba,Y. 1991
   Macroscopic Parameter Estimation from Incomplete Data with Metropolis-type Monte Carlo Algorithm,
   Proceedings of the Institute of Statistical Mathematics, **39**, 1-21, 1991.
   `[in Japanese, Paper]`

3. Iba,Y. 1991
   Metropolis-type Monte Carlo Algorithm and Quasi-Bayesian Estimation Procedure: An Application to a Change Point Problem,
   Proceedings of the Institute of Statistical Mathematics, **39**, 225-244, 1991.
   `[in Japanese, Letter]`

4. Iba, Y., Chikenji, G. and Kikuchi, M. 1998
   Simulation of Lattice Polymers with Multi-Self-Overlap Ensemble,
   Journal of Physical Society of Japan, **67**, 3327-3330, 1998.
   `[Letter]`

5. Iba, Y., Tokita, K. and Kikuchi, M. 1998
   Design Equation: A Novel Approach to Heteropolymer Design,
   Journal of Physical Society of Japan, **67**, 3985-3990, 1998.
   `[Full Paper]`

6. Iba,Y. 1998
   Mean Field Approximation in Bayesian Variable Selection,
   Proceedings of ICONIP'98, **1**, 530-533, ed. S.Usui and T.Omori, Ohmsha, Tokyo and IOS Press, Burke VA USA, 1998.
   `[Proceedings]`

7. Iba, Y. 1999
   Simulation of Spin Glass with Multi-System-Size Ensemble,
   Journal of Physical Society of Japan, **68**,1057-1058, 1999.
   `[Short Note]`

8. Iba, Y. 1999
   The Nishimori line and Bayesian Statistics,
   Journal of Physics A, Mathematical and General, **32**, 3875-3888, 1999.
   `[Full Paper]`

1.4.2 List of the Papers Added for Reference

1. Iba, Y, Akutsu, Y. and Kaneko, K. 1987
   Phase Transitions in 2-dimensional Stochastic Cellular Automata,
   Science on form (Sakura, 1985), 103-111, Reidel, Dordrecht, 1987.
   `[Proceedings]`

2. Iba,Y. 1988
   Stochastic Cellular Automaton Model for Recurrent Epidemics: Study of Spatio-Temporal Patterns,
   Proceedings of the Institute of Statistical Mathematics, **36**, 69-88, 1988.
   `[in Japanese, Paper]`

3. Iba,Y. 1995
   Gakusyu to Kaisou (Learning and Hierarchical Structures),
   Proceedings of Information Integration Workshop (IIW-95), – Beyond divide and conquer strategy –, 189-198, Real World Computing Partnership, 1995.
   `[in Japanese, Proceedings, Survey Paper]`

4. Iba,Y. 1996
   Information Integration and Fundamental Problems in Artificial Intelligence,
   Journal of Japanese Society for Artificial Intelligence, 19-26, **11**, 1996
   `[in Japanese, Invited (Tutorial) Paper]`

5. Iba, Y. and Tanaka-Yamawaki, M. 1996
   Statistical Analysis of Human Random Number Generators,
   Methodologies for the Conception, Design, and Application of Intelligent Systems, Proceedings of IIZUKA'96, **2**, 467-472, ed. T. Yamakawa, and G. Matsumoto, World Scientific, 1996.
   `[Proceedings]`

6. Iba,Y. 1996
   Markov Chain Monte Carlo Algorithms and Their Applications to Statistics,
   Proceedings of the Institute of Statistical Mathematics, **44**, 49-84, 1996.
   `[in Japanese, Research Review]`

7. Chikenji,G., Kikuchi,M. and Iba,Y. 1999
   Multi-Self-Overlap Ensemble for Protein Folding: Ground State Search and Thermodynamics,
   Physical Review Letters, **83**, 1886-1889, 1999.
   `[Letter]`

Bibliography

[1] **D. Geman and S. Geman**: Bayesian image analysis, *Disordered Systems and Biological Organization*, NATO ASI series F20, ed. E. Bienenstock et al. (Springer-Verlag, Berlin, 1986).

[2] **Y. Iba**: Bayesian statistics and statistical mechanics, *Cooperative Dynamics in Complex Physical Systems*, ed. H. Takayama (Springer-Verlag, Berlin, 1989) p.235-236.

[3] **E. Marinari and G. Parisi**: Simulated tempering: a new Monte Carlo scheme, Europhys. Lett **19** (1992) 451.

[4] **B. A. Berg and T. Neuhaus**: Multicanonical algorithms for first order phase transitions, Phys. Lett. **B267** (1991) 249.
B. A. Berg and T. Neuhaus: Multicanonical ensemble: a new approach to simulate first-order phase transitions, Phys. Rev. Lett. **68** (1992) 9-12.

[5] **B. A. Berg and T. Celik**: New approach to spin-glass simulations, Phys. Rev. Lett. **69** (1992) 2292.

[6] Brief notes on the algorithm is given in the following abstract and survey paper:
Y. Iba: Proceedings of the Institute of Statistical Mathematics **41** (1993) 65 [in Japanese]. An English version is included in Chap. 11 of this thesis as an appendix.
Y. Iba: Bussei Kenkyu (Kyoto) **60** (1993) 677 [in Japanese].

[7] **K. Hukushima and Y. Nemoto**: Exchange Monte Carlo method and application to spin glass simulations, J. Phys. Soc. Jpn. **65** (1996) 1604.
K. Hukushima, H. Takayama and H. Yoshino: Exchange Monte Carlo dynamics in the SK model, J. Phys. Soc. Jpn. **67** (1998) 12.

[8] **C. J. Geyer**: Markov chain Monte Carlo maximum likelihood, *Computing Science and Statistics*, ed. E. M. Keramides (Interface Foundation, Fairfax Station Va., 1991) p.156.

[9] **K. Kimura and K. Taki**: On a time-homogeneous parallel annealing algorithm, IEICE Technical Report **NC-90-1** (1990) 1 [in Japanese].

K. Kimura and K. Taki: Time-homogeneous parallel annealing algorithm, Proc. Comp. Appl. Math. (IMACS'91) **13** (1990) 827-828 [I have never seen this version of the paper.].
M. Hirosawa, M, R. J. Feldmann, D. Rawn, M. Ishikawa, M. Hoshida, and G. Micheals: Folding simulation using temperature parallel simulated annealing, *Proceedings of the international conference on fifth generation computer systems*(ICOT, 1992) p.300.
M. Ishikawa and M. Kanehisa: Butsuri **48** No.5 (1993) 341-343 [in Japanese].

[10] **Y. Iba**: An introduction to statistical physics for working statisticians and computer scientists, ISM Research Memorandum No.635, February, 1997 [in Japanese, third version. The original version is written in 1990.].

[11] **Y. Iba**: Markov chain Monte Carlo algorithms and their applications to statistics, Proceedings of the Institute of Statistical Mathematics **44** (1996) 49-84 [in Japanese].

[12] **P. N. Vorontsov-Velyaminov, A. V. Broukhno, T. V. Kuznetsova and A. P. Lyubartsev:** Free energy calculations by expanded ensemble method for lattice and continuous polymers, J. Phys. Chem. **100** (1996) 1153-1158.

[13] **G. Chikenji, M. Kikuchi, and Y. Iba**: Multi-self-overlap ensemble for protein folding: Ground state search and thermodynamics, Phys. Rev. Lett. **83** (1999) 1886-1889