

Springer Series in Statistics

Jun S. Liu

**Monte Carlo
Strategies in
Scientific
Computing**

科学计算中的蒙特卡罗策略

Springer

世界图书出版公司

www.wpcbj.com.cn

Jun S. Liu

Monte Carlo Strategies in Scientific Computing

With 56 Figures

Springer

世界图书出版公司

Jun S. Liu
Department of Statistics
Harvard University
Cambridge, MA 02138
jliu@hustat.harvard.edu
USA

Library of Congress Cataloging-in-Publication Data

Liu, Jun S.

Monte Carlo strategies in scientific computing / Jun S. Liu

p. cm. — (Springer series in statistics)

Includes bibliographical references and index.

ISBN 0-387-95230-6 (hc. : alk. paper)

1. Science—Statistical methods. 2. Monte Carlo method. I. Title. II. Series

Q180.55.S7 L58 2001

501.519282—dc21

00-069243

ISBN 0-387-95230-6

© 2001 Springer-Verlag New York, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

This reprint has been authorized by Springer-Verlag (Berlin/Heidelberg/New York) for sale in the People's Republic of China only and not for export therefrom.

www.springer-ny.com

Springer-Verlag New York Berlin Heidelberg
A member of BertelsmannSpringer Science+Business Media GmbH

Springer Series in Statistics

Advisors:

P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg,

I. Olkin, N. Wermuth, S. Zeger

Springer

New York

Berlin

Heidelberg

Hong Kong

London

Milan

Paris

Tokyo

Springer Series in Statistics

- Andersen/Borgan/Gill/Keiding*: Statistical Models Based on Counting Processes.
- Atkinson/Riani*: Robust Diagnostic Regression Analysis.
- Berger*: Statistical Decision Theory and Bayesian Analysis, 2nd edition.
- Borg/Groenen*: Modern Multidimensional Scaling: Theory and Applications.
- Brockwell/Davis*: Time Series: Theory and Methods, 2nd edition.
- Chan/Tong*: Chaos: A Statistical Perspective.
- Chen/Shao/Ibrahim*: Monte Carlo Methods in Bayesian Computation.
- David/Edwards*: Annotated Readings in the History of Statistics.
- Devroye/Lugosi*: Combinatorial Methods in Density Estimation.
- Efromovich*: Nonparametric Curve Estimation: Methods, Theory, and Applications.
- Eggermont/LaRiccias*: Maximum Penalized Likelihood Estimation, Volume I: Density Estimation.
- Fahrmeir/Tutz*: Multivariate Statistical Modelling Based on Generalized Linear Models, 2nd edition.
- Fan/Yao*: Nonlinear Time Series: Nonparametric and Parametric Methods.
- Farebrother*: Fitting Linear Relationships: A History of the Calculus of Observations 1750-1900.
- Federer*: Statistical Design and Analysis for Intercropping Experiments, Volume I: Two Crops.
- Federer*: Statistical Design and Analysis for Intercropping Experiments, Volume II: Three or More Crops.
- Ghosh/Ramamoorthi*: Bayesian Nonparametrics.
- Glaz/Naus/Wallenstein*: Scan Statistics.
- Good*: Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses, 2nd edition.
- Gouriéroux*: ARCH Models and Financial Applications.
- Gu*: Smoothing Spline ANOVA Models.
- Györfi/Köhler/Krzyżak/Walk*: A Distribution-Free Theory of Nonparametric Regression.
- Haberman*: Advanced Statistics, Volume I: Description of Populations.
- Hall*: The Bootstrap and Edgeworth Expansion.
- Härdle*: Smoothing Techniques: With Implementation in S.
- Harrell*: Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis.
- Hart*: Nonparametric Smoothing and Lack-of-Fit Tests.
- Hastie/Tibshirani/Friedman*: The Elements of Statistical Learning: Data Mining, Inference, and Prediction.
- Hedayat/Sloane/Stufken*: Orthogonal Arrays: Theory and Applications.
- Heyde*: Quasi-Likelihood and its Application: A General Approach to Optimal Parameter Estimation.
- Huet/Bouvier/Poursat/Jolivet*: Statistical Tools for Nonlinear Regression: A Practical Guide with S-PLUS and R Examples, 2nd edition.
- Ibrahim/Chen/Sinha*: Bayesian Survival Analysis.
- Jolliffe*: Principal Component Analysis, 2nd edition.

(continued after index)

To my wife Wei

Preface

An early experiment that conceives the basic idea of Monte Carlo computation is known as "Buffon's needle" (Dörrie 1965), first stated by Georges Louis Leclerc Comte de Buffon in 1777. In this well-known experiment, one throws a needle of length l onto a flat surface with a grid of parallel lines with spacing D ($D > l$). It is easy to compute that, under ideal conditions, the chance that the needle will intersect one of the lines is $2l/\pi D$. Thus, if we let p_N be the proportion of "intersects" in N throws, we can have an estimate of π as

$$\hat{\pi} = \lim_{N \rightarrow \infty} \frac{2l}{p_N D},$$

which will "converge" to π as N increases to infinity. Numerous investigators actually used this setting to estimate π . The idea of simulating random processes so as to help evaluate certain quantities of interest is now an essential part of scientific computing.

A systematic use of the Monte Carlo method for real scientific problems appeared in the early days of electronic computing (1945-55) and accompanied the development of the world's first programmable "super" computer, MANIAC (Mathematical Analyzer, Numerical Integrator and Computer), at Los Alamos during World War II. In order to make a good use of these fast computing machines, scientists (Stanislaw Ulam, John von Neumann, Nicholas Metropolis, Enrico Fermi, etc.) invented a statistical sampling-based method for solving numerical problems concerning random neutron diffusion in fissile material in atomic bomb designs and for estimating eigenvalues of the Schrödinger equation. The basic idea underlying

the method was first brought up by Ulam and deliberated between him and von Neumann in a car when they drove together from Los Alamos to Lamy. Allegedly, Nick Metropolis coined the name "Monte Carlo," which played an essential role in popularizing the method.

In the early 1950s, statistical physicists (N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller) introduced a Markov-chain-based dynamic Monte Carlo method for the simulation of simple fluids. This method was later extended to cover more and more complex physical systems, including spin glass models, harmonic crystal, polymer models, etc. In the 1980s, statisticians and computer scientists developed Monte-Carlo-based methods for a wide variety of tasks such as combinatorial optimizations, nonparametric statistical inference (e.g., jackknife and bootstrap), likelihood computation with missing observations, statistical genetics analysis, Bayesian modeling and computations, and others. In the 1990s, the method began to play an important role in computational biology and was used to solve problems in sequence motif identification and the analysis of complex pedigree. Now, the list of application areas of Monte Carlo methods includes biology (Leach 1996, Karplus and Petsko 1990, Lawrence, Altschul, Boguski, Liu, Neuwald and Wootton 1993), chemistry (Alder and Wainwright 1959), computer science (Kirkpatrick, Gelatt and Vecchi 1983), economics and finance (Gouriéroux and Monfort 1997); engineering (Geman and Geman 1984), material science (Frenkel and Smit 1996), physics (Metropolis, Rosenbluth, Rosenbluth, Teller and Teller 1953, Goodman and Sokal 1989, Marinari and Parisi 1992), statistics (Efron 1979, Gelfand and Smith 1990, Rubin 1987, Tanner and Wong 1987), and many others. Among all Monte Carlo methods, *Markov chain Monte Carlo* (MCMC) provides an enormous scope for dealing with very complicated stochastic systems and has been the central pillar in the study of macromolecules and other physical systems. Recently, the MCMC methodology has drawn much attention from statisticians because the method enables them to entertain more sophisticated and realistic statistical models.

Being attracted by the extreme flexibility and power of the Monte Carlo method, many researchers in different scientific areas have contributed to its development. However, because a substantial amount of domain-specific knowledge is required in order to understand problems in any of these fields, communications among researchers in these fields are very limited. Many efforts have been devoted to the reinvention of techniques that have been developed in other fields. It is therefore desirable to develop a relatively general framework in which scientists in every field — e.g., theoretical chemists, statistical physicists, structural biologists, statisticians, econometricians, and computer scientists — can compare their Monte Carlo techniques and learn from each other. For a large number of scientists and engineers who employ Monte Carlo simulation and related global optimization techniques (such as simulated annealing) as an essential tool in their work, there is also a need to keep up to date with recent advances in Monte Carlo method-

ologies and to understand the nature and connection of various proposed methods. The aim of this book is to provide a self-contained, unified, and up-to-date treatment of the Monte Carlo method.

This book is intended to serve three audiences: researchers specializing in the study of Monte Carlo algorithms; scientists who are interested in using advanced Monte Carlo techniques; and graduate students in statistics, computational biology, and computer sciences who want to learn about Monte Carlo computations. The prerequisites for understanding most of the methods described in this book are rather minimal: a one-semester course on probability theory (Pitman 1993) and a one-semester course on theoretical statistics (Rice 1994), both at the undergraduate level. However, it would be more desirable if the reader has some background in a specific scientific field such as artificial intelligence, computational biology, computer vision, engineering, or Bayesian statistics in which heavy computations are involved. This book is most suitable for a second-year graduate-level course on Monte Carlo methods, with an emphasis on their relevance to scientific and statistical research.

The author is most grateful to his mentor and friend Wing Hung Wong for his many important suggestions, his overwhelming passion for Monte Carlo and scientific problems, and his continuous encouragement. The author is also grateful to Persi Diaconis for teaching him many things including Markov chain theory, group theory, and nonparametric Bayes methods, to both Susan Holmes and Persi for numerous enlightening conversations on Markov chain Monte Carlo and other related problems, to Donald B. Rubin for insights on the missing data formulation and the Bayesian thinking, to Jonathan Goodman for helpful comments on multigrid Monte Carlo, to Yingnian Wu and Songchun Zhu for their materials on pattern simulations and thoughts on conditional sampling, to Faming Liang for his supply of many examples and figures, and to Minghui Chen and David van Dyk for helpful comments. Several former graduate students in the statistics departments of Stanford and Harvard universities — Yuguo Chen, Lingyu Chen, Chiara Sabatti, Tanya Logvinenko, Zhaohui Qin and Juni Zhang — have contributed in many ways to the development of this book. Ms. Helen Tombropoulos has provided editorial assistance to the author both for this book and for many articles published earlier. Finally, the author is greatly indebted to his wife Wei for her love and her continuous support of his research activities these years. Part of the book was written when the author was on the faculty of the Statistics Department of Stanford University. This work was also partially supported by the National Science Foundation Grants DMS-9803649 and DMS-0094613.

Contents

Preface	vii
1 Introduction and Examples	1
1.1 The Need of Monte Carlo Techniques	1
1.2 Scope and Outline of the Book	3
1.3 Computations in Statistical Physics	7
1.4 Molecular Structure Simulation	9
1.5 Bioinformatics: Finding Weak Repetitive Patterns	10
1.6 Nonlinear Dynamic System: Target Tracking	14
1.7 Hypothesis Testing for Astronomical Observations	16
1.8 Bayesian Inference of Multilevel Models	18
1.9 Monte Carlo and Missing Data Problems	19
2 Basic Principles: Rejection, Weighting, and Others	23
2.1 Generating Simple Random Variables	23
2.2 The Rejection Method	24
2.3 Variance Reduction Methods	26
2.4 Exact Methods for Chain-Structured Models	28
2.4.1 Dynamic programming	29
2.4.2 Exact simulation	30
2.5 Importance Sampling and Weighted Sample	31
2.5.1 An example	31
2.5.2 The basic idea	33
2.5.3 The “rule of thumb” for importance sampling	34

2.5.4	Concept of the weighted sample	36
2.5.5	Marginalization in importance sampling	37
2.5.6	Example: Solving a linear system	38
2.5.7	Example: A Bayesian missing data problem	40
2.6	Advanced Importance Sampling Techniques	42
2.6.1	Adaptive importance sampling	42
2.6.2	Rejection and weighting	43
2.6.3	Sequential importance sampling	46
2.6.4	Rejection control in sequential importance sampling	48
2.7	Application of SIS in Population Genetics	49
2.8	Problems	51
3	Theory of Sequential Monte Carlo	53
3.1	Early Developments: Growing a Polymer	55
3.1.1	A simple model of polymer: Self-avoid walk	55
3.1.2	Growing a polymer on the square lattice	56
3.1.3	Limitations of the growth method	59
3.2	Sequential Imputation for Statistical Missing Data Problems	60
3.2.1	Likelihood computation	60
3.2.2	Bayesian computation	62
3.3	Nonlinear Filtering	64
3.4	A General Framework	67
3.4.1	The choice of the sampling distribution	69
3.4.2	Normalizing constant	69
3.4.3	Pruning, enrichment, and resampling	71
3.4.4	More about resampling	72
3.4.5	Partial rejection control	75
3.4.6	Marginalization, look-ahead, and delayed estimate	76
3.5	Problems	77
4	Sequential Monte Carlo in Action	79
4.1	Some Biological Problems	79
4.1.1	Molecular Simulation	79
4.1.2	Inference in population genetics	81
4.1.3	Finding motif patterns in DNA sequences	84
4.2	Approximating Permanents	90
4.3	Counting 0-1 Tables with Fixed Margins	92
4.4	Bayesian Missing Data Problems	94
4.4.1	Murray's data	94
4.4.2	Nonparametric Bayes analysis of binomial data	95
4.5	Problems in Signal Processing	98
4.5.1	Target tracking in clutter and mixture Kalman filter	98
4.5.2	Digital signal extraction in fading channels	102
4.6	Problems	103

5	Metropolis Algorithm and Beyond	105
5.1	The Metropolis Algorithm	106
5.2	Mathematical Formulation and Hastings's Generalization	111
5.3	Why Does the Metropolis Algorithm Work?	112
5.4	Some Special Algorithms	114
5.4.1	Random-walk Metropolis	114
5.4.2	Metropolized independence sampler	115
5.4.3	Configurational bias Monte Carlo	116
5.5	Multipoint Metropolis Methods	117
5.5.1	Multiple independent proposals	118
5.5.2	Correlated multipoint proposals	120
5.6	Reversible Jumping Rule	122
5.7	Dynamic Weighting	124
5.8	Output Analysis and Algorithm Efficiency	125
5.9	Problems	127
6	The Gibbs Sampler	129
6.1	Gibbs Sampling Algorithms	129
6.2	Illustrative Examples	131
6.3	Some Special Samplers	133
6.3.1	Slice sampler	133
6.3.2	Metropolized Gibbs sampler	133
6.3.3	Hit-and-run algorithm	134
6.4	Data Augmentation Algorithm	135
6.4.1	Bayesian missing data problem	135
6.4.2	The original DA algorithm	136
6.4.3	Connection with the Gibbs sampler	137
6.4.4	An example: Hierarchical Bayes model	138
6.5	Finding Repetitive Motifs in Biological Sequences	139
6.5.1	A Gibbs sampler for detecting subtle motifs	140
6.5.2	Alignment and classification	141
6.6	Covariance Structures of the Gibbs Sampler	143
6.6.1	Data Augmentation	143
6.6.2	Autocovariances for the random-scan Gibbs sampler	144
6.6.3	More efficient use of Monte Carlo samples	146
6.7	Collapsing and Grouping in a Gibbs Sampler	146
6.8	Problems	151
7	Cluster Algorithms for the Ising Model	153
7.1	Ising and Potts Model Revisit	153
7.2	The Swendsen-Wang Algorithm as Data Augmentation	154
7.3	Convergence Analysis and Generalization	155
7.4	The Modification by Wolff	157
7.5	Further Generalization	157
7.6	Discussion	158

7.7	Problems	159
8	General Conditional Sampling	161
8.1	Partial Resampling	161
8.2	Case Studies for Partial Resampling	163
8.2.1	Gaussian random field model	163
8.2.2	Texture synthesis	165
8.2.3	Inference with multivariate t-distribution	169
8.3	Transformation Group and Generalized Gibbs	171
8.4	Application: Parameter Expansion for Data Augmentation	174
8.5	Some Examples in Bayesian Inference	176
8.5.1	Probit regression	176
8.5.2	Monte Carlo bridging for stochastic differential equation	178
8.6	Problems	181
9	Molecular Dynamics and Hybrid Monte Carlo	183
9.1	Basics of Newtonian Mechanics	184
9.2	Molecular Dynamics Simulation	185
9.3	Hybrid Monte Carlo	189
9.4	Algorithms Related to HMC	192
9.4.1	Langevin-Euler moves	192
9.4.2	Generalized hybrid Monte Carlo	193
9.4.3	Surrogate transition method	194
9.5	Multipoint Strategies for Hybrid Monte Carlo	195
9.5.1	Neal's window method	195
9.5.2	Multipoint method	197
9.6	Application of HMC in Statistics	198
9.6.1	Indirect observation model	199
9.6.2	Estimation in the stochastic volatility model	201
10	Multilevel Sampling and Optimization Methods	205
10.1	Umbrella Sampling	206
10.2	Simulated Annealing	209
10.3	Simulated Tempering	210
10.4	Parallel Tempering	212
10.5	Generalized Ensemble Simulation	215
10.5.1	Multicanonical sampling	216
10.5.2	The $1/k$ -ensemble method	217
10.5.3	Comparison of algorithms	218
10.6	Tempering with Dynamic Weighting	219
10.6.1	Ising model simulation at sub-critical temperature	221
10.6.2	Neural network training	222

11 Population-Based Monte Carlo Methods	225
11.1 Adaptive Direction Sampling: Snooker Algorithm	226
11.2 Conjugate Gradient Monte Carlo	227
11.3 Evolutionary Monte Carlo	228
11.3.1 Evolutionary movements in binary-coded space	230
11.3.2 Evolutionary movements in continuous space	231
11.4 Some Further Thoughts	233
11.5 Numerical Examples	235
11.5.1 Simulating from a bimodal distribution	235
11.5.2 Comparing algorithms for a multimodal example	236
11.5.3 Variable selection with binary-coded EMC	237
11.5.4 Bayesian neural network training	239
11.6 Problems	242
12 Markov Chains and Their Convergence	245
12.1 Basic Properties of a Markov Chain	245
12.1.1 Chapman-Kolmogorov equation	247
12.1.2 Convergence to stationarity	248
12.2 Coupling Method for Card Shuffling	250
12.2.1 Random-to-top shuffling	250
12.2.2 Riffle shuffling	251
12.3 Convergence Theorem for Finite-State Markov Chains	252
12.4 Coupling Method for General Markov Chain	254
12.5 Geometric Inequalities	256
12.5.1 Basic setup	257
12.5.2 Poincaré inequality	257
12.5.3 Example: Simple random walk on a graph	259
12.5.4 Cheeger's inequality	261
12.6 Functional Analysis for Markov Chains	263
12.6.1 Forward and backward operators	264
12.6.2 Convergence rate of Markov chains	266
12.6.3 Maximal correlation	267
12.7 Behavior of the Averages	269
13 Selected Theoretical Topics	271
13.1 MCMC Convergence and Convergence Diagnostics	271
13.2 Iterative Conditional Sampling	273
13.2.1 Data augmentation	273
13.2.2 Random-scan Gibbs sampler	275
13.3 Comparison of Metropolis-Type Algorithms	277
13.3.1 Peskun's ordering	277
13.3.2 Comparing schemes using Peskun's ordering	279
13.4 Eigenvalue Analysis for the Independence Sampler	281
13.5 Perfect Simulation	284
13.6 A Theory for Dynamic Weighting	287
13.6.1 Definitions	287

13.6.2	Weight behavior under different scenarios	288
13.6.3	Estimation with weighted samples	291
13.6.4	A simulation study	292
A	Basics in Probability and Statistics	295
A.1	Basic Probability Theory	295
A.1.1	Experiments, events, and probability	295
A.1.2	Univariate random variables and their properties	296
A.1.3	Multivariate random variable	298
A.1.4	Convergence of random variables	300
A.2	Statistical Modeling and Inference	301
A.2.1	Parametric statistical modeling	301
A.2.2	Frequentist approach to statistical inference	302
A.2.3	Bayesian methodology	304
A.3	Bayes Procedure and Missing Data Formalism	306
A.3.1	The joint and posterior distributions	306
A.3.2	The missing data problem	308
A.4	The Expectation-Maximization Algorithm	310
	References	313
	Author Index	333
	Subject Index	338

1

Introduction and Examples

1.1 The Need of Monte Carlo Techniques

An essential part of many scientific problems is the computation of integral

$$I = \int_D g(\mathbf{x}) d\mathbf{x},$$

where D is often a region in a high-dimensional space and $g(\mathbf{x})$ is the target function of interest. If we can draw independent and identically distributed (i.i.d.) random samples $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)}$ uniformly from D (by a computer), an approximation to I can be obtained as

$$\hat{I}_m = \frac{1}{m} \{g(\mathbf{x}^{(1)}) + \dots + g(\mathbf{x}^{(m)})\}.$$

The *law of large numbers* states that the average of many independent random variables with common mean and finite variances tends to stabilize at their common mean (see the Appendix); that is,

$$\lim_{m \rightarrow \infty} \hat{I}_m = I, \text{ with probability 1.}$$

Its convergence rate can be assessed by the *central limit theorem* (CLT):

$$\sqrt{m}(\hat{I}_m - I) \rightarrow N(0, \sigma^2), \text{ in distribution,}$$

where $\sigma^2 = \text{var}\{g(\mathbf{x})\}$. Hence, the “error term” of this Monte Carlo approximation is $O(m^{-1/2})$, regardless of the dimensionality of \mathbf{x} . This basic

setting underlies the potential role of the Monte Carlo methodology in science and statistics.

In the simplest case when $D = [0, 1]$ and $I = \int_0^1 g(x)dx$, one can approximate I by

$$\bar{I}_m = \frac{1}{m} \{g(b_1) + \cdots + g(b_m)\},$$

where $b_j = j/m$. This method can be called the *Riemann approximation*. When g is reasonably smooth, the Riemann approximation gives us an error rate of $O(m^{-1})$, better than that of the Monte Carlo method. More sophisticated methods such as Simpson's rule and the Newton-Cotes rules give better numerical approximations (Thisted 1988). However, a fatal defect of these deterministic methods is that they do not scale well as the dimensionality of D increases. For example, in a 10-dimensional space with $D = [0, 1]^{10}$, we will have to evaluate $O(m^{10})$ grid points in order to achieve an accuracy of $O(m^{-1})$ in the Riemann approximation of I . In contrast, the naive Monte Carlo approach, which draws $x^{(1)}, \dots, x^{(m)}$ uniformly from D , has an error rate $O(m^{-1/2})$ regardless of the dimensionality of D , at least theoretically.

Although the "error rate" of a Monte Carlo integration scheme remains the same in high-dimensional problems, two intrinsic difficulties arise: (a) when the region D is large in high-dimensional space, the variance σ^2 , which measures how "uniform" the function g is in region D , can be formidably large; (b) one may not be able to produce uniform random samples in an arbitrary region D . To overcome these difficulties, researchers often employ the idea of *importance sampling* in which one generates random samples $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)}$ from a nonuniform distribution $\pi(\mathbf{x})$ that puts more probability mass on "important" parts of the state space D . One can estimate integral I as

$$\hat{I} = \frac{1}{m} \sum_{j=1}^m \frac{g(\mathbf{x}^{(j)})}{\pi(\mathbf{x}^{(j)})},$$

which has a variance $\sigma_\pi^2 = \text{var}_\pi\{g(\mathbf{x})/\pi(\mathbf{x})\}$. In the most fortunate case, we may choose $\pi(\mathbf{x}) \propto g(\mathbf{x})$ when g is non-negative and I is finite, which results in an exact estimate of I . But in no known application of the Monte Carlo method has this "luckiest situation" ever occurred. More realistically, we may hope to find a good "candidate" π which will explore more in regions where the value of g is high. In such a situation, generating random draws from π can be a challenging problem.

Demands for sampling from a nonuniform distribution π are also seen from another set of problems in bioinformatics, computational chemistry, physics, structural biology, statistics, etc. In these problems, the desired probability distribution $\pi(\mathbf{x})$ of a complex system, where \mathbf{x} is often called a *configuration* of the system, arises from basic laws in physics and statistical