

Jun S. Liu

# Monte Carlo Strategies in Scientific Computing

With 56 Figures



Springer



# Contents

|                                                                |            |
|----------------------------------------------------------------|------------|
| <b>Preface</b>                                                 | <b>vii</b> |
| <b>1 Introduction and Examples</b>                             | <b>1</b>   |
| 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         |
| <b>2 Basic Principles: Rejection, Weighting, and Others</b>    | <b>23</b>  |
| 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        |
| <b>3</b> | <b>Theory of Sequential Monte Carlo</b>                               | <b>58</b> |
| 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        |
| <b>4</b> | <b>Sequential Monte Carlo in Action</b>                               | <b>79</b> |
| 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       |



|          |                                                                  |            |
|----------|------------------------------------------------------------------|------------|
| <b>5</b> | <b>Metropolis Algorithm and Beyond</b>                           | <b>105</b> |
| 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        |
| <b>6</b> | <b>The Gibbs Sampler</b>                                         | <b>129</b> |
| 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        |
| <b>7</b> | <b>Cluster Algorithms for the Ising Model</b>                    | <b>153</b> |
| 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        |
| <b>8</b>  | <b>General Conditional Sampling</b>                                      | <b>161</b> |
| 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 equa-<br>tion . . . . . | 178        |
| 8.6       | Problems . . . . .                                                       | 181        |
| <b>9</b>  | <b>Molecular Dynamics and Hybrid Monte Carlo</b>                         | <b>183</b> |
| 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        |
| <b>10</b> | <b>Multilevel Sampling and Optimization Methods</b>                      | <b>205</b> |
| 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        |



|                                                                 |            |
|-----------------------------------------------------------------|------------|
| <b>11 Population-Based Monte Carlo Methods</b>                  | <b>225</b> |
| 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        |
| <b>12 Markov Chains and Their Convergence</b>                   | <b>245</b> |
| 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        |
| <b>13 Selected Theoretical Topics</b>                           | <b>271</b> |
| 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        |
| <b>A</b> | <b>Basics in Probability and Statistics</b>                | <b>295</b> |
| 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        |
|          | <b>References</b>                                          | <b>313</b> |
|          | <b>Author Index</b>                                        | <b>333</b> |
|          | <b>Subject Index</b>                                       | <b>338</b> |