Example Syllabus – ISC 5228
Markov Chain Monte Carlo Simulations
Fall Semester 2006

Instructor: Prof. Bernd A. Berg, 443 DSL, phone 644-1959 or -1010

Meeting times : TBA, 152 DSL.

Credit: 3 hours.

Prerequisites: Calculus I and II, Scientific Programing, or consent of the instructor.

Course Description:

With the emergence of high performance computing simple Monte Carlo (MC) and Markov Chain Monte Carlo (MCMC) simulations have become a major enabling technique in applied sciences. The goal of this course is to provide a self-contained presentation. This includes not only Monte Carlo methods themselves, but also its statistical foundations and the statistics needed to analyze MCMC data, which are autocorrelated. The course is addressed to applied scientists, who need "ready to work recipes" and running computer code (which will be provided for all subjects covered). Illustrative examples may be drawn from statistical physics or from topics of other participating departments, for instance from simulations of biomolecules. Students are encouraged to bring in their own examples, related to their planned or actual research work.

Text (required):

Markov Chain Monte Carlo Simulations and Their Statistical Analysis (With Web-Based Fortran Code), Bernd A. Berg, World Scientific 2004, ISBN 981-238-935-0.

Alternate Texts (not required):

Markov Chain Monte Carlo: Innovations and Applications (Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, Vol. 7), W.S. Kendall, F. Liang, and J.-S. Wang, World Scientific 2005, ISBN 981-256-427-6.

Markov Chain Monte Carlo in Practice – Interdisciplinary Statistics, W.R. Gilks, S. Richardson and D.J. Spiegelhalter, Chapman and Hall/CRC, 1996, ISBN 041-205-551-1.

A Guide to Monte Carlo Simulations in Statistical Physics, David Landau and Kurt Binder, Cambridge University Press, 2000, ISBN 052-165-366-5.

Monte Carlo Strategies in Scientific Computing (Springer Series in Statistics), Jun S. Liu, Springer, 2001, ISBN 038-795-230-6.

Objectives:

The students should become able to implement the methods in their own computer programs and to apply them to their own research projects. Towards the end of the semester each student will have to demonstrate her or his understanding by presenting a final project to the class. On the way to the goal that the students use MCMC simulations in own research, there will be numerous assignments, which define and test the required skills. In particular the assignments will include (but are not be limited to):

- Install and run pseudo-random number generators.
- Make simple graphical illustration (e.g., gnuplot) for probability densities and cumulative distribution functions.
- Apply proper difference tests (Gaussian, Student, F-Test) to estimators of suitable random variables.
- Estimate confidence bounds from simple MC simulations, including applications of the statistical bootstrap.
- Perform a jackknife analysis of a sample of independent random variables.

- Set up MCMC simulations using the Metropolis, the heatbath or other suitable updating schemes.
- Illustrate "self-averaging" of MCMC data with a least on example from statistical physics.
- Perform binning (blocking) of MCMC data and a subsequent jackknife calculation of their error bars.
- Re-weight MCMC data.
- Calculate autocorrelations and the integrated autocorrelation time from MCMC data.
- Evaluate the efficiency of distinct MCMC updating schemes against one another.
- Explain the difference between a self-consistent and a "reasonable" error analysis of MCMC data.
- Determine fit parameters from independent data of MCMC simulations.
- Determine fit parameters from autocorrelated data of MCMC simulations.
- Be able to give an overview of applications of MCMC simulations.
- Setup at least one advanced simulation, e.g., simulated tempering on several CPUs using the Message Passing Interface (MPI) language.

Outline:

The corresponding chapters of the book will be reading assignments as posted on the web.

1. Sampling, statistics and computer code (4 lectures)
   - probability distributions, sampling and random numbers
   - organization of the computer code
   - Gaussian distribution
   - confidence intervals
   - order statistics and heapsort
   - functions and expectation values of random variables
   - sums of independent random variables and characteristic functions
   - linear transformation of random variables
   - the central limit theorem and binning
2. Error analysis for independent random variables (5 lectures)
   - Gaussian confidence intervals and error bars
   - estimators of the variance and bias
   - statistical error bar routines
   - the Gaussian difference test
   - the χ^2 and sample variance distribution
   - Gosset's student distribution and student difference test
   - the error of the error bar and the variance ratio test (F-test)
   - When are distributions consistent? (χ^2 and Kolmogorov tests)
   - The jackknife approach
   - determination of parameters (linear and Levenberg-Marquardt fitting)
3. Markov Chain Monte Carlo (6 lectures)
   - preliminaries: The canonical ensemble and the two-dimensional Ising model
   - lattice labeling
   - sampling and re-weighting
   - importance sampling
   - the Metropolis algorithm
   - the heath bath algorithm
   - illustrations of Monte Carlo simulations I: discrete systems
   - illustrations of Monte Carlo simulations II: continuous systems
4. Error analysis for Markov chain data (4 lectures)
   - autocorrelations
   - integrated autocorrelation time and binning
   - self-consistent versus reasonable error analysis
   - analysis of MCMC data (examples)

- comparison of MCMC algorithms
- fitting of MCMC data

5. Advanced Topics (7 lectures, the instructor will choose a subset)
   - simulations of generalized ensembles
   - free energy and entropy calculations
   - event driven simulations
   - cluster algorithms
   - molecular dynamics
   - hybrid Monte Carlo
   - kinetic Monte Carlo
   - large scale simulations
   - parallel computing
   - **to be announced (may be influenced by students)**
6. Presentation of Final Projects (5 lectures, starting whenever the first project is ready)

The estimates of the numbers of lectures are approximate and include time for tests and discussions.

Grading/Evaluation:

The course grade will be based to 70% on assignments (mainly computational) and to 30% on a final project, which each student is expected to present in class. Regular class attendance is required. At least half of the assignments will be given unannounced during class hours and a score of zero points will be recorded if you are absent without a legitimate excuse. The other assignments will be given as homework. The topic of the final project should be discussed with the instructor and be decided by the eighth week or earlier. This includes mandatory visits of the office hours of the instructor. Projects which relate MCMC calculations to your actual research are most desireable.

The course grade dividing lines are: $A > 90\%$, $A^- > 85\%$, $B^+ > 80\%$, $B > 70\%$, $B^- > 65\%$, $C^+ > 60\%$, $C > 50\%$, $C^- > 45\%$, $D > 30\%$ and F for less or equal 30%.

Honor Code: A copy of the University Academic Honor Code can be found in the current Student Handbook. You are bound by this in all of your academic work. It is based on the premise that each student has the responsibility 1) to uphold the highest standards of academic integrity in the students own work, 2) to refuse to tolerate violations of academic integrity in the University community, and 3) to foster a high sense of integrity and social responsibility on the part of the University community. Out of class you are encouraged to work together on assignments but plagiarizing of the work of others or study manuals is academically dishonest. Please see the following web site for a complete explanation of the Academic Honor Code.

http://www.fsu.edu/Books/Student-Handbook/codes/honor.html
http://www.fsu.edu/Books/Student-Handbook/

ADA: Students with disabilities needing academic accommodations should: 1) register with and provide documentation to the Student Disability Resource Center (SDRC); 2) bring a letter to the instructor from SDRC indicating you need academic accommodations. This should be done within the first of class. This and other class materials are available in alternative format upon request.