BIASED METROPOLIS-HEATBATH ALGORITHM

Alexei Bazavov and Bernd A. Berg

Department of Physics,
Florida State University

May 1, 2007

Introduction

Notation

- Probability density function (PDF) $P(x)$, $Q(x)$
- Assume for clarity $-1 \leqslant x \leqslant 1$
- Cumulative distribution function (CDF)

$$F(x) = \frac{\int_{-1}^{x} dx' \, P(x')}{\int_{-1}^{1} dx' \, P(x')}$$

- r is a uniformly distributed random number $r \in [0, 1)$
- In a Markov process subindex o corresponds to current state, n – to a new state
- U, V, T – SU(N) matrices in fundamental representation

Goal: efficiently sample $P(x)$

Metropolis Algorithm (MA)

- Sampling procedure:
  1. Generate x_n uniformly in the domain
  2. Accept with probability

$$\gamma_{o\to n} = \min\left\{1, \frac{P(x_n)}{P(x_o)}\right\}$$

  3. Otherwise $x_n = x_o$
- *Acceptance rate* (AR) is the number of accepted over the number of proposed changes
- Low AR leads to a high degree of autocorrelation

Ideal Heatbath (IHB)

- The "filter"
$$x = F^{-1}(r)$$
converts a uniformly distributed random number r into x distributed with $P(x)$
- In many cases numerical evaluation of F^{-1} is too slow to be an option

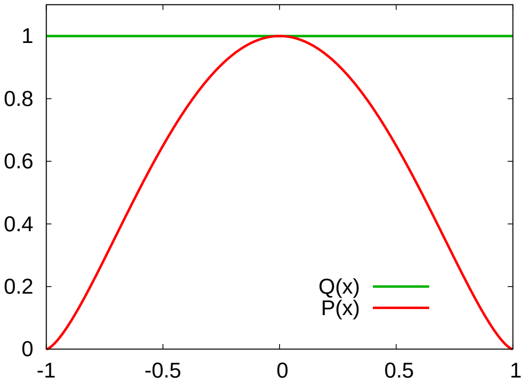
Heatbath (HB)

- "Envelope" the desired distribution $P(x)$ with some other PDF $Q(x)$ which can be sampled efficiently (e.g. with an ideal heatbath)
- Von Neumann rejection method:
  1. Sample x_n from $Q(x)$
  2. Accept with probability

$$\gamma_{o \to n} = \frac{P(x_n)/Q(x_n)}{(P(x)/Q(x))_{max}}$$

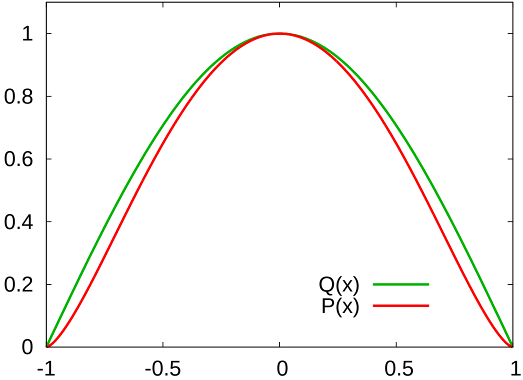
  3. Repeat until accepted (RUA)
- *Trial rate* (TR) is the number of trials needed for x_n to be accepted

Example: $P(x) = (1 - x^2)^{3/2}$



- Enveloping PDF $Q(x) = 1$
- TR is equal to the ratio of areas: $16/3\pi \simeq 1.70$

Example: $P(x) = (1 - x^2)^{3/2}$



- Enveloping PDF $Q(x) = \cos(\pi x/2)$
- TR is equal to the ratio of areas: $32/3\pi^2 \simeq 1.08$

General Approach

The generalized transition probabilities (Hastings, 1970):

$$W_{o\to n} = Q_{o\to n}\,\gamma_{o\to n},$$

- $Q_{o\to n}$ – proposal probability
- $\gamma_{o\to n}$ – acceptance probability

$$\gamma_{o\to n} = \frac{f[\min\{(P_o\,Q_{o\to n})/(P_n\,Q_{n\to o}),(P_n\,Q_{n\to o})/(P_o\,Q_{o\to n})\}]}{1+(P_o\,Q_{o\to n})/(P_n\,Q_{n\to o})}$$

with $0 \leqslant f(x) \leqslant 1 + x$ for $0 \leqslant x \leqslant 1$
A simple possibility:

$$f(x) = 1 + x$$

General Approach

Finally

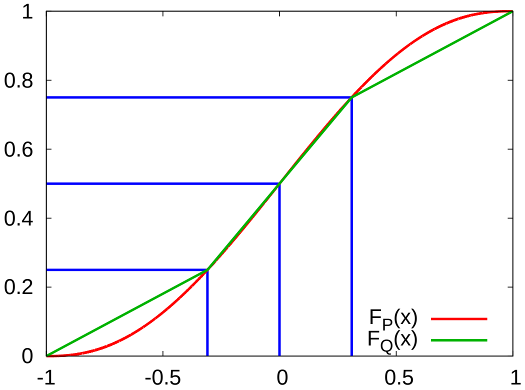
$$\gamma_{o \to n} = \min \left\{ 1, \frac{P_n}{P_o} \frac{Q_{n \to o}}{Q_{o \to n}} \right\}$$

For the previous example

$$\gamma_{o \to n} = \min \left\{ 1, \frac{P(x_n)}{P(x_o)} \frac{Q(x_o)}{Q(x_n)} \right\}$$

The choice of $Q(x)$ influences only AR, with any $Q(x)$ a Markov process with such transition probabilities results in sampling $P(x)$ distribution

Example: $P(x) = (1-x^2)^{3/2}$



- $F_Q(x)$ is a piece-wise linear approximation of $F_P(x)$ ($m = 4$ steps)
- $Q(x \in [x_{i-1}, x_i)) = 1/(m\,\Delta x_i)$, $i = 1, \ldots, m$

BMHA

- Construct BMHA table: $x_0 < x_1 < \ldots < x_m$
- Sampling procedure:
  1. Find i_o such that $x_o \in [x_{i_o-1}, x_{i_o})$
  2. Generate integer i_n uniformly from 1 to m
  3. Generate x_n uniformly in the interval $[x_{i_n-1}, x_{i_n})$
  4. Accept with probability

$$\gamma_{o \to n} = \min\left\{1, \frac{P(x_n)}{P(x_o)} \frac{\Delta x_{i_n}}{\Delta x_{i_o}}\right\}$$

  5. Otherwise $x_n = x_o$
- Proposal probability of the heatbath is combined with Metropolis-type acceptance probability

SU(2) Gauge Theory

$$P(U) = \exp\{\beta S(U)\}, \quad S(U) = \frac{1}{2}\sum_{i=1}^{6} \mathrm{Re}\mathrm{Tr}[T_i\, U]$$

$$T = \sum_{i=1}^{6} T_i, \quad \text{let} \quad \alpha^2 = \det||T||, \quad 0 \leqslant \alpha \leqslant 6$$

$$\text{then} \quad \tilde{T} = \frac{1}{\alpha}\, T \in \mathrm{SU}(2), S(U) = \frac{1}{2}\,\alpha\, \mathrm{Re}\mathrm{Tr}[\tilde{T}\, U]$$

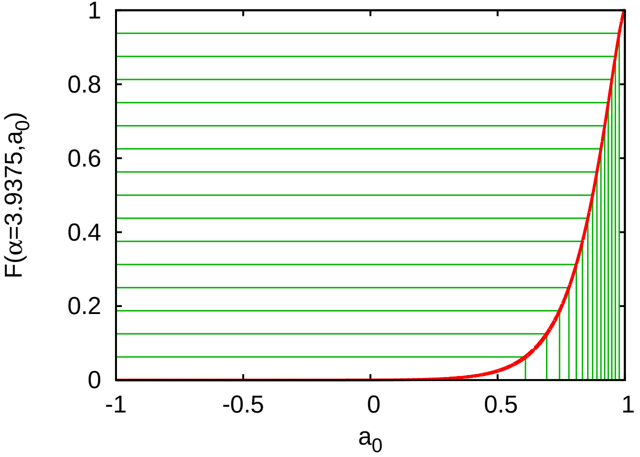
$$P(\alpha, U)dU = \exp\left\{\frac{\beta\,\alpha}{2}\mathrm{Re}\mathrm{Tr}[\tilde{T}\, U]\right\} dU$$

$$V = \tilde{T}\, U, \quad dV = dU, \quad V = a_0 I + i\vec{a}\,\vec{\sigma}$$

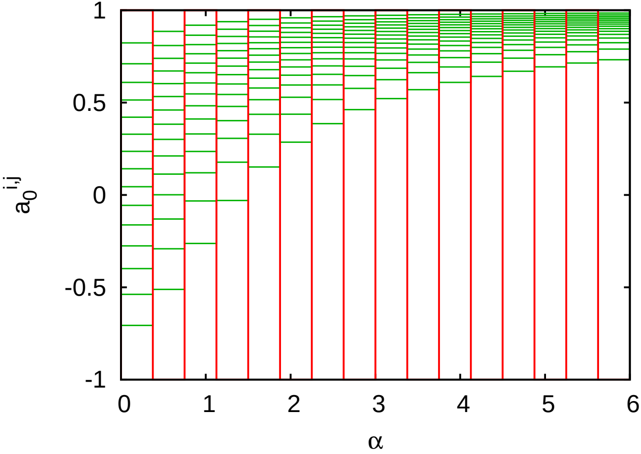
$$dV = \sqrt{1 - a_0^2}\, da_0\, d\Omega_{\vec{a}}, \quad \mathrm{Re}\mathrm{Tr}[V] = 2a_0$$

$$P(\alpha, a_0) = \sqrt{1 - a_0^2}\, \exp\{\beta\,\alpha\, a_0\}$$

SU(2) Gauge Theory



SU(2) Gauge Theory

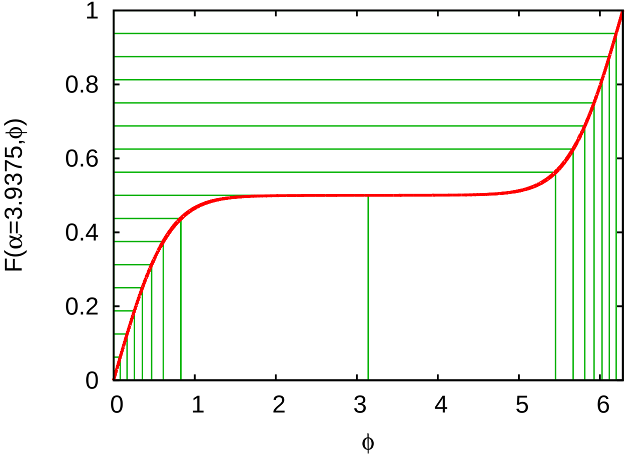


SU(2) Gauge Theory

Lattice: 4×16^3
Coupling: $\beta_g = 2.3$
Sweeps: $16384 + 32 \times 20480$
CDF discretization: 32×128

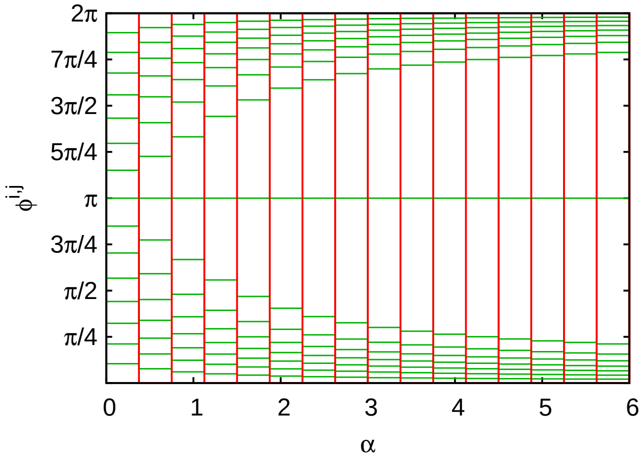
	HB (FHKP)	Metropolis	BMHA
CPU time	194,873 [s]	181,321 [s]	199,244 [s]
TR/AR	1.043	0.111	0.975
$\langle \mathrm{Tr}(U_\Box)/2 \rangle$	0.603147 (17)	0.603066 (52)	0.603111 (21)
τ_{int}	49.8 (3.5)	409 (66)	48.2 (3.8)

U(1) Gauge Theory



$$P(\alpha,\phi) = \exp\{\beta\,\alpha\ \cos\phi\}$$

U(1) Gauge Theory



U(1) Gauge Theory

Lattice: 4×16^3
Coupling: $\beta_g = 1.0$
Sweeps: $16384 + 32 \times 20480$
CDF discretization: 32×128

	Metropolis	BMHA
CPU time	84,951 [s]	107,985 [s]
AR	0.286	0.972
$\langle \cos \phi_{\Box} \rangle$	0.59103 (16)	0.59106 (12)
τ_{int}	341 (26)	142 (10)

SU(2) F-A Gauge Theory

$$P(U) = \exp\{\beta_f S_f(U) + \beta_a S_a(U)\}, \quad S_a(U) = \frac{1}{3}\sum_{i=1}^{6}(\mathrm{ReTr}[T_i\, U])^2$$

How to proceed:

- For proposal probabilities use BMHA table which takes into account only the fundamental part (BMHA-fund)
- Approximate the adjoint part by neglecting fluctuations in T_i:

$$T = \sum_{i=1}^{6} T_i, \quad T_i \to T_i' = \frac{1}{6}\, T$$

Proposal PDF

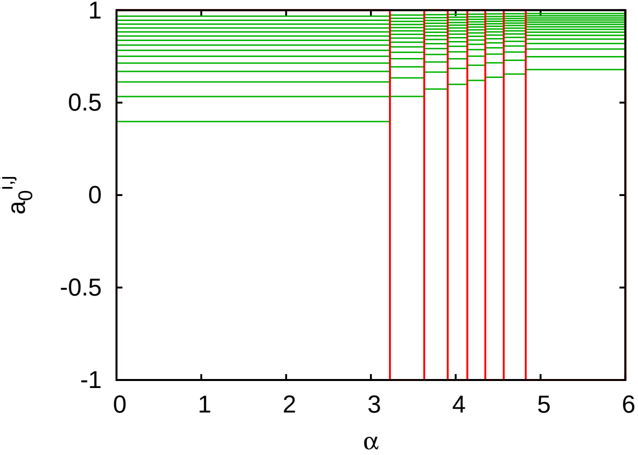
$$Q(\alpha, a_0) = \sqrt{1 - a_0^2}\, \exp\left\{\beta_f\, \alpha\, a_0 + \frac{2\beta_a}{9}\, \alpha^2\, a_0^2\right\}$$

SU(2) F-A Gauge Theory

Lattice: 4^4
Coupling: $(\beta_f, \beta_a) = (1.5, 0.9)$
Sweeps: $1000 + 32 \times 1000$
CDF discretization: 32×128

	Metropolis	BMHA-fund	BMHA
AR	0.065	0.65	0.82
$\langle U^f_\square \rangle$	0.3451(15)	0.34636(52)	0.34694(62)
$\langle U^a_\square \rangle$	0.6368(15)	0.63798(47)	0.63853(56)
$\tau_{\text{int}}(\langle U^f_\square \rangle)$	100.2 (8.6)	19.5 (1.7)	19.8 (2.5)
$\tau_{\text{int}}(\langle U^a_\square \rangle)$	95.9 (8.0)	17.1 (1.4)	16.5 (2.2)

SU(2) Gauge Theory



CDF and BMHA tables can be built out of data: take into account how parameters are distributed (apriori unknown)

Summary

BMHA

- comparable in performance with existing HB (SU(2)) or better (U(1) and SU(2) F-A)
- easy to construct
- generalizable for multivariate PDFs
- has uniform speed
- can be constructed from data (!)
- easily combined with multicanonical algorithm