# IMA Summer Program

# Classical and Quantum Approaches in Molecular Modeling

## Lecture 5: Molecular Sampling

Robert D. Skeel
Department of Computer Science, and of Mathematics
Purdue University

http://bionum.cs.purdue.edu/2007July24.pdf

Acknowledgment: NIH

# Sampling

Create a stochastic (or deterministic!) ergodic Markov chain to generate $x^1, x^2, \ldots, x^n, \ldots$ having the desired distribution.

- Monte Carlo methods are unbiased in the limit $N_{\text{trials}} \to \infty$ but waste information.

- Molecular dynamics with extended Hamiltonians and/or stochastic terms has a bias due to finite $\Delta t$.

Sampling can be *accelerated* by techniques such as replica exchange and multicanonical/Wang-Landau sampling.

# Canonical ensemble

Let $\beta = 1/k_{\mathrm{B}}T$.
The probability density function factors as

$$\text{const}\, \mathrm{e}^{-\beta p^{\mathsf{T}} M^{-1} p/2} \cdot \rho(x) \quad \text{where } \rho(x) = \text{const}\, \mathrm{e}^{-\beta U(x)}.$$

The constant in $\rho(x)$ is unknown.

# Outline

I. Markov chain Monte Carlo methods

II. Hybrid Monte Carlo methods

III. Molecular dynamics

IV. Replica exchange method

V. Multicanonical/Wang-Landau sampling

VI. Practicalities

# A simple Markov chain Monte Carlo method

Given $x$, one step is as follows:

Pick an atom $i$ at random.
Generate i.i.d. values $\Delta_{\mathrm{x}}$, $\Delta_{\mathrm{y}}$, $\Delta_{\mathrm{z}}$ uniform on $[-\frac{1}{2}\Delta, \frac{1}{2}\Delta]$
$x' = x$ with $[\Delta_{\mathrm{x}}, \Delta_{\mathrm{y}}, \Delta_{\mathrm{z}}]^{\mathsf{T}}$ added to $\vec{r}_i$
Accept $x'$ with probability

$$\min\{1, \rho(x')/\rho(x)\}$$

—*Metropolis acceptance criterion.*

Otherwise, choose $x$ as the new value.

*Note.* A rejected move counts as a step in the chain.

These trial moves have *symmetric* conditional p.d.f.:

$$\rho_{\mathrm{t}}(x|x') = \rho_{\mathrm{t}}(x'|x).$$

# MCMC Convergence

Each configuration $x^n$ of the Markov chain has a density $\rho^n(x)$.
Only in the limit $n \to \infty$ might $\rho^n(x) \to \rho(x)$: *convergence.*

**Proposition.** convergence $\Leftrightarrow$ stationarity & ergodicity.

*Stationarity* means $\rho^n(x) = \rho(x) \Rightarrow \rho^{n+1}(x) = \rho(x)$

*Ergodicity* means that all subsets of positive measure will be visited with probability $> 0$ in finite time.

The Markov chain given previously is ergodic.

# Stationarity

**Proposition.** detailed balance $\Rightarrow$ stationarity.

*Detailed balance*$^*$ means $\rho(x|x')\rho(x') = \rho(x'|x)\rho(x)$
where $\rho(x'|x)$ is the conditional p.d.f. for a complete step.

**Proposition.** symmetric trial moves & Metropolis criterion $\Rightarrow$ detailed balance

# Metropolis-Hastings criterion

accommodates nonsymmetric trial moves by accepting with probability

$$\min\left\{1, \frac{\rho(x')\rho_{\rm t}(x|x')}{\rho(x)\rho_{\rm t}(x'|x)}\right\}.$$

Example is a trial move given by

$$x' = x + \Delta t D \nabla \log \rho(x) + \sqrt{2\Delta t} D^{1/2} Z$$

where $D$ is a constant diagonal matrix and $Z$ is a set of independent standard Gaussian random numbers.

I would call this scheme a . . .

# "Brownian dynamics sampler"

Basic idea is similar to

force-bias MC (1978)

and almost identical to

smart MC (1978).

Abstracted idea has been called

Metropolis-adjusted Langevin algorithm (1994).

# Outline

I. Markov chain Monte Carlo methods

II. Hybrid Monte Carlo methods

III. Molecular dynamics

IV. Replica exchange method

V. Multicanonical/Wang-Landau sampling

VI. Practicalities

# Hybrid Monte Carlo

Hybrid Monte Carlo uses MD to generate possible moves.
Given $x$ :

(1) Generate $p$ from const $\exp(-\frac{1}{2}\beta p^{\mathsf{T}} M^{-1} p)$.

(2) Obtain $x'$, $p'$ from short MD trajectory.

(3) Accept $x'$ with probability

$$\min\left\{1, \frac{\exp(-\beta H(x', p'))}{\exp(-\beta H(x, p))}\right\},$$

where $H(x, p) = \frac{1}{2} p^{\mathsf{T}} M^{-1} p + U(x)$.

(4) If rejected, choose $x$.

# Convergence of HMC

It is enough that the integrator

be reversible and volume-preserving.

Counterintuitively, $\langle \Delta H \rangle > 0$ , $\Delta H = H(x', p') - H(x, p)$.
Indeed, $\langle \Delta H \rangle \propto \Delta t^{2p} N$.

To get a 50% acceptance rate, we need

$\langle \Delta H \rangle = 0.9099\, k_{\mathrm{B}} T$,

which implies

$\Delta t \propto N^{-1/2p}$

# Outline

I. Markov chain Monte Carlo methods

II. Hybrid Monte Carlo methods

III. Molecular dynamics

IV. Replica exchange method

V. Multicanonical/Wang-Landau sampling

VI. Practicalities

# Deterministic MD

Instantaneous temperature $T(p) = p^{\mathsf{T}} M^{-1} p / k_{\mathrm{B}} N_{\mathrm{d}}$
where $N_{\mathrm{d}} =$ number of DOFs.
Nosé-Hoover augments Newton's equations with a thermostat:

$$\frac{\mathrm{d}}{\mathrm{d}t} x = M^{-1} p, \qquad \frac{\mathrm{d}}{\mathrm{d}t} p = -\nabla U(x) - \frac{p_{\mathrm{s}}}{Q} p,$$

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{\mathrm{s}} = N_{\mathrm{d}} k_{\mathrm{B}} (T(p) - T).$$

where $Q =$ thermal inertia.
If $T(p) > T$, the value of $p_{\mathrm{s}}$ will increase,
and eventually $p_{\mathrm{s}}$ will be positive,
causing $p$—and $T(p)$—to decrease.
(And the opposite happens if $T(p) < T$.)

Nosé-Hoover generates canonical ensemble via

$$\langle A(x,p)\rangle = \lim_{t\to\infty} \frac{1}{t}\int_0^t A(x(t),p(t))\,\mathrm{d}t.$$

It is not Hamiltonian but has a conserved quantity

$$\frac{1}{2}p^{\mathsf{T}}M^{-1}p + U(x) + \frac{1}{2Q}p_{\mathrm{s}}^2 + N_{\mathrm{d}}k_{\mathrm{B}}T\ln s$$

where

$$\frac{\mathrm{d}}{\mathrm{d}t}s = \frac{1}{Q}p_{\mathrm{s}}s.$$

Drift in this "extended energy" may be excessive, however.

# Nosé-Poincaré

It is defined by the extended Hamiltonian $H(x, s, \bar{p}, p_{\mathrm{s}}) =$

$$s\left(\frac{1}{2}s^{-2}\bar{p}^{\mathsf{T}}M^{-1}\bar{p} + U(x) + \frac{1}{2Q}p_{\mathrm{s}}^2 + N_{\mathrm{d}}k_{\mathrm{B}}T\ln s - E\right)$$

where $E$ is chosen to make $H$ initially zero.

$$\langle A\rangle = \lim_{t\to\infty}\frac{1}{t}\int_0^t A(x(t'), s(t')^{-1}\bar{p}(t'))\,\mathrm{d}t'.$$

# Stochastic MD

Recall Langevin dynamics

$$M\frac{\mathrm{d}}{\mathrm{d}t^2}x = F(x) - CM\frac{\mathrm{d}}{\mathrm{d}t}x + (2k_{\mathrm{B}}TCM)^{1/2}\frac{\mathrm{d}}{\mathrm{d}t}W(t)$$

where

$C$ is a diagonal matrix of damping constants, e.g., $5\,\mathrm{ps}^{-1}$, and

$W(t)$ is a set of $3N$ independent canonical Wiener processes.

# Outline

I. Markov chain Monte Carlo methods

II. Hybrid Monte Carlo methods

III. Molecular dynamics

IV. Replica exchange method

V. Multicanonical/Wang-Landau sampling

VI. Practicalities

# Temperature replica exchange

Large conformational barriers can be surmounted
by raising temperature,
and this can be done
while maintaining Boltzmann-Gibbs sampling
using replica-exchange aka parallel tempering (1995).

# Method

Simulate an ensemble of systems at temperatures
$T_1 < T_2 < \cdots < T_\mu$ where $T_1$ is the desired temperature.
Periodically, choose $\nu$ at random from $1, 2, \ldots, \mu - 1$,
and consider swapping configurations $x_{(\nu)}$ and $x_{(\nu+1)}$.
Probability of the swapped state relative the unswapped state is

$$r = \frac{\rho_\nu\left(x_{(\nu+1)}\right)}{\rho_\nu\left(x_{(\nu)}\right)} \frac{\rho_{\nu+1}\left(x_{(\nu)}\right)}{\rho_{\nu+1}\left(x_{(\nu+1)}\right)}$$

where $\rho_\nu(x)$ is the probability density for temperature $T_\nu$.
Accept the exchange with a probability

$$\min\{1, r\}.$$

Then continue sampling.

# Shortcoming

Probability of rejection increases with system size $N$.

Number of replicas needed to prevent excessive rejections increases as $N^{1/2}$.

# Hamiltonian replica exchange

A generalization of replica exchange.

As an example, write $U = U^{\mathrm{pp}} + U^{\mathrm{pw}} + U^{\mathrm{ww}}$
where the splitting represents protein–protein, protein–water, and water–water interactions, respectively.

Then consider
$U_\nu = \gamma_\nu U^{\mathrm{pp}} + U^{\mathrm{pw}} + U^{\mathrm{ww}}, \quad 1 = \gamma_1 > \gamma_2 > \cdots > \gamma_\mu$
—hot solute, cold solvent.

# Outline

I. Markov chain Monte Carlo methods

II. Hybrid Monte Carlo methods

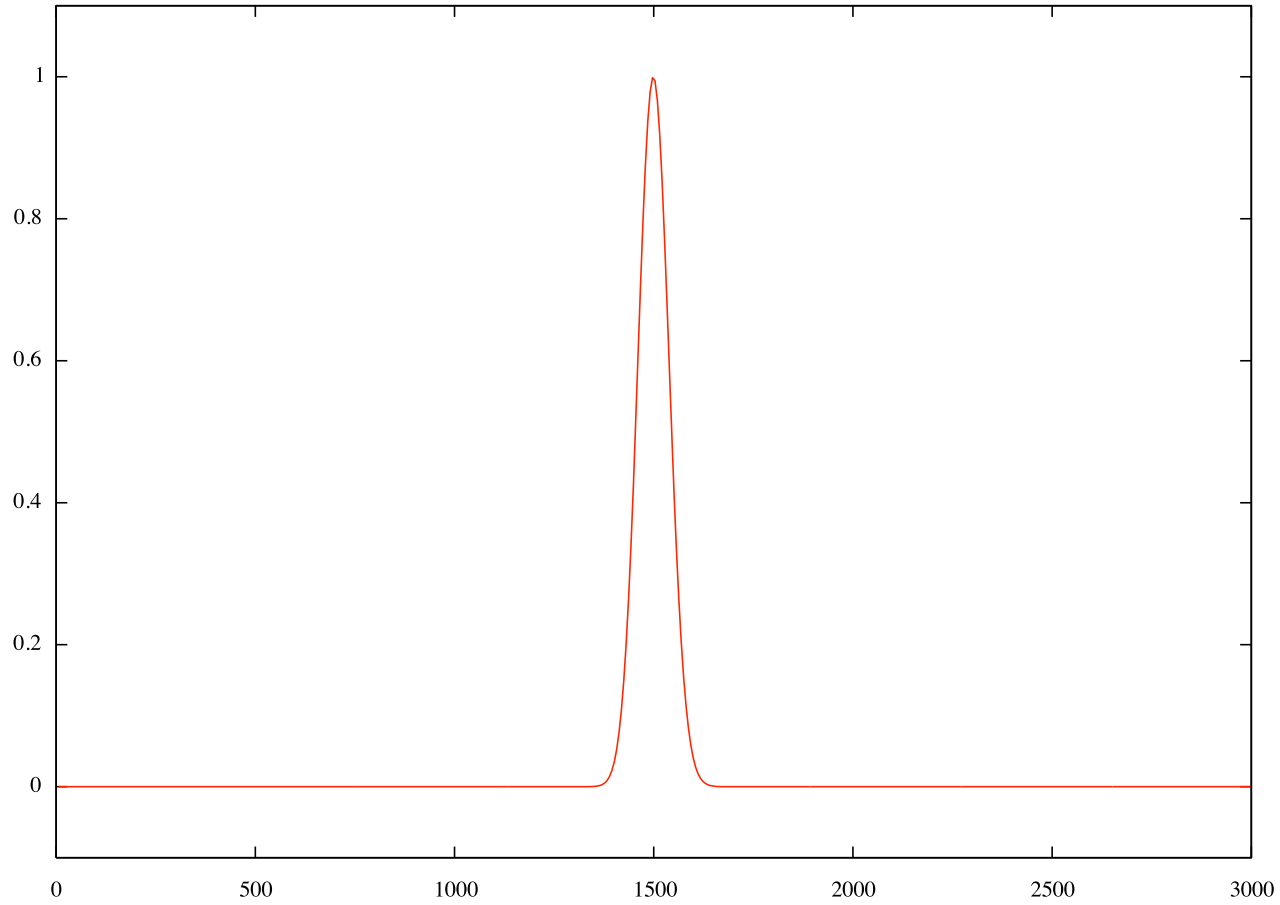
III. Molecular dynamics

IV. Replica exchange method

V. Multicanonical/Wang-Landau sampling

VI. Practicalities

# Energy distribution for canonical ensemble




Spread is $\propto \sqrt{N} k_{\mathrm{B}} T$.

# Multicanonical sampling

Energy barriers can be overcome by sampling from a density for which $U(x)$ has a flat distribution.

For suitable ranges $E_{\min} \leq U(x) \leq E_{\max}$, there is such a density,

$$\rho_{\text{multi}}(x) = \text{const}/g(U(x)),$$

where $g(E)$ is the (configurational) density of states,

$$g(E) = \text{const} \int \delta(U(x) - E)\mathrm{d}x.$$

# Reweighting

If we have an approximation to $g(E)$,
canonical ensemble averages can be calculated by reweighting

$$\langle A(x) \rangle = \frac{\langle A(x) \mathrm{e}^{-\beta U(x)} / g(U(x))^{-1} \rangle_{\mathrm{multi}}}{\langle \mathrm{e}^{-\beta U(x)} / g(U(x))^{-1} \rangle_{\mathrm{multi}}},$$

valid independent of the accuracy of $g(E)$.
(If $g(E)$ has good accuracy, there is an alternative.)

# Wang-Landau scheme

Discretize using histograms:

partition $E_{\min} \leq E \leq E_{\max}$ into $J$ subintervals,

define basis functions $1_j(E) = \begin{cases} 1, & \text{in } j\text{th subinterval,} \\ 0, & \text{elsewhere.} \end{cases}$

Construct approximations, $g_{(1)}(E)$, $g_{(2)}(E)$, $\ldots$, of decreasing *granularity*:

$$\log g_{(1)}(E) = \sum_{j=1}^{J} N_j^{(1)} 1_j(E),$$

$$\log g_{(2)}(E) = \log g_{(1)}(E) + \frac{1}{2} \sum_{j=1}^{J} N_j^{(2)} 1_j(E),$$

et cetera.

# Inner loop

Initialize: $g^0(E) = g_{(k)}(E)$.

Inner loop: for $n = 1, 2, \ldots$

Generate proposal $\bar{x}^n$ from $x^{n-1}$

Choose $x^n = \bar{x}^n$ or $x^{n-1}$ based on

Metropolis criterion for density $= \text{const}/g^{n-1}(U(x))$

\# increment histogram value

$\log g^n(E) = \log g^{n-1}(E) + \left(\frac{1}{2}\right)^k \sum_{j=1}^{J} 1_j(U(x^n)) 1_j(E)$,

until well sampled.

Set $g_{(k+1)}(E) = g^n(E)$.

not a Markov chain

Scheme tries to create a flat energy distribution,

the resistance encountered is a correction to $g_{(k)}(E)$.

# Outline

I. Markov chain Monte Carlo methods

II. Hybrid Monte Carlo methods

III. Molecular dynamics

IV. Replica exchange method

V. Multicanonical/Wang-Landau sampling

VI. Practicalities

# Practicalities

The practicalities of doing such calculations involve three steps:

**structure building** Setting up the input files is best done interactively with scripts and visual feedback.
visualization programs: RasMol, VMD, PyMOL, ...

**simulation** Generating dynamics or sampling trajectories is best done in background or remotely.
simulation programs: CHARMM, Amber, Gromacs, NAMD, LAMMPS, NWChem, Tinker, ...

**analysis** Analyzing trajectory data.

# Simulation specifications

- Specify molecular system & surroundings
- Specify computational tasks
- Select computational model:
  uncontrolled approximations and error tolerances
  - internal forces
  - external forces, e.g., temperature and pressure control
  - dynamics (sampling or real)
- (Override defaults for performance parameters)
- Design simulation protocol

# References


- M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, 1987,

- D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd edition, 2002.

- A. R. Leach, *Molecular Modelling: Principles and Applications*, 2nd edition, 2001,

- T. Schlick, *Molecular Modeling and Simulation: An Interdisciplinary Guide*, 2002,

- *Journal of Chemical Physics*