

Improved Hybrid Monte Carlo method for conformational sampling



Jesús A. Izaguirre
with Scott Hampton

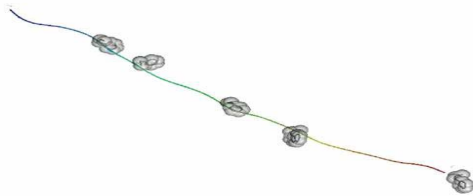
Department of Computer Science and Engineering
University of Notre Dame

July 1, 2003

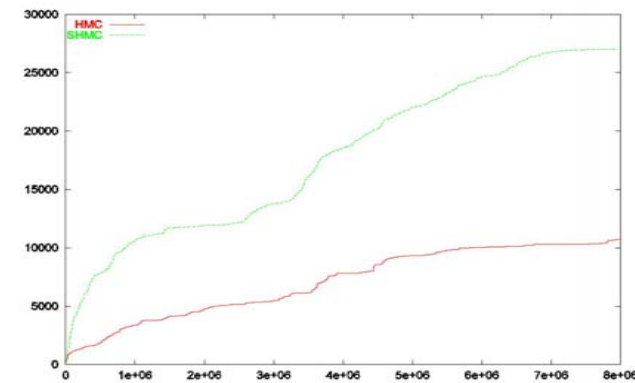
This work is partially supported by two NSF grants (CAREER and BIOCOMPLEXITY)
and two grants from University of Notre Dame

Overview

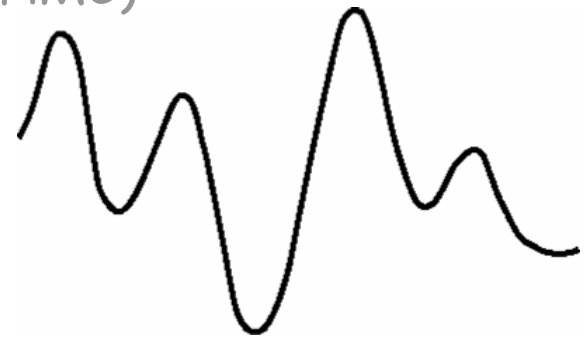
1. Methods for sampling conformational space



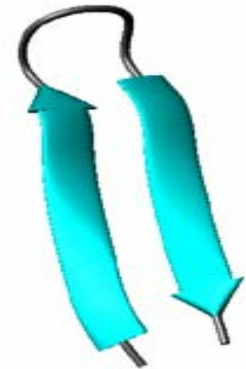
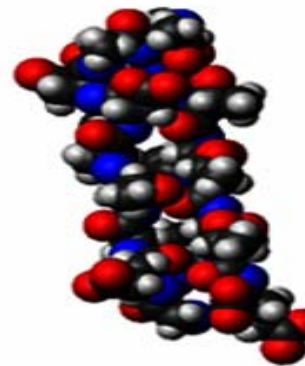
3. Evaluation of SHMC



2. New Shadow HMC (SHMC)

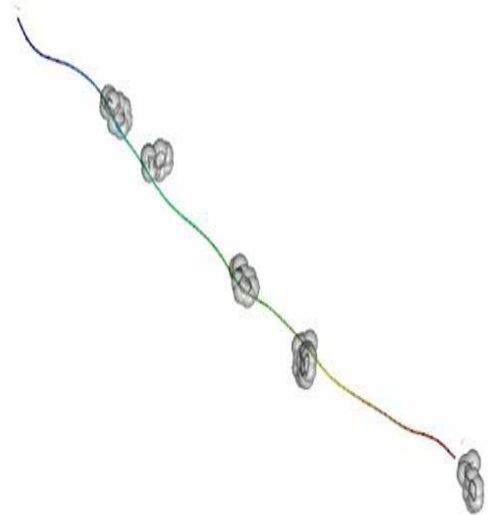


4. Discussion



Questions related to sampling

- Sampling
 - Compute equilibrium averages in NVT (or other) ensemble
 - Examples:
 - Equilibrium distribution of solvent molecules in vacancies
 - Free energies
 - Pressure
 - Characteristic conformations

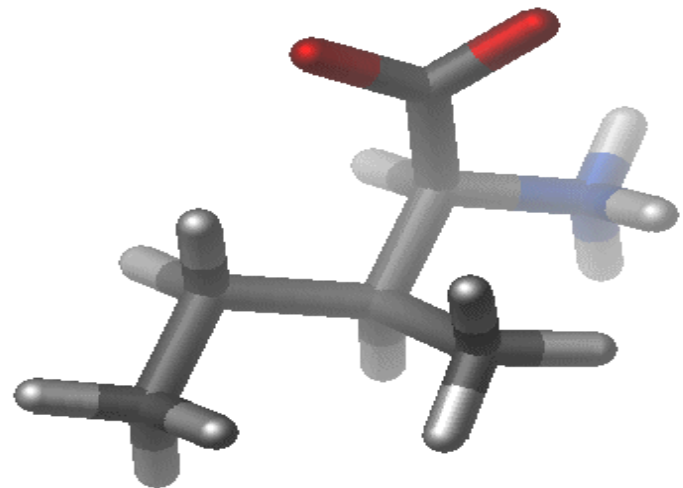


Classical molecular dynamics

- Newton's equations of motion: $\mathbf{M}\mathbf{q}'' = -\nabla U(\mathbf{q}) = \mathbf{F}(\mathbf{q}). \dots (1)$

- Atoms
- Molecules
- CHARMM force field

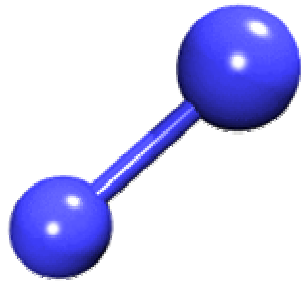
(Chemistry at Harvard
Molecular Mechanics)



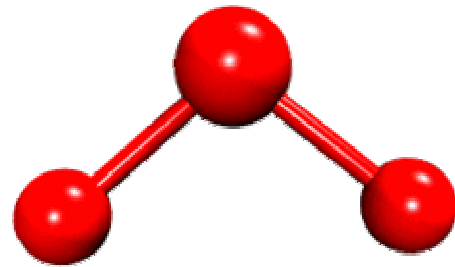
Bonds, angles and torsions

Energy Terms Described in the CHARMM forcefield

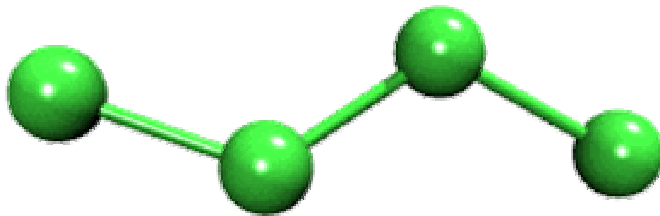
Bond



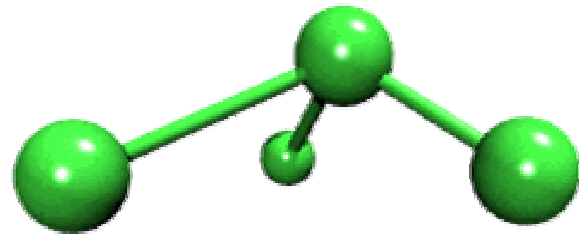
Angle




Dihedral



Improper



Energy Functions



$$\begin{aligned}
 U(\vec{R}) = & \underbrace{\sum_{\text{bonds}} k_i^{\text{bond}} (r_i - r_0)^2}_{U_{\text{bond}}} + \underbrace{\sum_{\text{angles}} k_i^{\text{angle}} (\theta_i - \theta_0)^2}_{U_{\text{angle}}} + \\
 & \underbrace{\sum_{\text{dihedrals}} k_i^{\text{dihe}} [1 + \cos(n_i \phi_i + \delta_i)]}_{U_{\text{dihedral}}} + \\
 & \underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{\text{nonbond}}}
 \end{aligned}$$

U_{bond} = oscillations about the equilibrium bond length

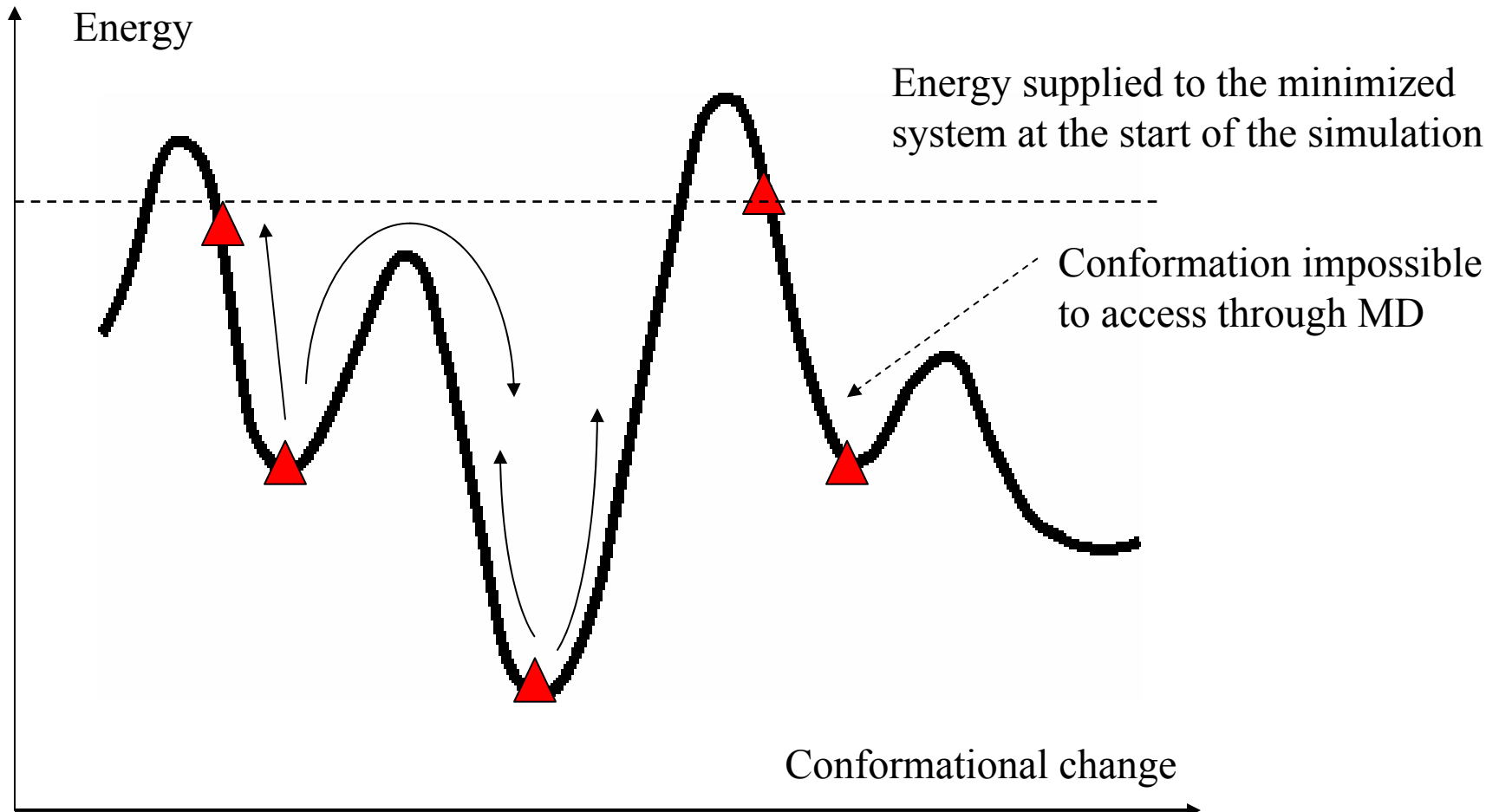
U_{angle} = oscillations of 3 atoms about an equilibrium angle

U_{dihedral} = torsional rotation of 4 atoms about a central bond

U_{nonbond} = non-bonded energy terms (electrostatics and Lennard-Jones)

Molecular Dynamics - what does it mean?

MD = change in conformation over time using a forcefield





Hybrid Monte Carlo I

We can sample from a distribution with density $\rho(x)$ by simulating a Markov Chain with the following transitions:

(1) From the current state, x , a candidate state x' is drawn from a proposal distribution $S(x, x')$. The proposed state is accepted with probability

$$\min\left(1, \frac{p(x')S(x', x)}{p(x)S(x, x')}\right)$$

(2) If the proposal distribution is symmetric, $S(x, x') = S(x', x)$, then the acceptance probability only depends on $p(x')/p(x)$



Hybrid Monte Carlo II

- (1) Proposal functions must be *reversible* :
if $x' = s(x)$ then $x = s(x')$
- (2) Proposal functions must *preserve volume*:
the Jacobian of the proposal must have absolute value 1

Examples:

- (a) Valid proposal: $x' = -x$
- (b) Invalid proposals: $x' = 1/x$ (Jacobian not 1)
 $x' = x + 5$ (not reversible)



Hybrid Monte Carlo III

Hybrid Monte Carlo:

- Apply stochastic step (e.g., regenerate momenta)
- Use reversible symplectic integrator for MD to generate the next proposal in MC:
 - Hamiltonian dynamics preserve volume in phase space, and so do symplectic integrators (determinant of Jacobian of map is 1)
 - It is simple to make symplectic integrators time reversible
- Apply Metropolis MC acceptance criterion



Hybrid Monte Carlo IV

Advantages of HMC:

- HMC can propose and accept distant points in phase space
 - Make sure new SHMC has high enough accuracy
- HMC can move in a biased way, rather than in a random walk like MC (distance n vs \sqrt{n})
 - Make L long enough in SHMC
- HMC is a rigorous sampling method: systematic sampling errors due to finite step size in MD are eliminated by the Metropolis step of HMC.
 - Make sure bias is eliminated by SHMC



Hybrid Monte Carlo V

The probability of acceptance in HMC decays exponentially with system size N and time step δt . For Verlet/leapfrog is:

$$P(N, \delta t) = 1 - \operatorname{erf}\left(0.5 \sqrt{\frac{N \delta t^2}{k_B T}}\right).$$

Analysis of N replicas of multivariate Gaussian distributions shows that HMC takes $O(N^{5/4})$ with time step $\delta t = O(N^{-1/2})$

(**Kennedy & Pendleton, 91**)

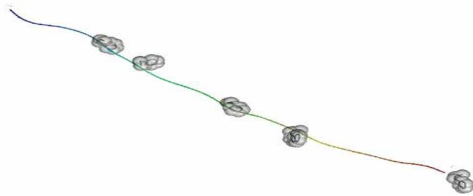


Hybrid Monte Carlo VII

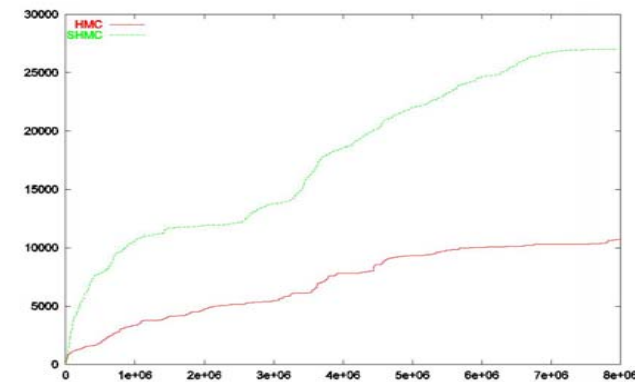
- The key problem in scaling is the *accuracy* of the MD integrator
- Higher order MD integrators could help scaling
- **Creutz and Gocksch (1989)** proposed higher order symplectic methods to improve scaling of HMC
- In MD, however, these methods are more expensive than the gain due to the scaling. They need several force evaluations per step
 - *$O(N)$ electrostatic methods may make higher order integrators in HMC feasible for large N*

Overview

1. Methods for sampling conformational space



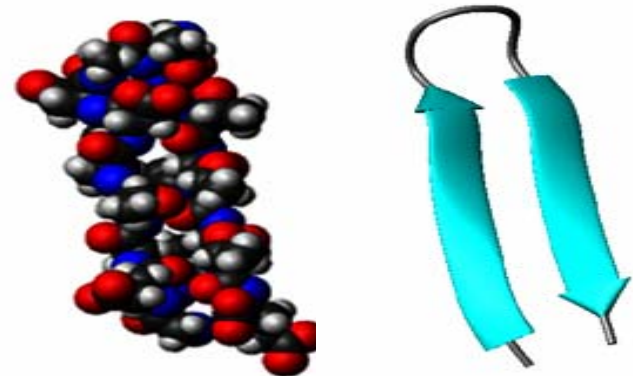
3. Evaluation of SHMC



2. New Shadow HMC (SHMC)



4. Discussion





Improved HMC

- Symplectic integrators conserve exactly (within roundoff error) a modified Hamiltonian that for short MD simulations (such as in HMC) stays close to the true Hamiltonian **Sanz-Serna & Calvo 94**
- Our idea is to use highly accurate approximations to the modified Hamiltonian in order to improve the scaling of HMC



Shadow Hamiltonian

Work by **Skeel & Hardy (2001)**, shows how to compute an arbitrarily accurate approximation to the modified Hamiltonian integrated by symplectic integrators based on splitting, which they call *shadow Hamiltonian*:

$$\mathcal{H}(q,p) = \frac{1}{2}p^T M^{-1}p + U(q)$$

$$\mathcal{H}_{2m}^S(q,p) = \mathcal{H}(q,p) + O(\delta t^{2m})$$

The \mathcal{H}^S of order $2m$ requires m extra copies of past positions and velocities, and requires to propagate one extra degree of freedom with the momenta, called β , using the MD integrator.

For example, $\mathcal{H}_8^S(z) = \mathcal{H}(z) + f(z^{n-1}, z^{n-2}, z^{n-3}, z^{n-4})$, where $z = [q \ p \ \beta]$.



Example Shadow Hamiltonian (partial)

$$SH_s = A_{10} + \frac{5}{42} * A_{30} + \frac{13}{105} * A_{32} - \frac{2}{7} * A_{12} - \frac{19}{210} * A_{14} - \frac{1}{140} * A_{34}$$

$$\begin{aligned} A_{10} &= X_1 \cdot V_0 - V_1 \cdot X_0 - \frac{1}{2}(\beta^{n+1} - \beta^{n-1}) \\ &= \left(\frac{1}{2M}(X^{n+1} - X^{n-1}) \cdot V^n\right) - \left(\frac{1}{2M}(V^{n+1} - V^{n-1}) \cdot X^n\right) - \frac{1}{2}(\beta^{n+1} - \beta^{n-1}) \end{aligned}$$



SHMC Algorithm

- (1) Draw random values for the momenta p from Gaussian distribution. Start from current positions and momenta (q, p)
- (2) Perform an MD simulation of length L and time step δt to get new positions and momenta (q', p') using a symplectic reversible integrator such as leapfrog
- (3) Accept new state (q', p') with probability $P(q', p') = \min(1, \exp(-\beta\delta\mathcal{H}^s))$,
where $\delta\mathcal{H}^s = \mathcal{H}^s(q', p') - \mathcal{H}^s(q, p)$
- (4) Repeat (1) - (3) M times
- (5) Do a reweighting to obtain proper canonical distributions:

$$\langle A \rangle_{NVT} = \frac{\frac{1}{M} \sum_{i=1}^M \frac{A_i}{\exp[-\beta\mathcal{H}^s(q_i, p_i)]} \exp[-\beta\mathcal{H}(q_i, p_i)]}{\frac{1}{M} \sum_{i=1}^M \frac{1}{\exp[-\beta\mathcal{H}^s(q_i, p_i)]} \exp[-\beta\mathcal{H}(q_i, p_i)]}$$

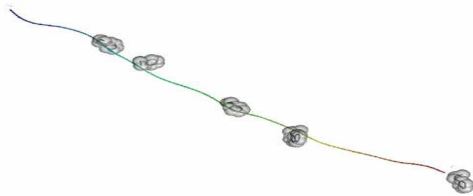


SHMC

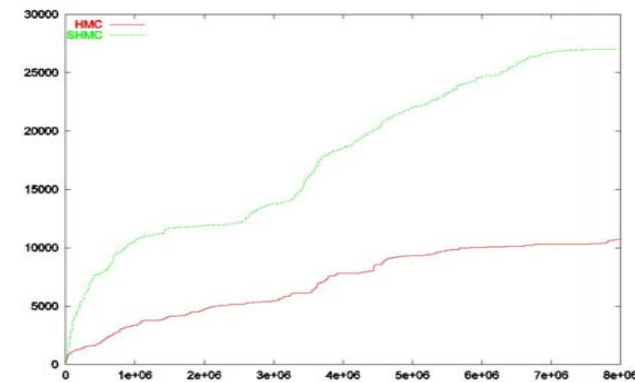
- Nearly linear scalability of acceptance rate with system size N
- Computational cost of SHMC, $O(N^{(1+1/2m)})$ where m is accuracy order of integrator
- Extra storage (m copies of q and p)
- Moderate overhead (10% for medium protein such as BPTI)

Overview

1. Methods for sampling conformational space



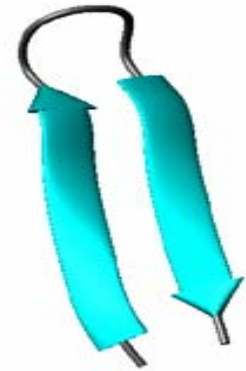
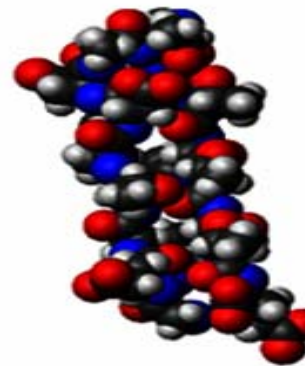
3. Evaluation of SHMC



2. New Shadow HMC (SHMC)



4. Discussion





Evaluating MC methods I

1. Is SHMC sampling from desired distribution?
 - Does it preserve detailed balance?
 - Used simple model systems that can be solved analytically. Compared to analytical results and HMC. Examples: *Lennard-Jones liquid, butane*
 - Is it ergodic?
 - Impossible to prove for realistic problems. Instead, show *self-averaging* of properties. *Computed self-averaging of non-bonded forces and potential energy*

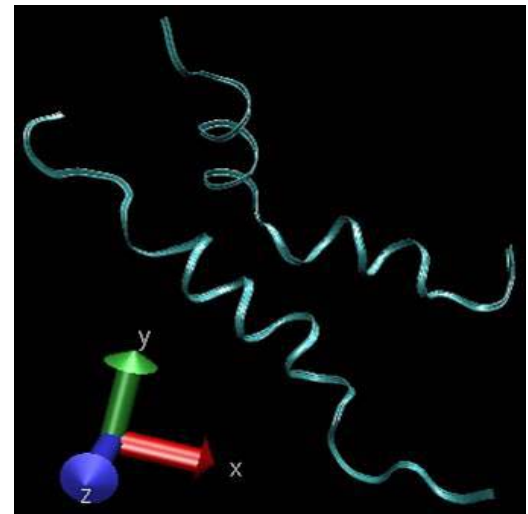
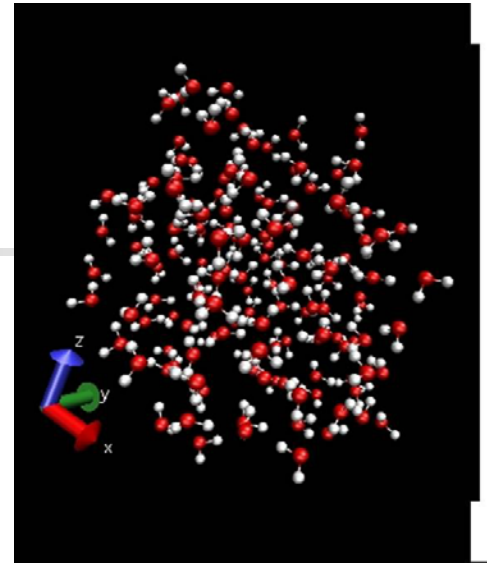
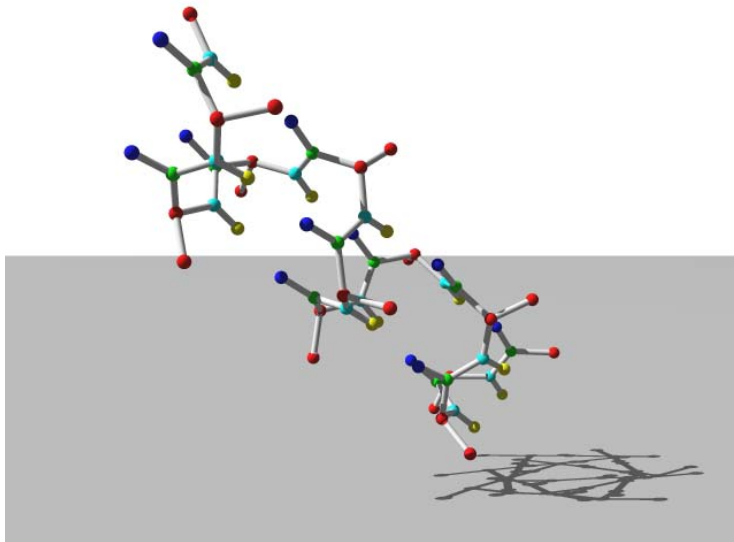


Evaluating MC methods II

2. Is system equilibrated?
 - Average values of set of properties fluctuate around mean value
 - Convergence to steady state from
 - Different initial conditions
3. Are statistical errors small?
 - Runs about 10 times longer than slowest relaxation in system
 - Estimated statistical errors by block averaging
 - Computed properties (torsion energy, pressure, potential energy)
 - Vary system sizes (4 - 1101 atoms)
4. What are the sampling rates?
 - Cost (in seconds) per new conformation
 - Number of conformations discovered



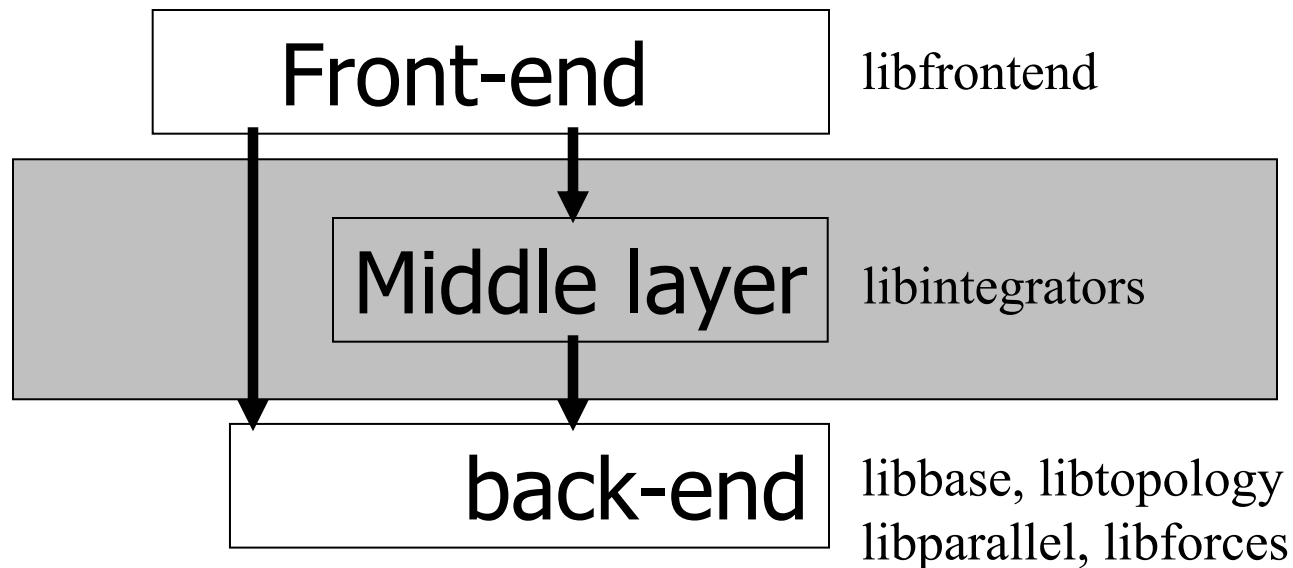
Systems tested





ProtoMol: a framework for MD

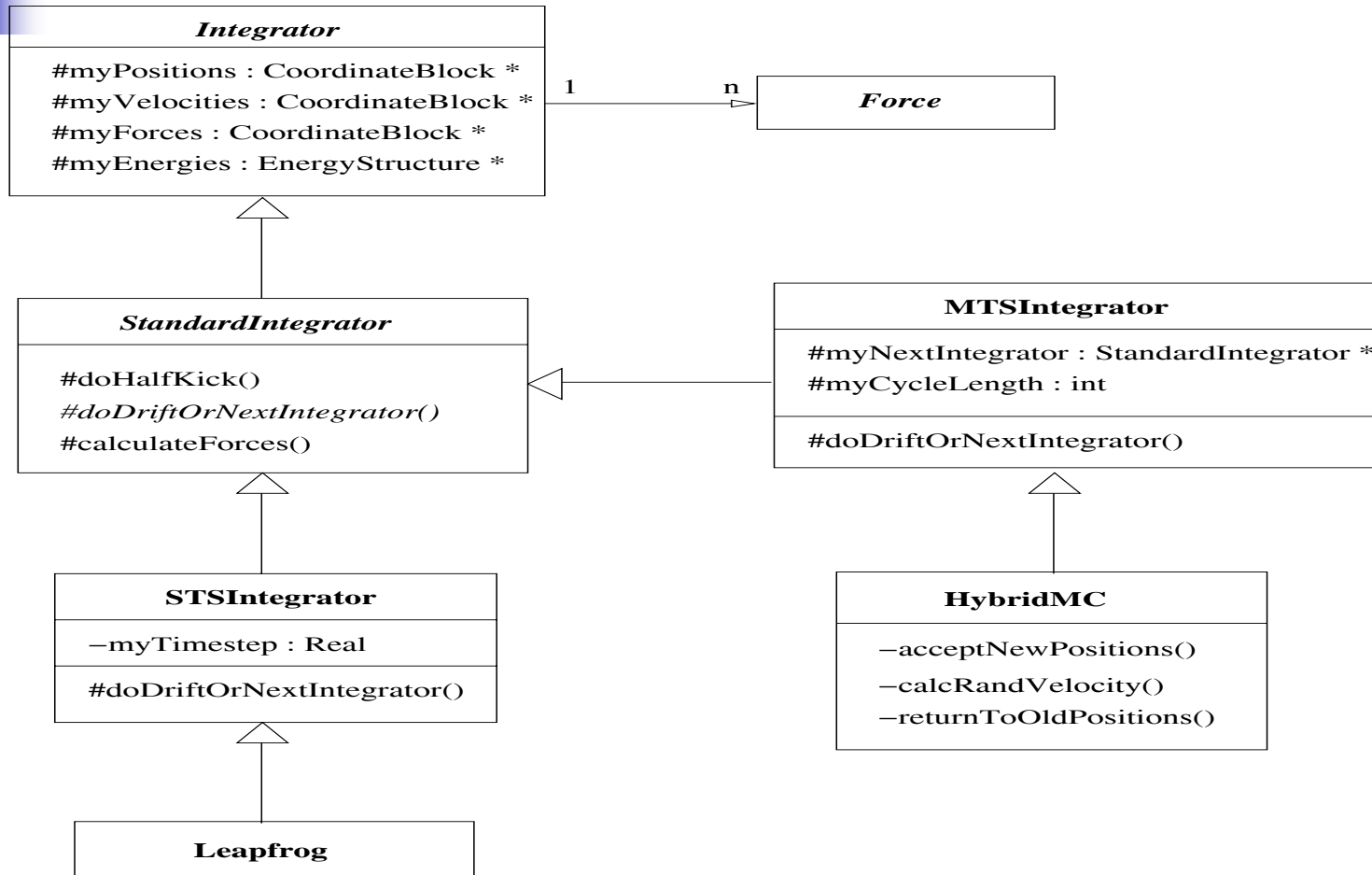
Matthey, et al, *ACM Tran. Math. Software (TOMS)*, submitted



Modular design of **ProtoMol** (Prototyping Molecular dynamics).

Available at <http://www.cse.nd.edu/~lcls/protomol>

SHMC implementation



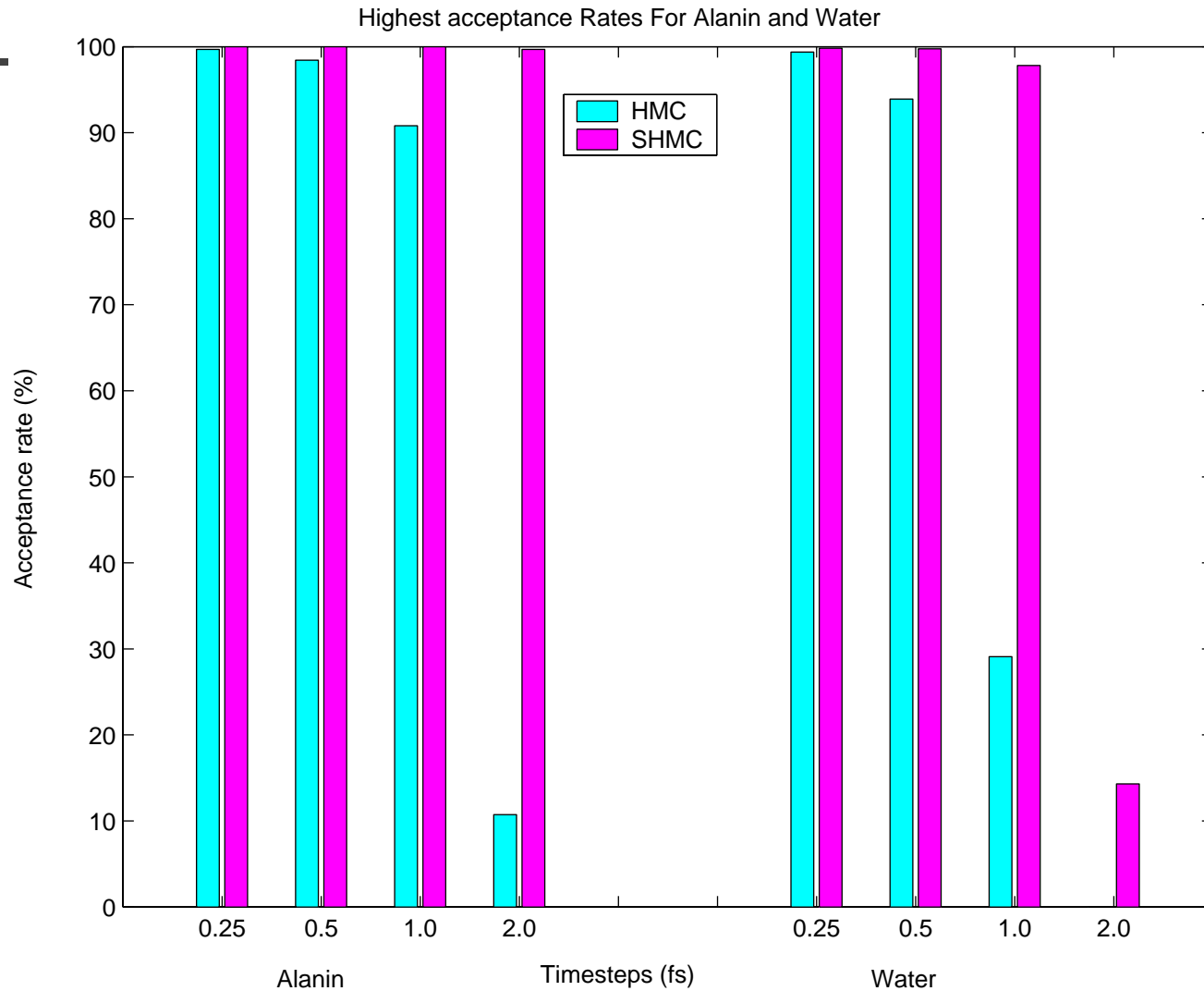


Experiments: acceptance rates I

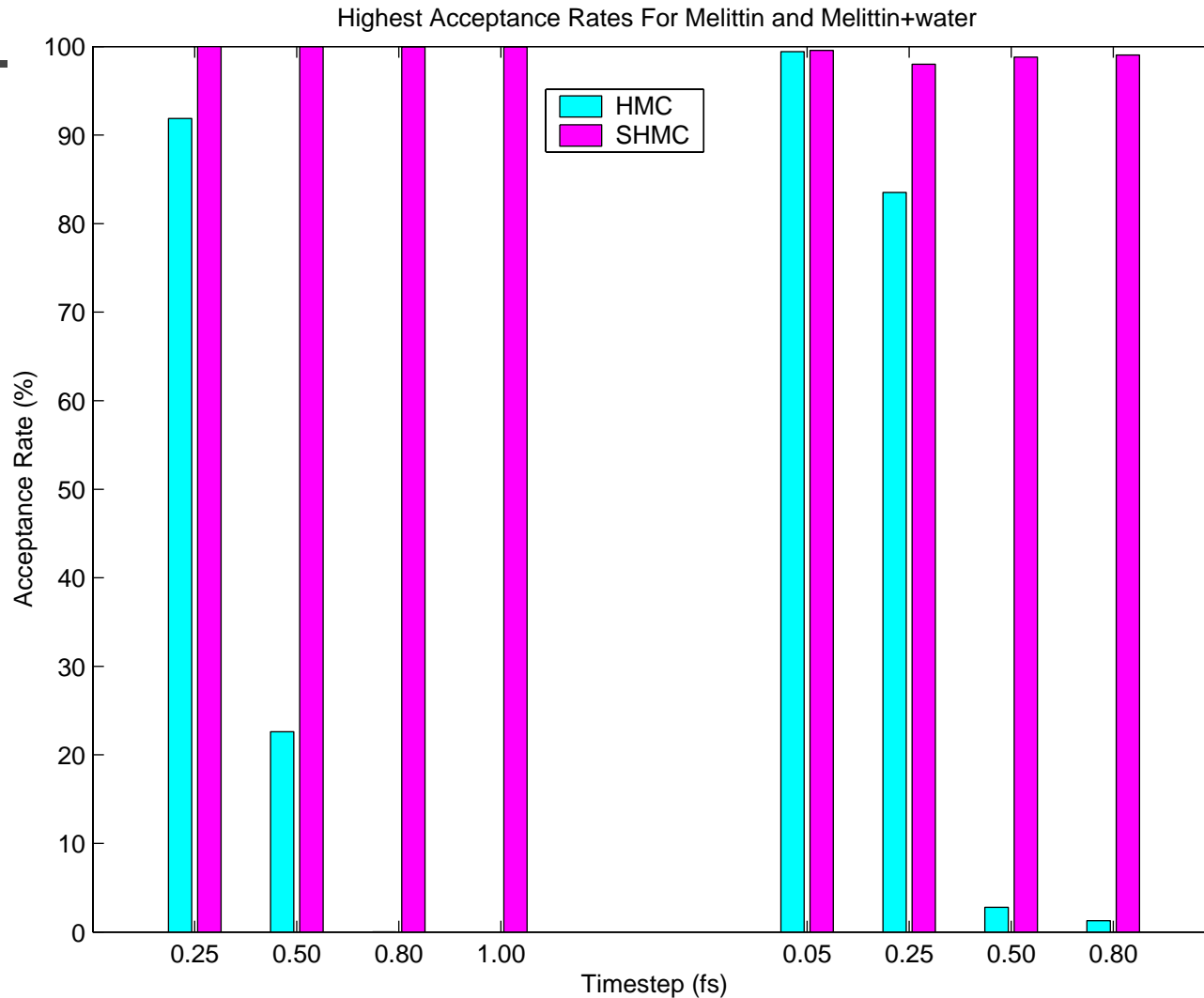
Numerical experiments confirm the predicted behavior of the acceptance rate with system size. Here, for fixed acceptance rate, the maximum time step for HMC and SHMC is presented

<i>System</i>	<i>N</i>	H M C δt	S H M C δt
Butane	4	2	6
Alanine	66	0.5	2
Water	423	0.25	1.07
Melittin	828	0.15	1.03
Solvated Melittin	5143	0.05	0.8

Experiments: acceptance rates II



Experiments: acceptance rates III

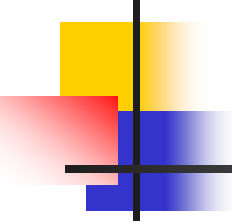




Average of observable

- Average torsion energy for extended atom Butane (CHARMM 28)
- Each data point is a 114 ns simulation
- Temperature = 300 K

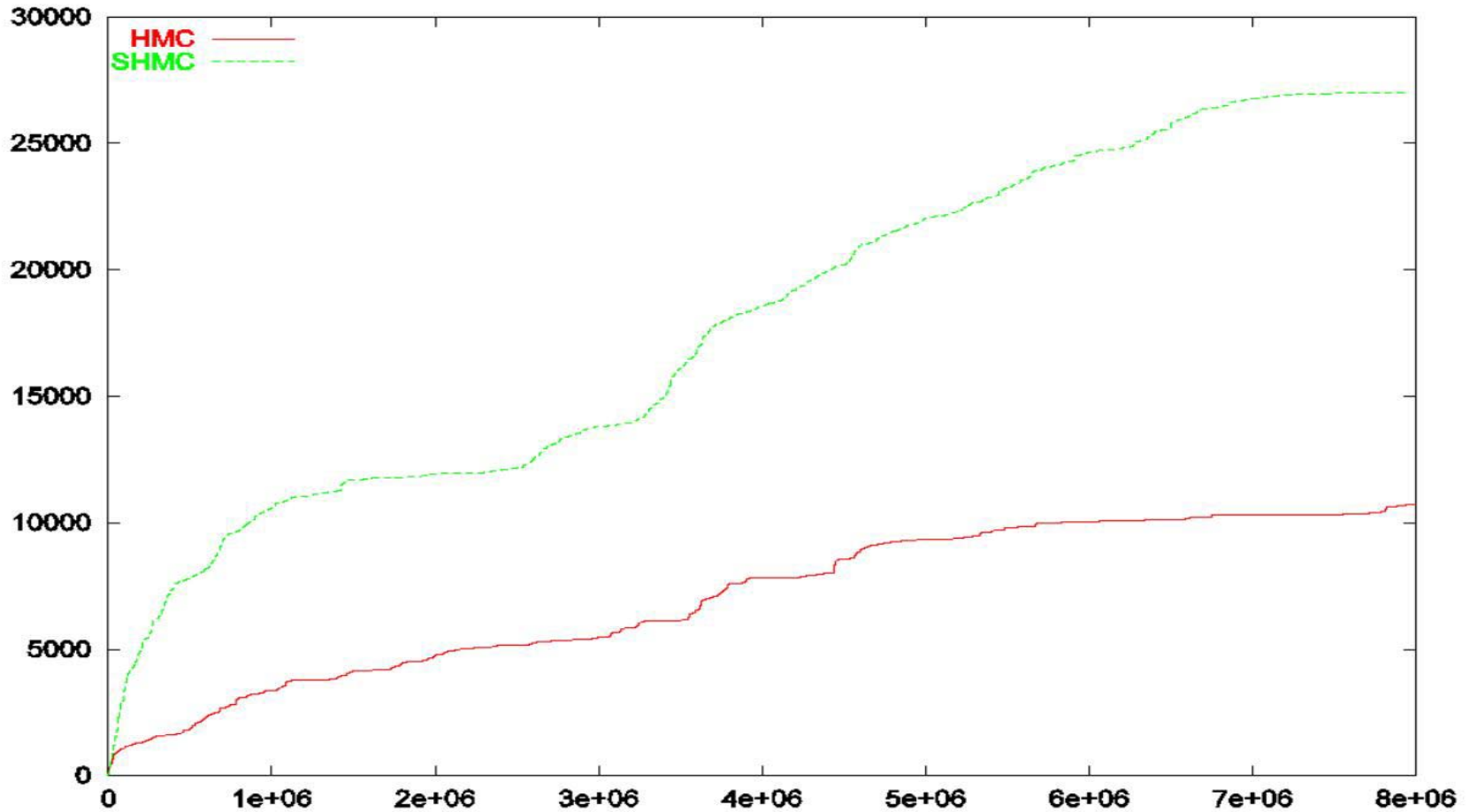
	L = 630 fs		L = 72 fs	
δt (fs)	H M C	S H M C	H M C	S H M C
8	0.65 ± 0.03	0.67 ± 0.02	0.65 ± 0.03	0.65 ± 0.08
(A c c)	(51%)	(99%)	(51%)	(99%)
6	0.63 ± 0.02	0.64 ± 0.02	0.62 ± 0.02	0.63 ± 0.02
(A c c)	(79%)	(100%)	(79%)	(100%)
3	0.63 ± 0.02	0.63 ± 0.02	0.63 ± 0.02	0.62 ± 0.02
(A c c)	(96%)	(100%)	(96%)	(100%)
1	0.62 ± 0.02	0.64 ± 0.02	0.64 ± 0.02	0.62 ± 0.02
(A c c)	(100%)	(100%)	(100%)	(100%)



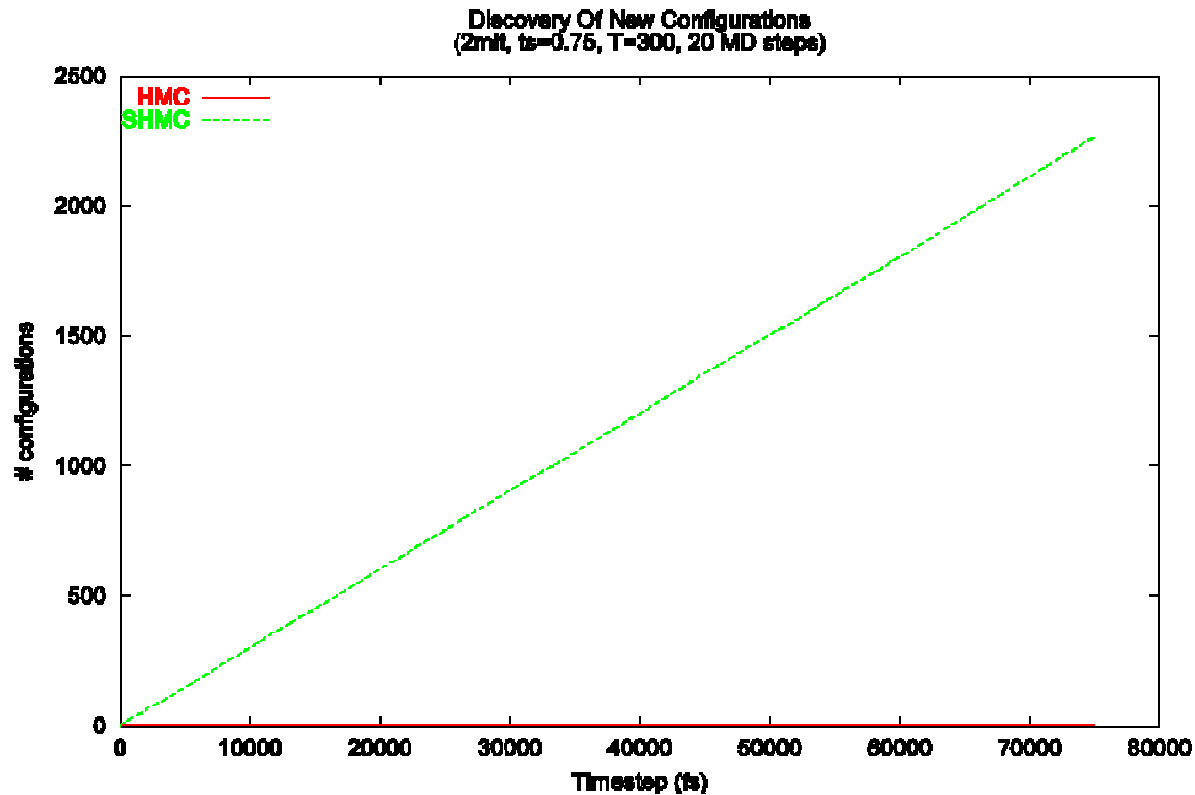
Sampling Metric (or how to count conformations)

- For each dihedral angle (not including Hydrogen) do this preprocessing:
 - Find local maxima, counting periodicity
 - Label 'wells' between maxima
- During simulation, for each dihedral angle $\text{Phi}[i]$:
 - Determine 'well' $\text{Phi}[i]$ occupies
 - Assign name of well to a conformation string
- String determines conformation (extends method by **McCammon *et al.*, 1999**)

Sampling rate for decalanine ($dt = 2$ fs)



Sampling rate for 2mlt





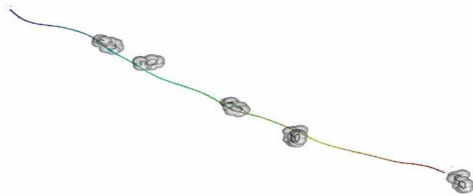
Sampling rate comparison

- C is number of new conformations discovered
- $Cost$ is total simulation time divided by C
- Each row found by sweeping through step size (for alanine, between 0.25 and 2 fs; for melittin and bpti between 0.1 and 1 fs) and simulation length L

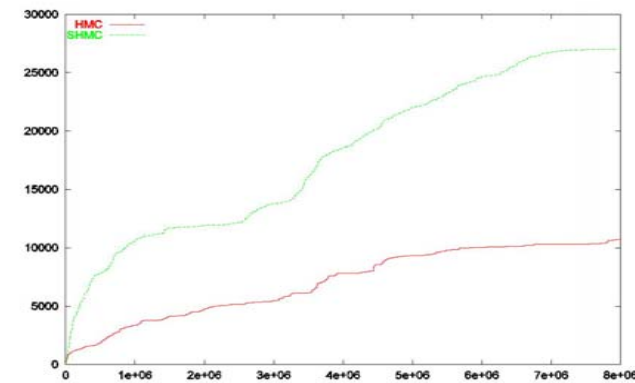
	HMC		SHMC		Speedup
<i>System</i>	C	Cost/ C	C	Cost/ C	
Alanine	985	4 s	992	0.86 s	4.8
Melittin	270	122 s	2340	16 s	7.6
BPTI	289	145 s	1146	16.9 s	8.6

Overview

1. Methods for sampling conformational space



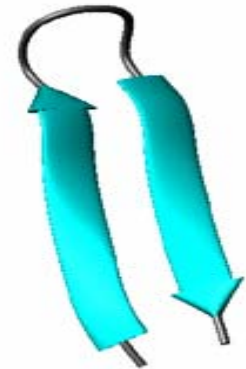
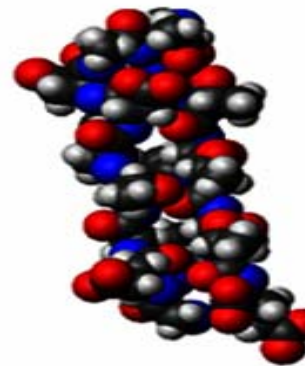
3. Evaluation of SHMC



2. New Shadow HMC (SHMC)



4. Discussion





Summary and discussion

- SHMC has a higher acceptance rate than HMC, particularly as system size and time step increase
- SHMC discovers new conformations more quickly
- SHMC requires extra storage and moderate overhead.
- For large time steps, weights may increase, which harms the variance. This dampens maximum speedup attainable
- More careful coding is needed for SHMC than HMC
- For large N , higher order integrators may be competitive with SHMC
- Instead of reweighting, one may modify the acceptance rule



Future work

- Multiscale problems for rugged energy surface
 - Multiple time stepping algorithms plus constraining
 - Temperature tempering and multicanonical ensemble (e.g., method of Fischer, Cordes, & Schutte 1999)
 - Potential smoothing
 - Combine multiple SHMC runs using method of histograms
 - Include other MC moves (e.g., change essential dihedrals or Chandler's moves)
- System size
 - Parallel multigrid or multipole $O(N)$ electrostatics
- Applications
 - Free energy estimation for drug design
 - Folding and metastable conformation



Acknowledgments

- Graduate student: Scott Hampton
- Dr. Thierry Matthey, co-developer of ProtoMol, University of Bergen, Norway
- Students in CSE 598K, "Computational Biology," Spring 2002
- Tamar Schlick for her delightful new book, *Molecular Modeling and Simulation - An Interdisciplinary Guide*
- Dr. Robert Skeel, Dr. Ruhong Zhou, and Dr. Christoph Schutte for valuable discussions
- Dr. Radford Neal's presentation "Markov Chain Sampling Using Hamiltonian Dynamics" (<http://www.cs.utoronto.ca>)
- Dr. Klaus Schulten's presentation "An introduction to molecular dynamics simulations" (<http://www.ks.uiuc.edu>)