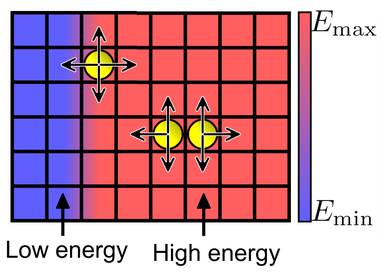
Karlsruhe Institute of Technology

Nanoscale and Biomolecular Simulation
Institute of Nanotechnology

Efficient time-resolved conformational sampling using a Kinetic Monte Carlo approach

Moritz Wolf[(1)], Konstantin Klenin[(2)], Timo Strunk[(1)], Wolfgang Wenzel[(1)]
[(1)]Institute of Nanotechnology, KIT, Germany
[(2)]Steinbuch Center for Computing, KIT, Germany

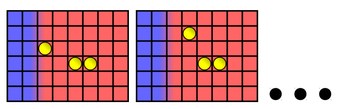
The Kinetic Monte Carlo method (KMC)[1]



Trivial systems:

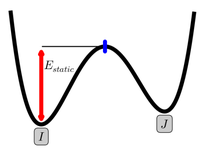
- 2D potential with discrete sites
- High barriers between this sites
- Short range interaction
- For each conformation of the system: Finite number of possible transitions

I. List of transitions depends on original conformation:



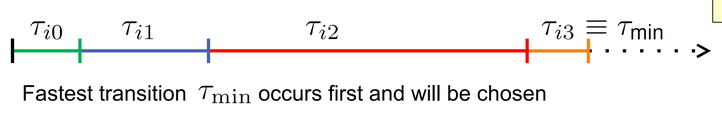
II. Using (H)TST to estimate rates:

$$k^{\text{HTST}} \sim e^{-\beta E_{\text{static}}}$$



III. For each transition i → j, a decay time τ_{ij} is calculated:

$$\tau_{ij} = -\frac{1}{r_{ij}} \cdot \ln(x-1) \qquad \text{, where } x \in [0,1)$$



Fastest transition $\tau_{\min}$ occurs first and will be chosen

IV. Compute expectation values:

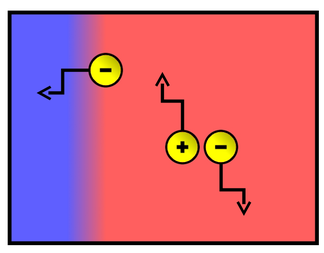
$$\langle E \rangle = \frac{1}{T} \int_0^T E(\tau)\, d\tau$$

Strategy:

I. Enumerate all possible transitions
II. Compute rates for each transition
III. Accept the fastest transition
IV. Repeat steps I-III to compute expectation values

☹ Discretized phase space needed
☹ All transitions must be known
☺ Easy to parallelize

The Diffusive Monte Carlo Method (DMC)



Complex Systems:

For high dimensional systems (e.g. proteins) complete transition lists cannot be assembled. Also the unknown number of dimensions makes it impossible to discretize the phase space. The resulting huge amount of states complicates the search for interesting low energy states.

Ideas:

- Replace full list by a dynamically generated representative subset of **N** final states
- Emphasize final states with high relative probabilities (**thermodynamic sampling**)
- New **rates** must correct sampling bias
- Sampling bias can be estimated only in a finite neighborhood Δ

New transition rates:

$$r_{I \to II} \equiv \frac{\sum_I e^{-\beta E}}{N e^{-\beta \min(E_I, E_{II})}}$$

$$r_{II \to I} \equiv \frac{\sum_{II} e^{-\beta E}}{N e^{-\beta \min(E_I, E_{II})}}$$

These rates lead to a prefactor in the detailed balance criterion that should be either monitored in a running simulation to be equal to one, or set to one by renormalizing the total escape rate.

Convergence conditions:

- **Complete sampling of the neighborhood**
- **Neighborhood must be barrier free**

☺ Continuous phase space
☹ A barrier-free subspace must be sampled correct
☺ Easy to parallelize
☺ Use of MC-sampling → long timescale transitions possible

1D, constant Potential

$V(x) = \text{const}$

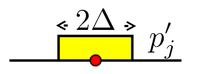
Diffusion

KMC: discretized space

DMC: random Population

Final states occur multiple times with probability:

$$p_j \sim \frac{e^{-\beta E_j}}{Z_i}$$



N finite area searches

KMC:

$$r_{ij} = \begin{cases} e^{-\beta(E_j - E_i)} & , E_i < E_j \\ 1 & , \text{else} \end{cases} \quad = 1$$

DMC:

$$r_{ij} = p'_j \begin{cases} \frac{1}{p_i} & , E_i < E_j \\ \frac{1}{p_j} & , \text{else} \end{cases} \quad = 1$$

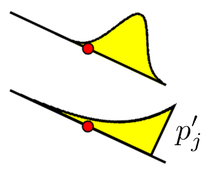
1D, locally linear potential

$V(x) = \text{linear}$

Simple energy landscape

KMC

DMC



N finite area searches

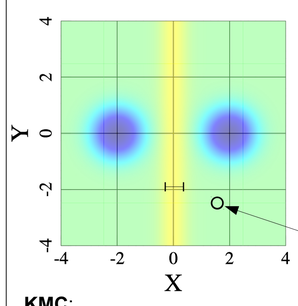
KMC:

$$r_{ij} \quad = \begin{cases} e^{-\beta(E_j - E_i)} & , E_i < E_j \\ 1 & , \text{else} \end{cases}$$

DMC:

$$r_{ij} = p'_j \begin{cases} \frac{1}{p_i} & , E_i < E_j \\ \frac{1}{p_j} & , \text{else} \end{cases} \quad = \begin{cases} e^{-\beta(E_j - E_i)} & , E_i < E_j \\ 1 & , \text{else} \end{cases}$$

Arbitrary landscape



Finite area search: using a rmsd potential to restrict sampling

Convergence condition:

- Sampled p_j must approximate exact p'_j
- Sampled area must be barrier free
  → $E_{\text{static}} \approx \Delta E$
  → Δ small vs. $\frac{\partial G}{\partial \vec{Q}}$

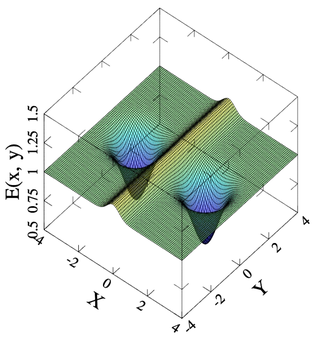
KMC:

$$r_{ij} \quad = \begin{cases} e^{-\beta E_{\text{static}}} & , E_i < E_j \\ 1 & , \text{else} \end{cases}$$

DMC:

$$r_{ij} = p'_j \begin{cases} \frac{1}{p_i} & , E_i < E_j \\ \frac{1}{p_j} & , \text{else} \end{cases} \quad = \begin{cases} e^{-\beta \Delta E} & , E_i < E_j \\ 1 & , \text{else} \end{cases}$$

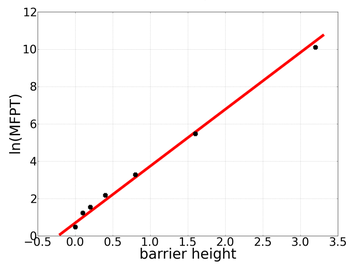
Simplified Model Problem



A two dimensional model system including two separated minima bordered by repulsive boundary conditions.

Kinetic:

Plotting the Mean First Passage Time versus the barrier height gives the expected exponential relation.
Also the calculation of the diffusion constant D give the right value, for a constant potential as well as for a linear one.



Multiplying this D with the time, a particle needs to go from one minimum to the other one and back, leads to the same results as using a shooting algorithm.

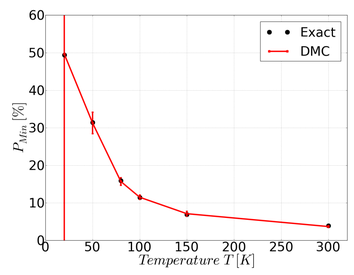
$$\langle \vec{R}^2 \rangle = 2dDt$$

Calculation of the diffusive constant for a constant and a linear potential:

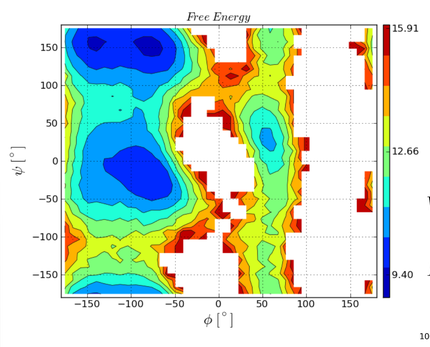
$$\langle R \rangle = \frac{FD}{kT} \cdot t$$

Thermodynamic:

The probability of staying in one of the minima gives the same results, using either the DMC method, or numerical integration.



Alanine dipeptide (N-acetyl-alanine-N'-methylamide)

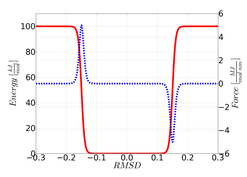


The Free Energy landscape of the alanine dipeptide shows four minima. Each two in the beta-sheet region and the helical sites.

The simulations were done in the AMBER99SB forcefield, using the GROMACS work package with its implicit solvent implementation. To restrict a single simulation to a certain subspace, we used PLUMED[2] to add a rmsd-based potential.

$$V(x) = \sigma\left(0.5\left(\tanh\left(\frac{x-a}{b}\right) - \tanh\left(\frac{x+a}{b}\right)\right) + 1\right)$$

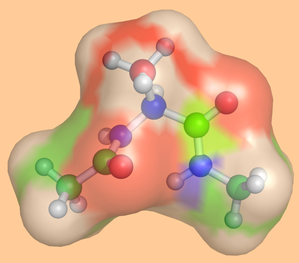
$$F(x) = -\frac{0.5\sigma}{b}\left(\left(\frac{1}{\cosh\left(\frac{x-a}{b}\right)}\right)^2 - \left(\frac{1}{\cosh\left(\frac{x+a}{b}\right)}\right)^2\right)$$



As we can't add a hard border potential to a MD simulation, this function was added to restrict the system to a finite area.

Conclusions & Outlook:

- Implementing a posteriori sampling quality check to see, if the parameters are chosen correctly
- Validating the DMC method in the alanine dipeptide system
- Running a DMC simulation of small proteins using MC-based simulations
- Benchmarking the calculation effort of DMC versus MD



References:

[1] Arthur F. Voter, “Introduction to kinetic monte carlo method” (Springer, December 4, 2008)

[2] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R.A. Broglia and M. Parrinello, PLUMED: a portable plugin for free energy calculations with molecular dynamics, Comp. Phys. Comm. 180, 1961 (2009)

Corresponding author address: Wolfgang Wenzel, Karlsruhe Institute of Technology Institute for Nanotechnology, Karlsruhe, Germany
Email: Wenzel@kit.edu, Moritz.Wolf@kit.edu

We thank the **Landesstiftung BW** for their support in the **HPC-5** program.