

Overview of fundamental concepts in molecular simulations

Neeraj Rai

Associate Professor and Ergon, Inc. Distinguished Professor
Dave C. Swalm School of Chemical Engineering
Mississippi State University

Newton said, “If I have seen further than others, it is because I've stood on the shoulders of giants.” These days we stand on each other's feet.

-Richard Hamming

Why do we need particle-based molecular simulations?

Why do we need particle-based molecular simulations?

- To understand physical phenomena
- To predict thermophysical and transport properties

*So, you have decided to use molecular simulation
for your research question, how should you
proceed?*

So, you have decided to use molecular simulation for your research question, how should you proceed?



Identify a
suitable
model

So, you have decided to use molecular simulation for your research question, how should you proceed?



Identify a
suitable
model

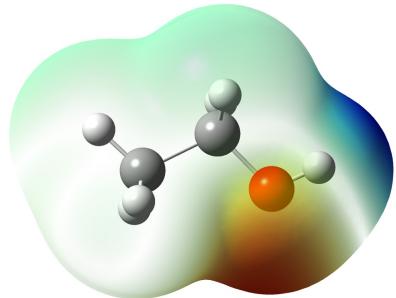
Identify a
suitable
method

So, you have decided to use molecular simulation for your research question, how should you proceed?

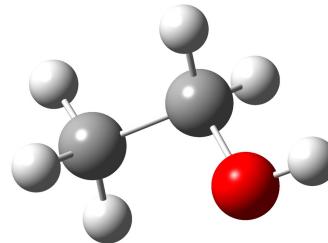


Identify a suitable model

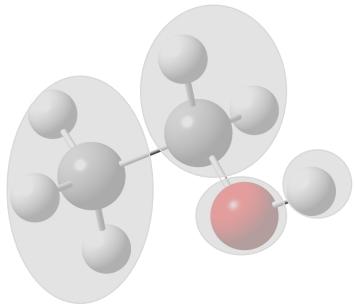
❖ Electronic Structure



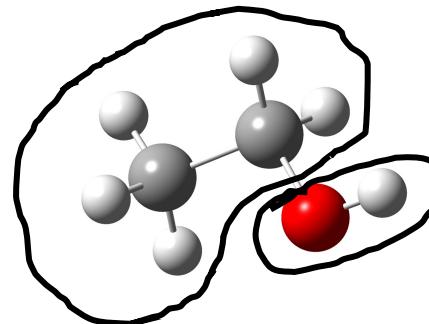
❖ All-atom



❖ United atom



❖ Coarse grained



Accuracy and cost are the two most important criteria for model selection

Identify a suitable method

**Molecular dynamics
(deterministic)**

**Monte Carlo
(Stochastic)**

- ❖ Transport properties
- ❖ Collective motion

- ❖ Phase equilibria
- ❖ Conformational sampling

Identify a suitable method

Molecular dynamics
(deterministic)

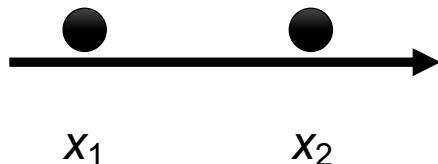
Monte Carlo
(Stochastic)

- ❖ Transport properties
- ❖ Collective motion
- ❖ Phase equilibria
- ❖ Conformational sampling

Molecular dynamics and Monte Carlo are two different approaches to sample the phase space, thus, referred to as sampling techniques

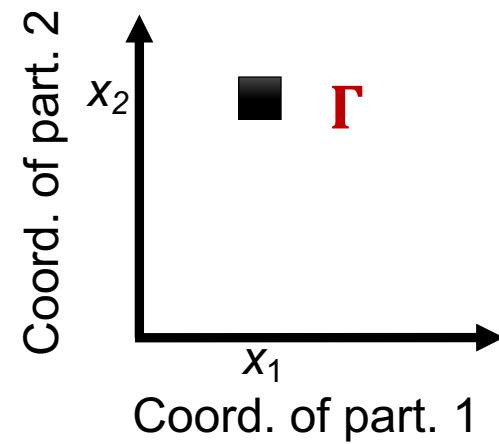
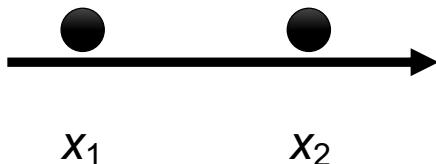
Phase space

Let's think of two particles in one-dimensional space with coordinates x_1 and x_2



Phase space

Let's think of two particles in one-dimensional space with coordinates x_1 and x_2



Configurational space

Phase space

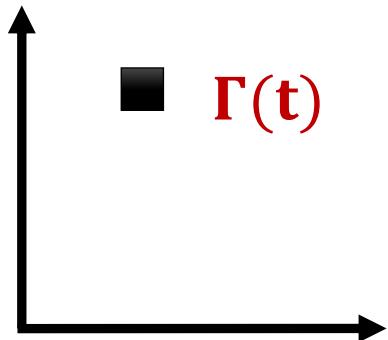
Let's think of N particles in three-dimensional space, we need:

Phase space

Let's think of N particles in three-dimensional space, we need:

$3N$ coordinates & $3N$ momenta

The $6N$ dimensional space is known as the phase space. The system at a given time is represented as a point ($\Gamma(t)$) in the phase space.



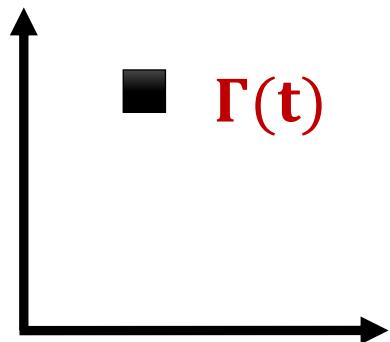
Phase space

Phase space

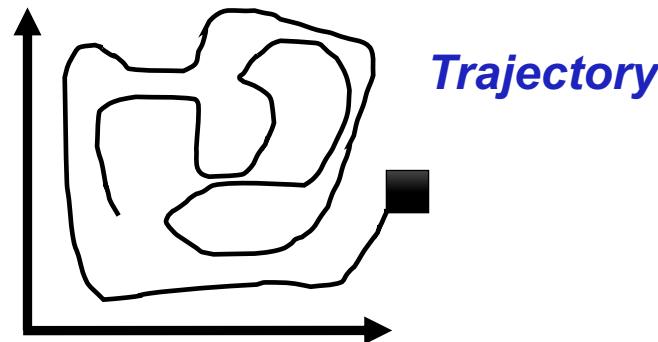
Let's think of N particles in three-dimensional space, we need:

$3N$ coordinates & $3N$ momenta

The $6N$ dimensional space is known as the phase space. The system at a given time is represented as a point ($\Gamma(t)$) in the phase space.

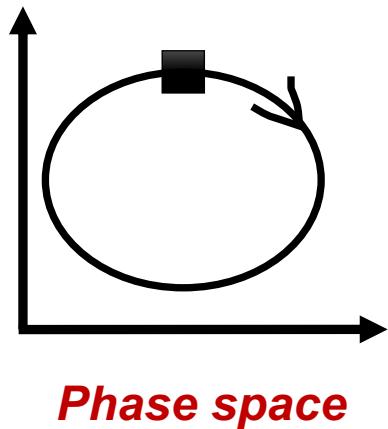


Phase space

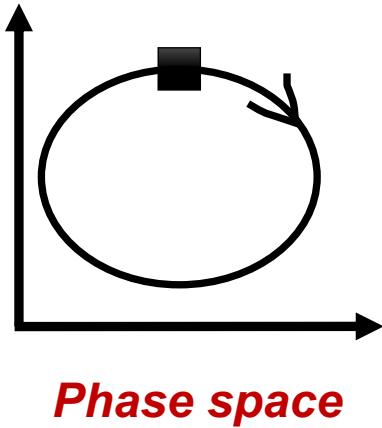


Phase space

Ergodicity



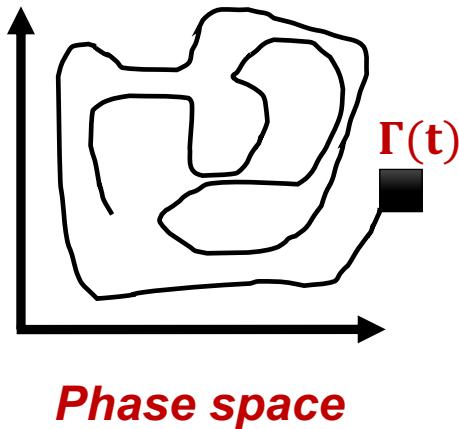
Ergodicity



This is a non-ergodic system!

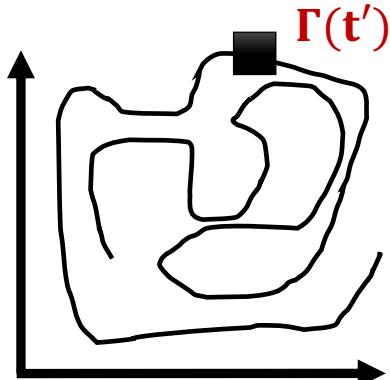
*Non-ergodicity can be an issue for both molecular dynamics
and Monte Carlo simulations.*

How to calculate thermodynamic or transport properties?



Do we have sufficient information to calculate the **total energy** of the system at this particular point ($\Gamma(t)$) in the phase space?

How to calculate thermodynamic or transport properties?



Phase space

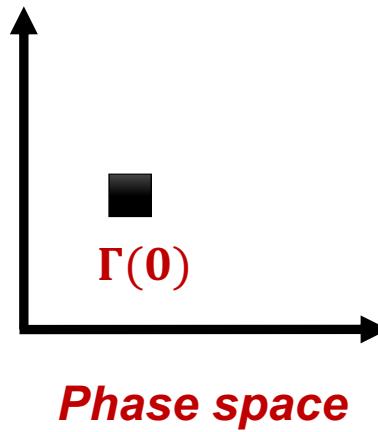
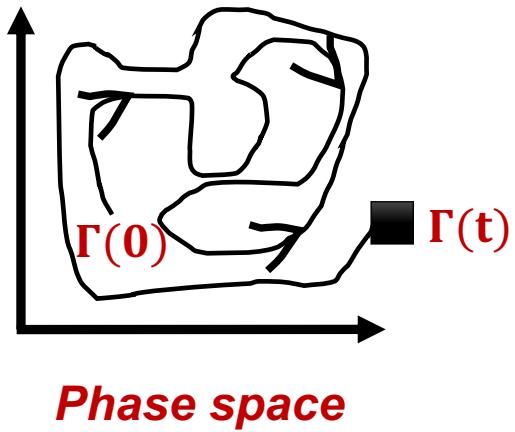
Do we have sufficient information to calculate total energy of the system at this particular point ($\Gamma(t')$) in the phase space?

Let A be any physical property (an observable), then a time average can be calculated as follows

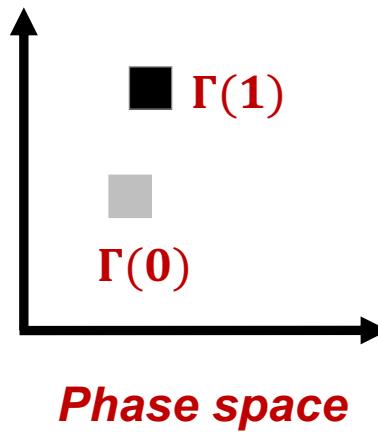
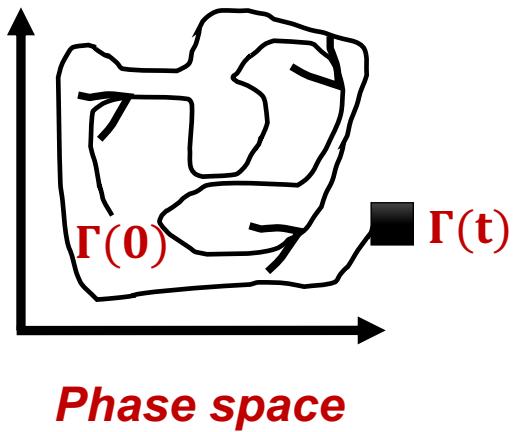
$$A_{obs} = \langle A(\Gamma(t)) \rangle_{time} = \frac{1}{\tau_{obs}} \sum_{\tau=1}^{t_{obs}} A(\Gamma(t))$$

where, τ_{obs} is the total number of timesteps in a molecular dynamics simulation

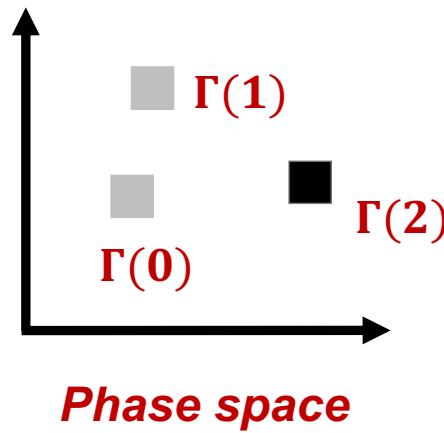
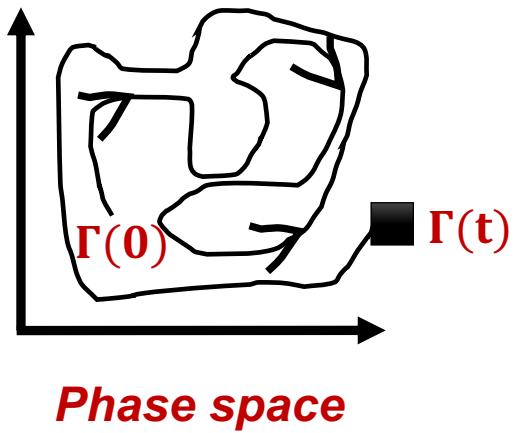
Time evolution vs stochastic sampling



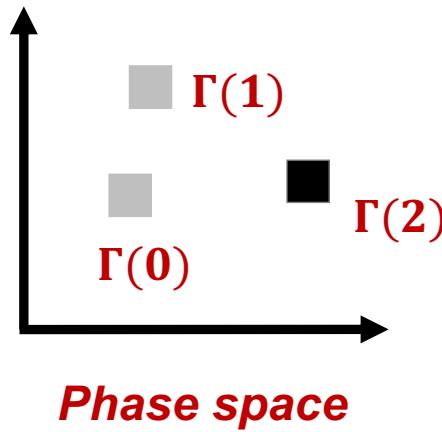
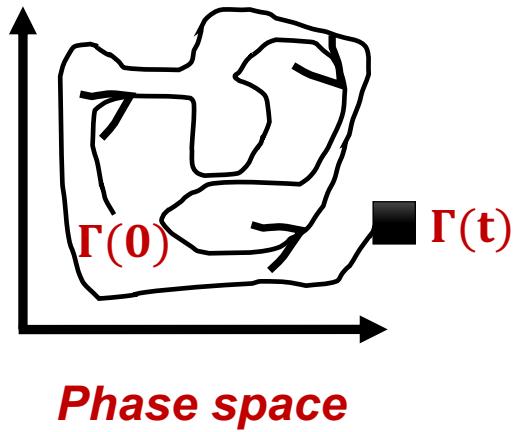
Time evolution vs stochastic sampling



Time evolution vs stochastic sampling



Time evolution vs stochastic sampling



$$A_{obs} = \langle A(\Gamma(t)) \rangle_{time} = \frac{1}{\tau_{obs}} \sum_{t=1}^{t_{obs}} A(\Gamma(t)) \quad A_{obs} = \langle A(\Gamma) \rangle_{ens} = \sum_{\Gamma} A(\Gamma) \rho_{ens}(\Gamma)$$

Gibbs ensemble postulate:
time average = ensemble average

Concept of ensemble

Concept of ensemble

“An ensemble is a collection of a very large number of systems, each constructed to be a replica on a thermodynamic (macroscopic) level of the particular thermodynamic system of interest”

-McQuarrie-

Partition function

The partition function is the central quantity in statistical mechanics. It is similar in importance as the wave function is to quantum mechanics or Gibbs function to thermodynamics.

Partition function

If we know partition function for a particular system, we have all the information necessary to determine thermodynamic properties of the system.

The microcanonical ensemble (NVE)

For microcanonical ensemble, the number of particles (N), volume (V), and total energy (E) is fixed.

Partition function is simply degeneracy of the system, $\Omega(N, V, E)$

$$S = k \ln(\Omega(N, V, E))$$

The canonical ensemble (NVT)

For canonical ensemble, the number of particles (N), volume (V), and temperature (T) is fixed.

$$\begin{aligned} Q(N, V, T) &= \sum_E \Omega(N, V, E) e^{-E(N, V)/kT} \\ S &= k \ln Q + kT \left(\frac{\partial \ln Q}{\partial T} \right)_{N, V} \\ p &= kT \left(\frac{\partial \ln Q}{\partial V} \right)_{N, T} \\ \mu &= -kT \left(\frac{\partial \ln Q}{\partial N} \right)_{V, T} \\ E &= kT^2 \left(\frac{\partial \ln Q}{\partial T} \right)_{V, N} \end{aligned}$$

The grand canonical ensemble (μVT)

For grand canonical ensemble, the chemical potential (μ), volume (V), and temperature (T) is fixed.

$$\Xi(\mu, V, T) = \sum_N Q(N, V, T) e^{-\mu N / kT}$$

The grand canonical ensemble (μVT)

For grand canonical ensemble, the chemical potential (μ), volume (V), and temperature (T) is fixed.

$$\Xi(\mu, V, T) = \sum_N Q(N, V, T) e^{\mu N / kT}$$

The isothermal isobaric ensemble (NpT)

For isothermal isobaric ensemble, the number of particles (N), pressure (p), and temperature (T) is fixed.

$$\Delta(N, p, T) = \sum_V Q(N, V, T) e^{-pV/kT}$$

Fluctuations

Fluctuations are of fundamental importance in statistical mechanics.

Fluctuations

In the canonical ensemble, fluctuations in energy are related to heat capacity, while in the grand canonical ensemble, fluctuations in number of particles is related to the compressibility.

Canonical

$$\sigma_E^2 = kT^2 C_v$$

$$\frac{\sigma_E}{\bar{E}} = \frac{(kT^2 C_v)^{\frac{1}{2}}}{\bar{E}} \sim O\left(\frac{1}{\sqrt{N}}\right)$$

Grand canonical

$$\sigma_N^2 = \frac{\bar{N}kT \kappa}{V}$$

$$\frac{\sigma_N}{\bar{N}} \sim O\left(\frac{1}{\sqrt{N}}\right)$$

Fluctuations

In the canonical ensemble, fluctuations in energy are related to heat capacity, while in the grand canonical ensemble, fluctuations in number of particles is related to the compressibility.

Canonical

$$\sigma_E^2 = kT^2 C_v$$

$$\frac{\sigma_E}{\bar{E}} = \frac{(kT^2 C_v)^{\frac{1}{2}}}{\bar{E}} \sim O\left(\frac{1}{\sqrt{N}}\right)$$

Grand canonical

$$\sigma_N^2 = \frac{\bar{N}kT \kappa}{V}$$

$$\frac{\sigma_N}{\bar{N}} \sim O\left(\frac{1}{\sqrt{N}}\right)$$

How fast fluctuations decay leads to transport properties such as diffusion and viscosity!

Fluctuations

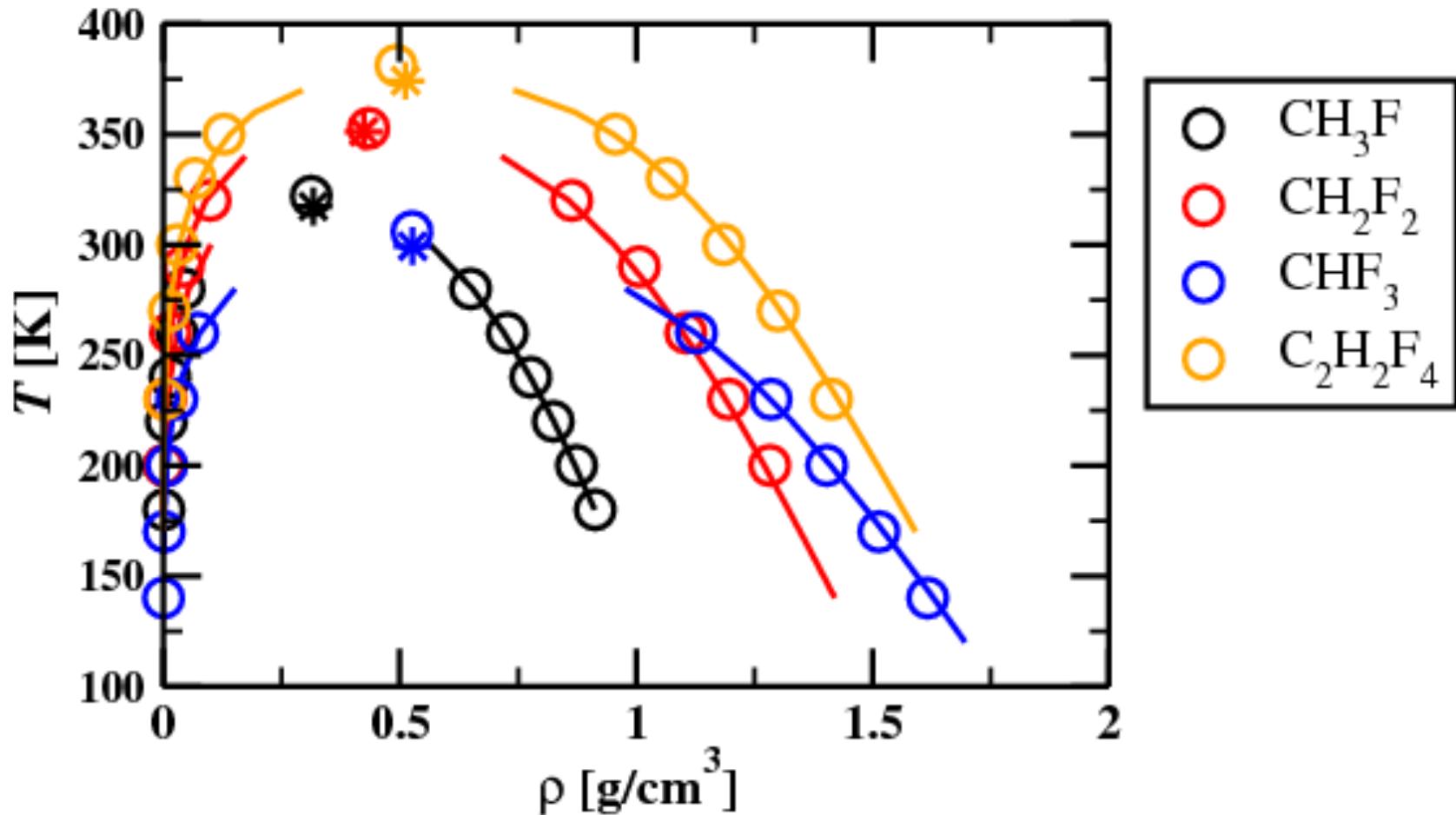
Another important aspect of fluctuations is the ensemble equivalence

$$Q(N, V, T) = \sum_E \Omega(N, V, E) e^{-E/kT}$$

$$Q(N, V, T) = \Omega(N, V, \bar{E}) e^{-\bar{E}/kT}$$

Phase equilibria

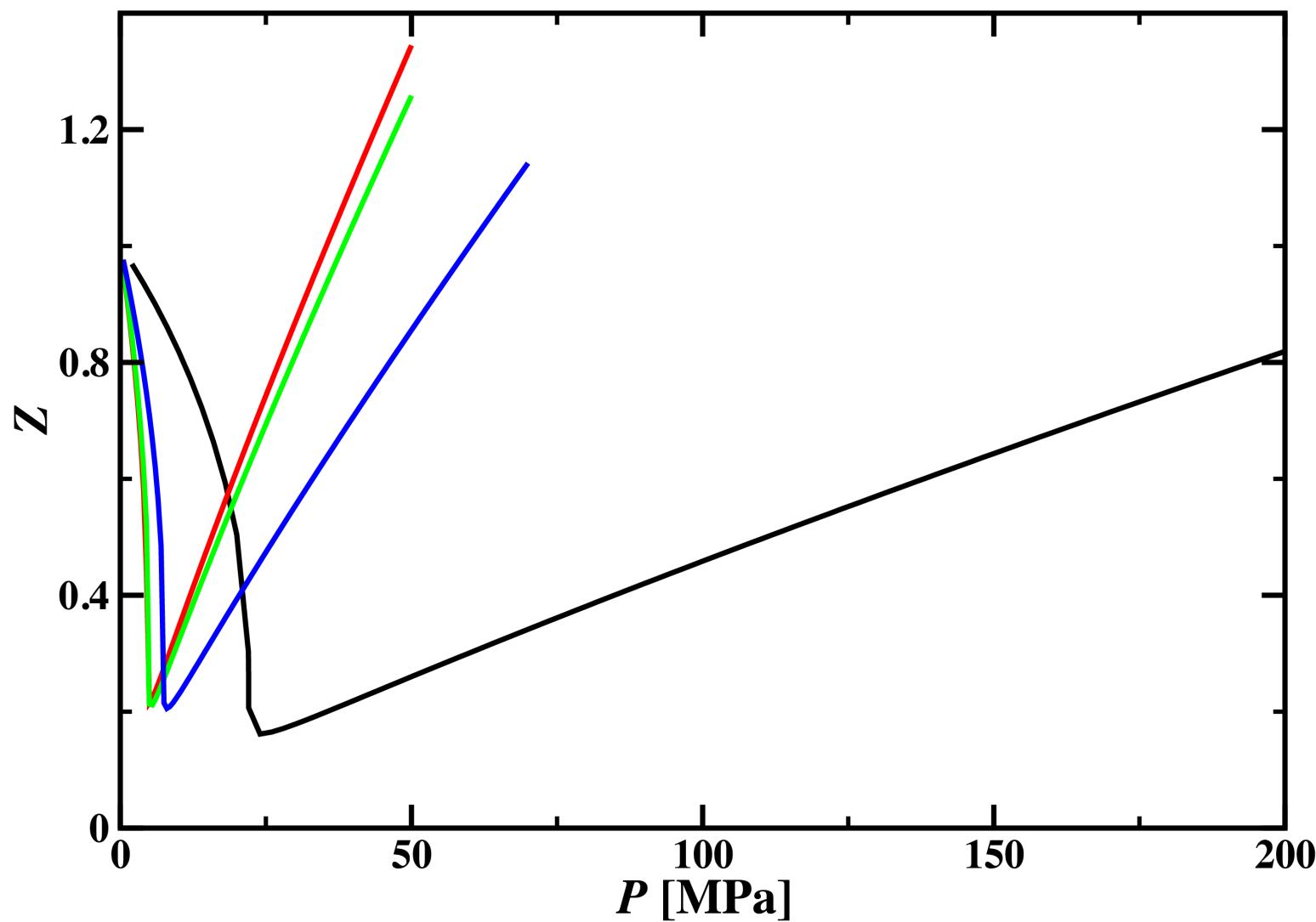
Vapor Liquid Coexistence Curve



Compressibility Factor

$$Z = PV/RT$$

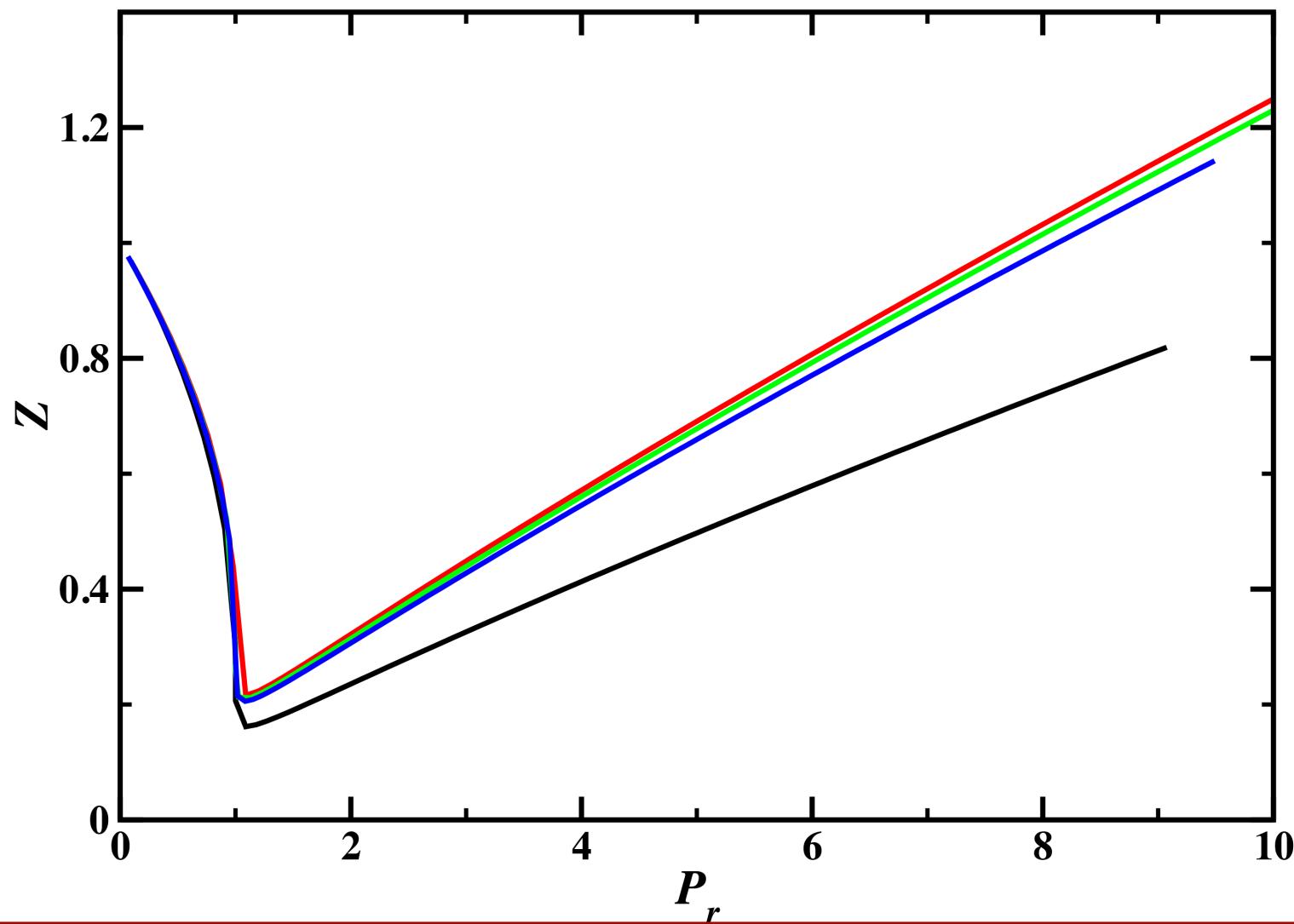
Compressibility Factor

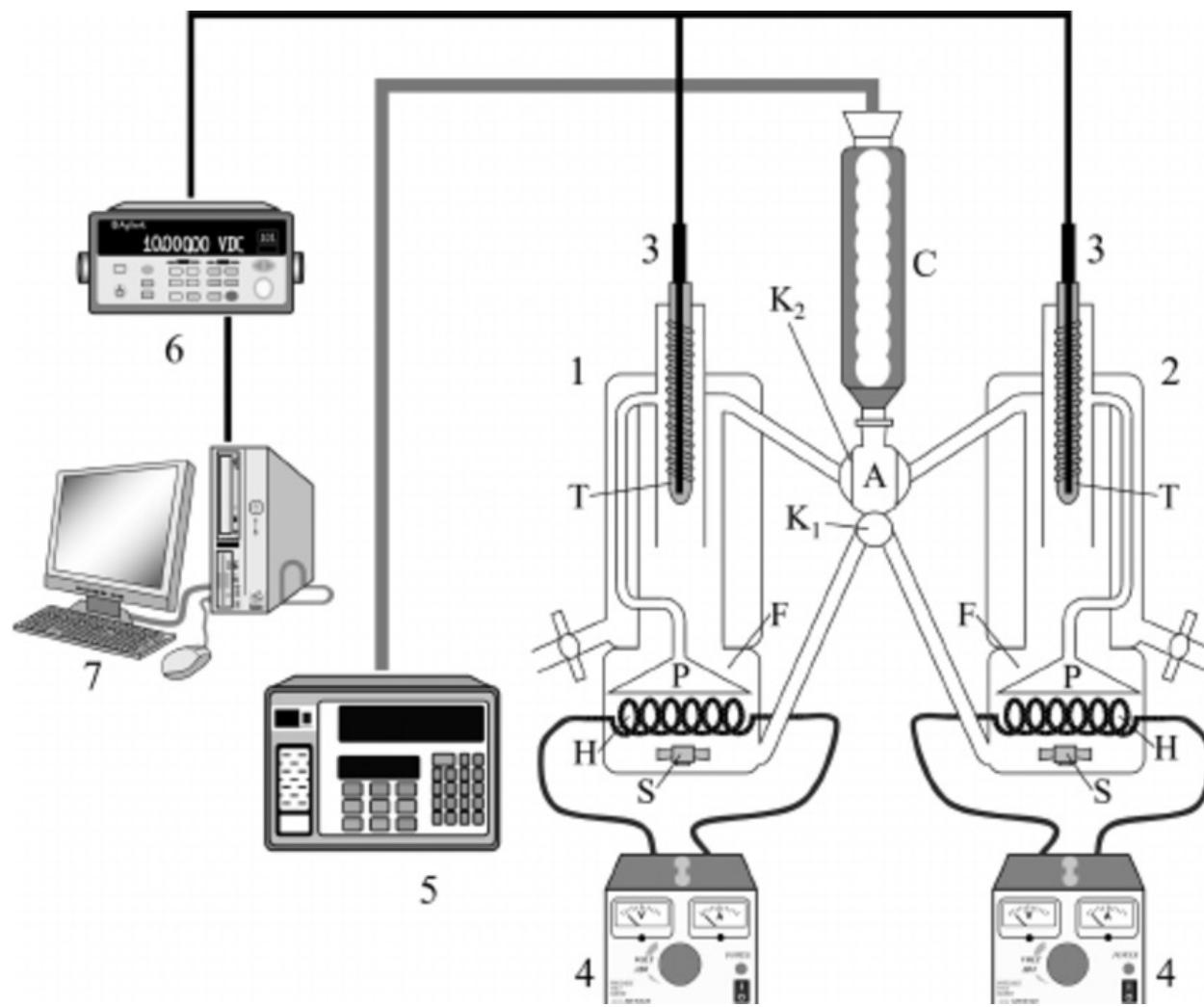


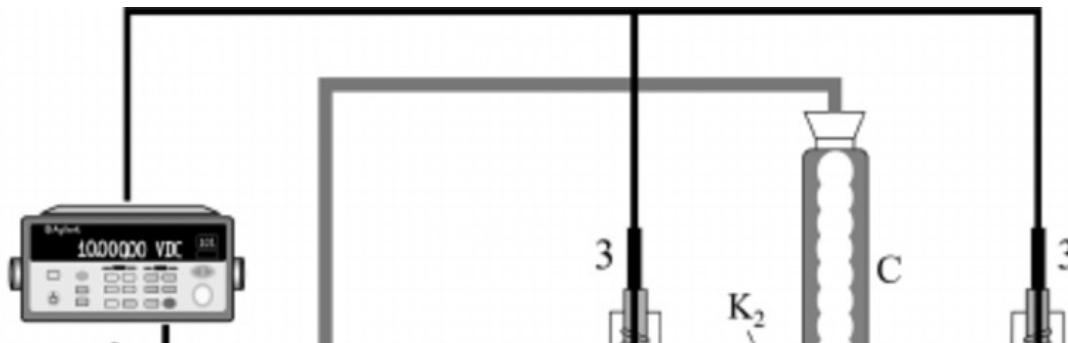
Reduced Pressure

$$P_r = P/P_C$$

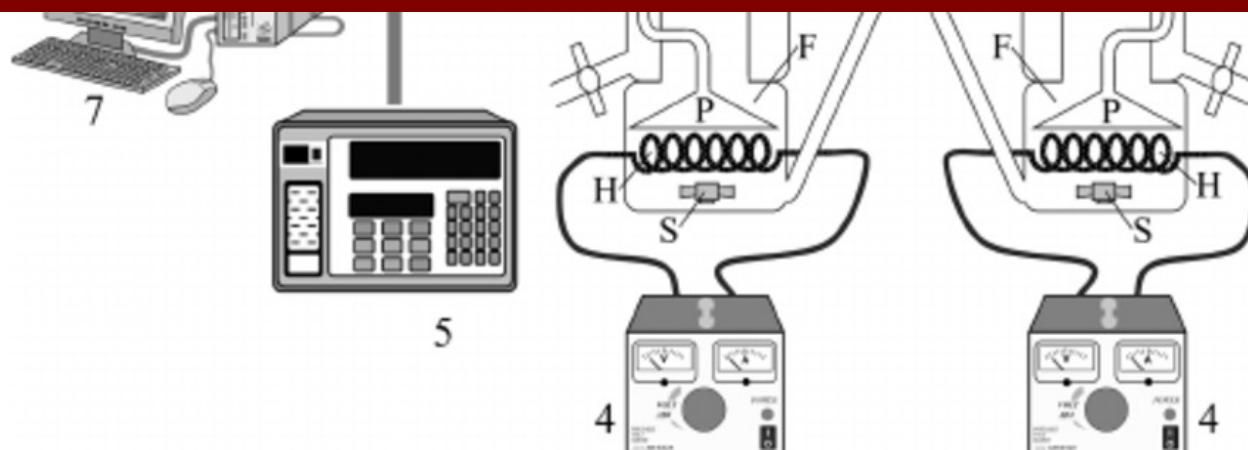
Compressibility Factor







Long equilibration time renders molecular dynamics techniques unsuitable for modeling vapor liquid equilibria



Metropolis Monte Carlo

Metropolis Monte Carlo

Gibbs Ensemble Monte Carlo (GEMC)

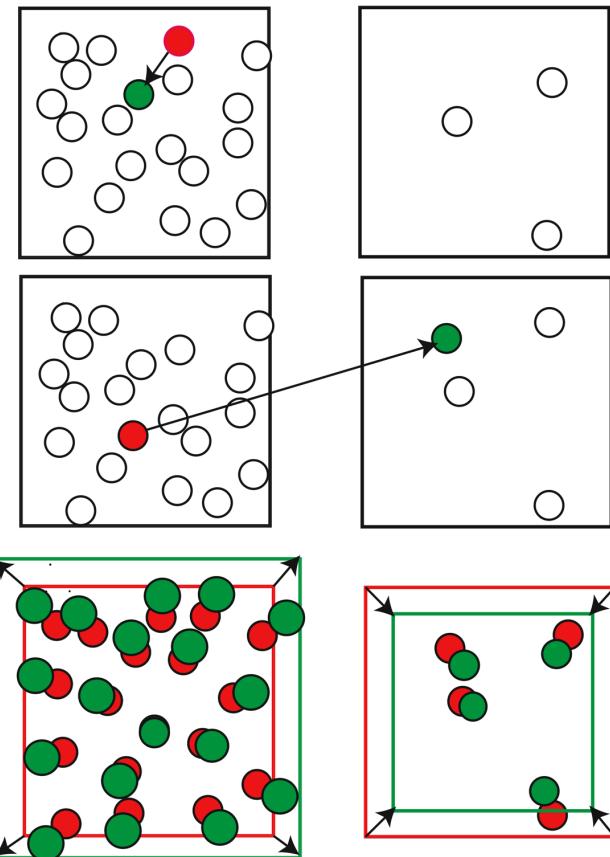
A. Z. Panagiotopoulos *Mol. Phys.*, 61, 813 (1987)

Main advantage:

Direct simulation of coexisting phases without a physical interface using a unified partition function with two separate sub-regions

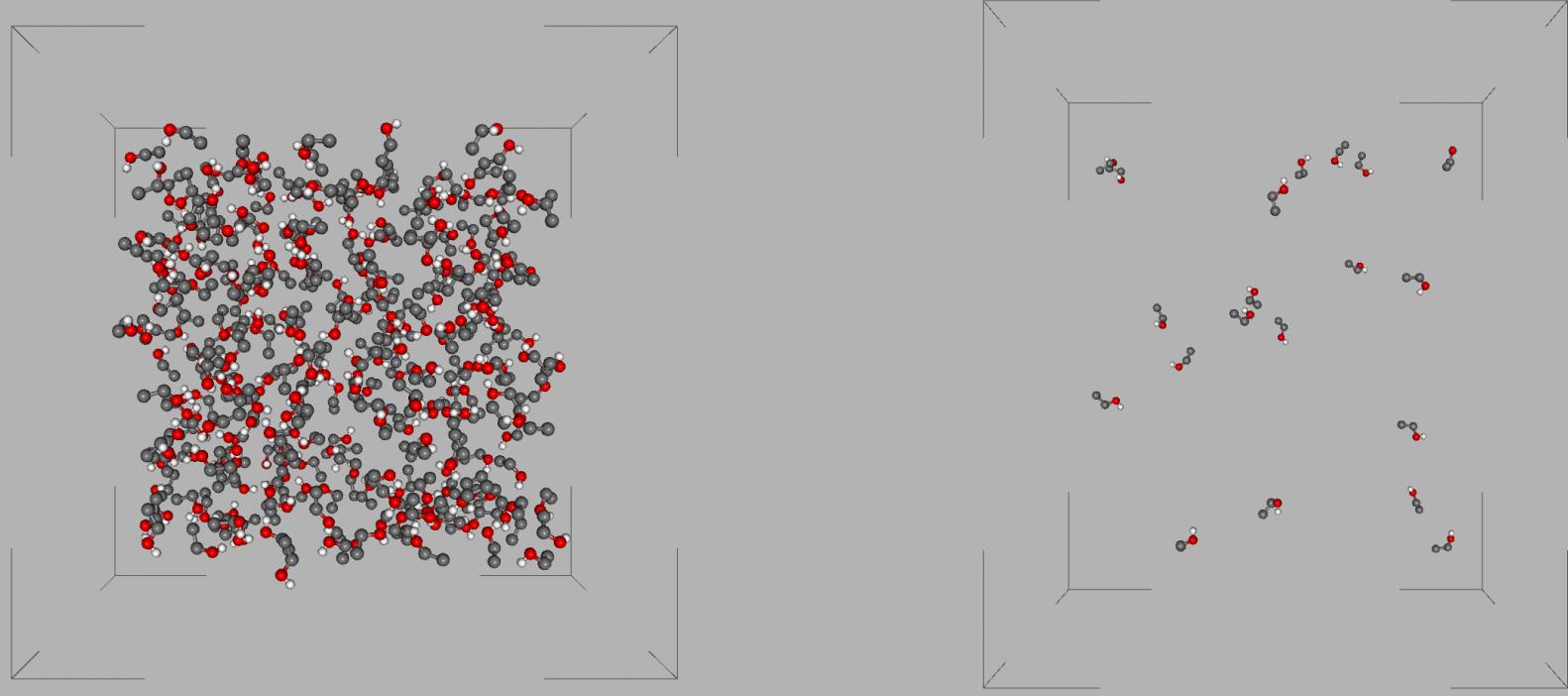
Three types of moves for sampling of classical mechanical phase space:

1. Translations, rotations, and conformational changes → thermal equilibrium
2. Particle swaps → phase equilibrium (chemical potential)
3. Volume exchanges → mechanical equilibrium (pressure)

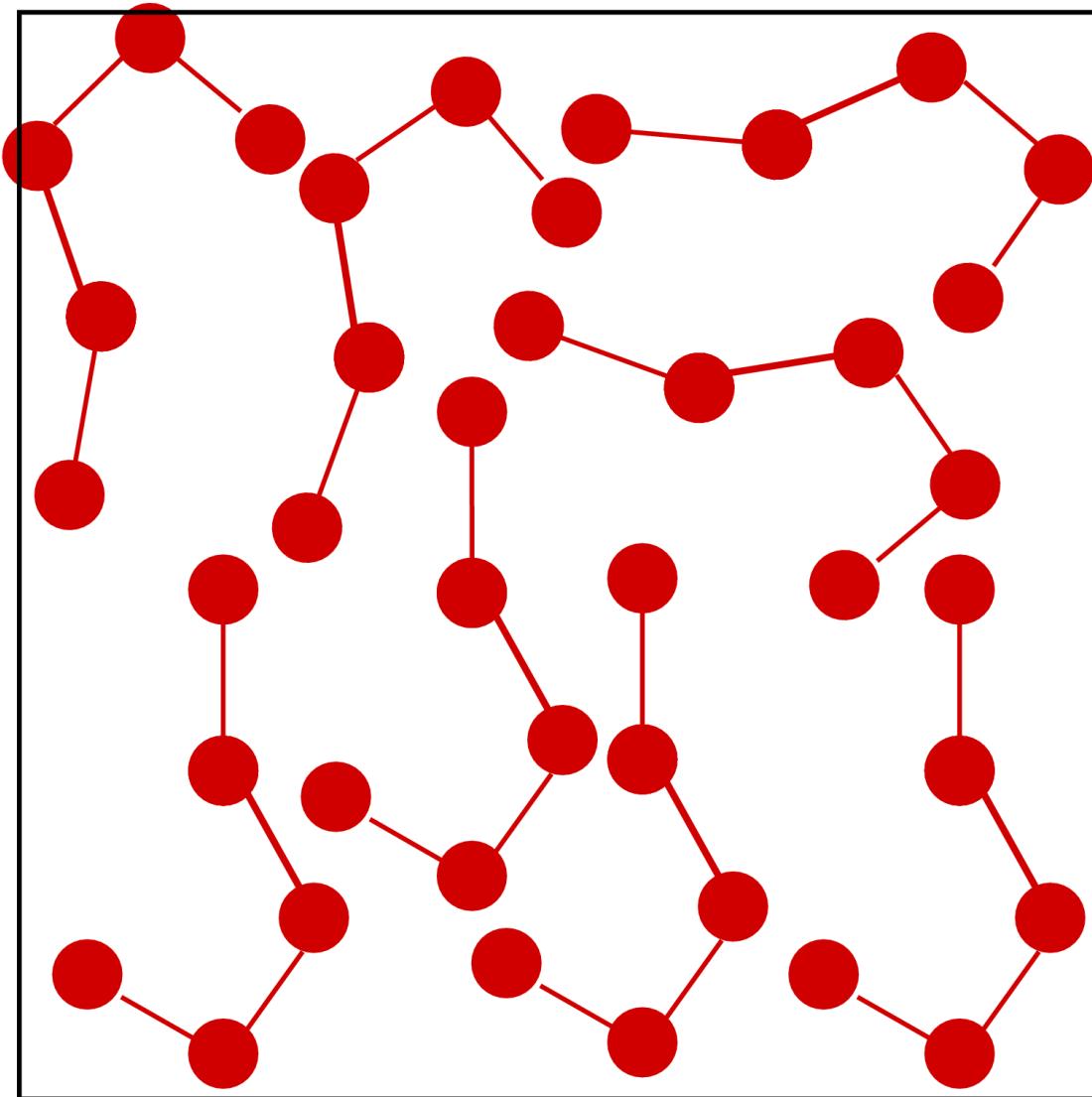


GEMC Simulation of Ethanol at 420 K

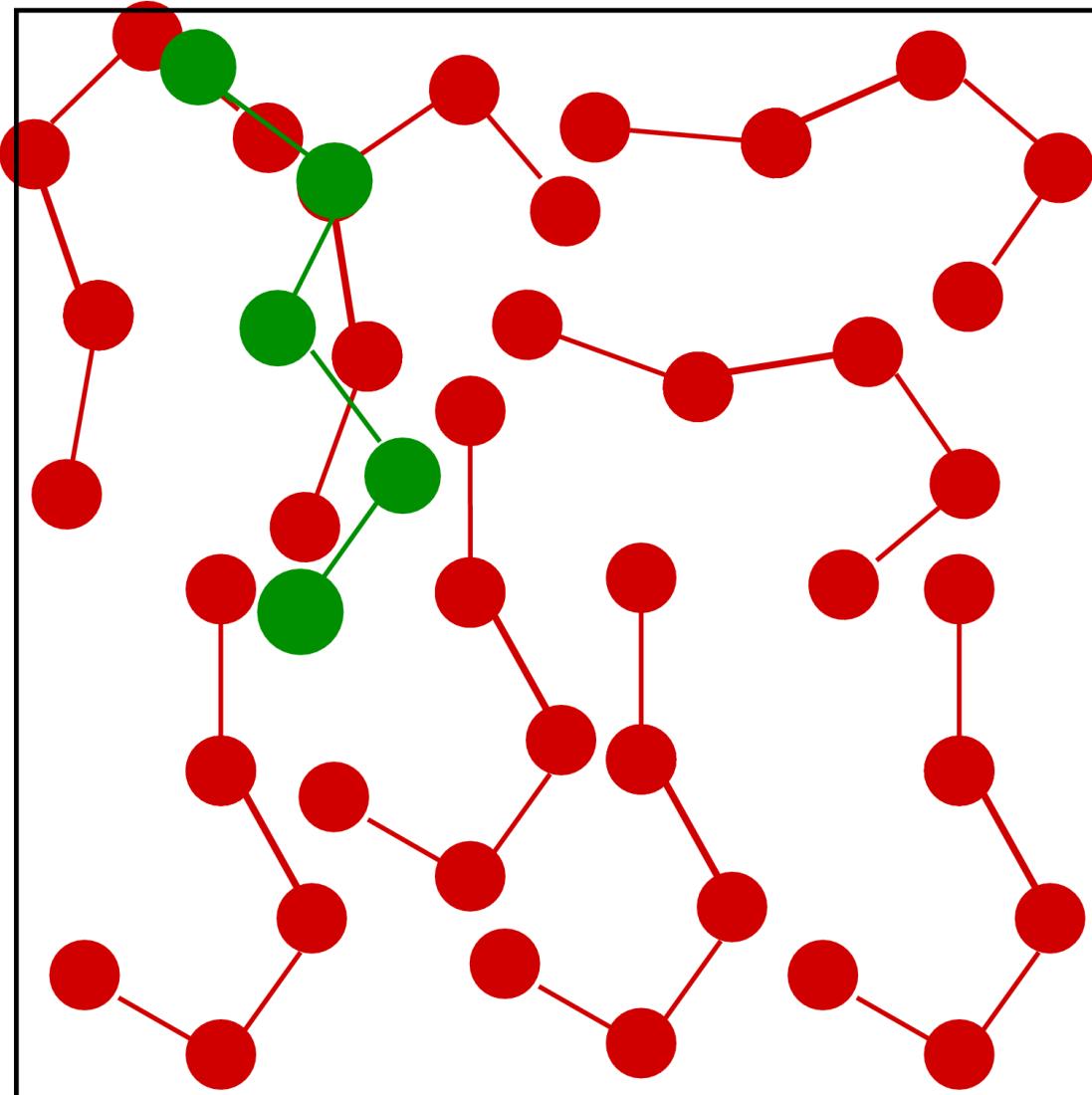
GEMC Simulation of Ethanol at 420 K



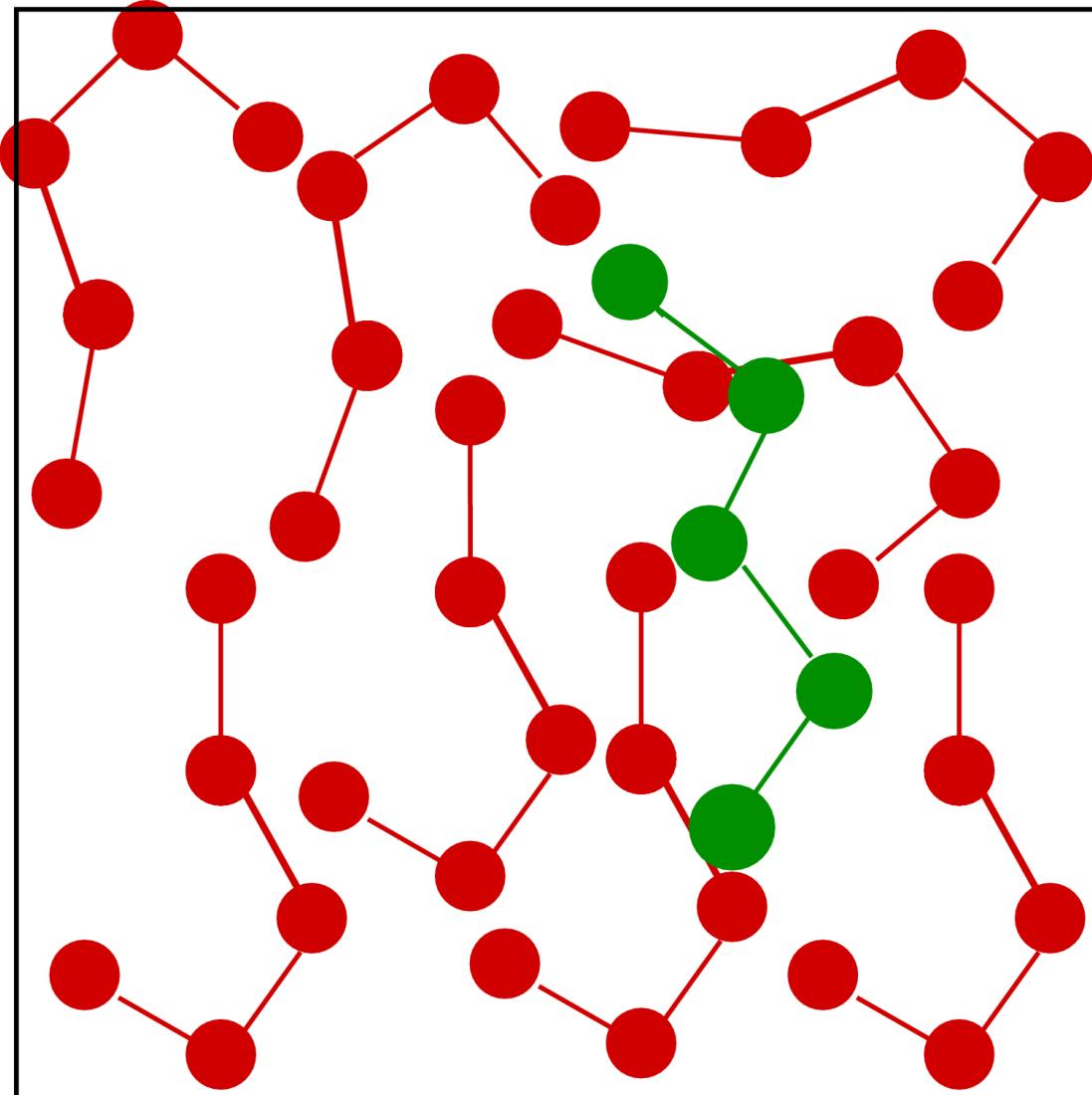
GEMC Bottleneck: Insertion Difficulty



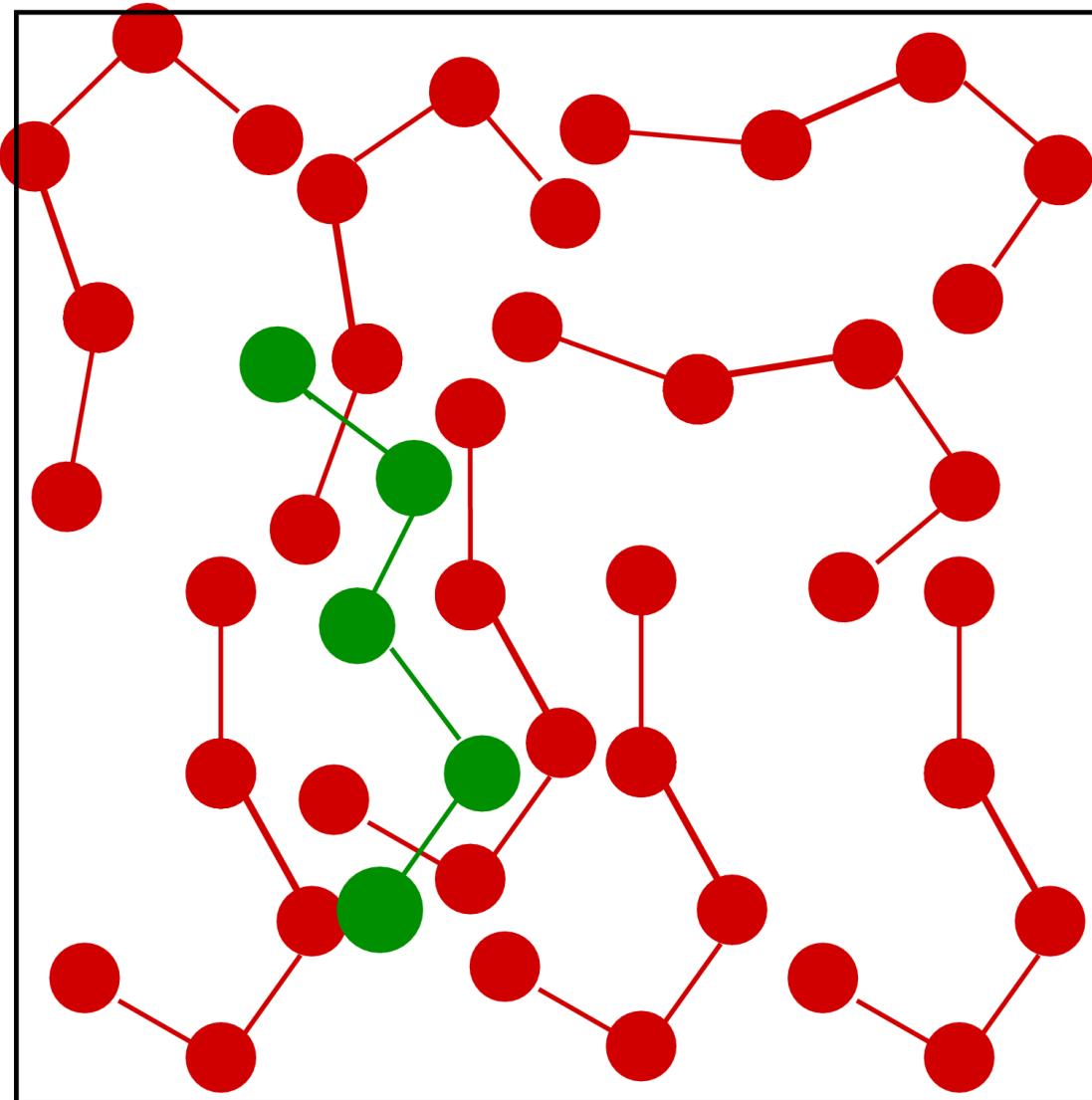
GEMC Bottleneck: Insertion Difficulty



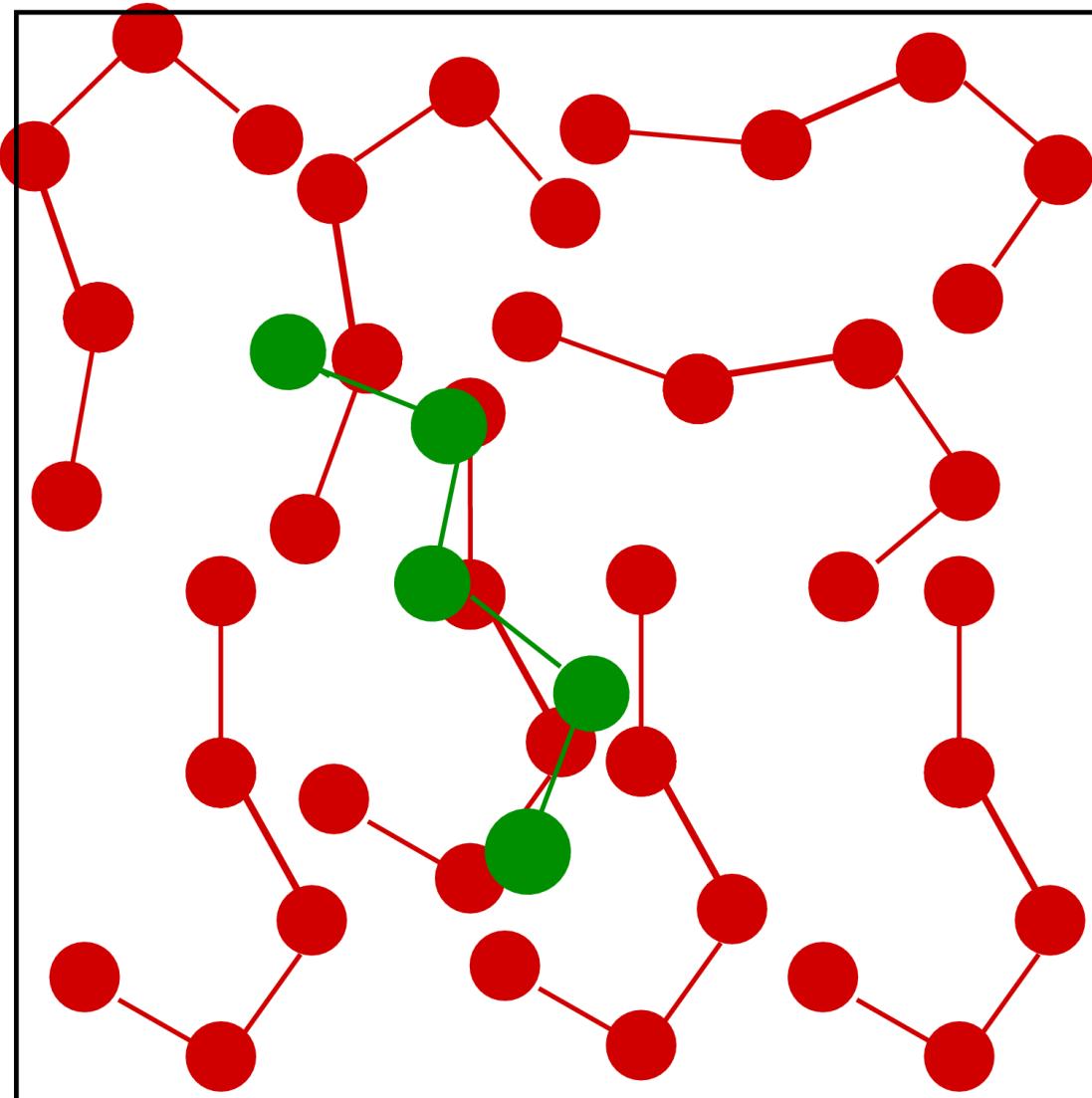
GEMC Bottleneck: Insertion Difficulty



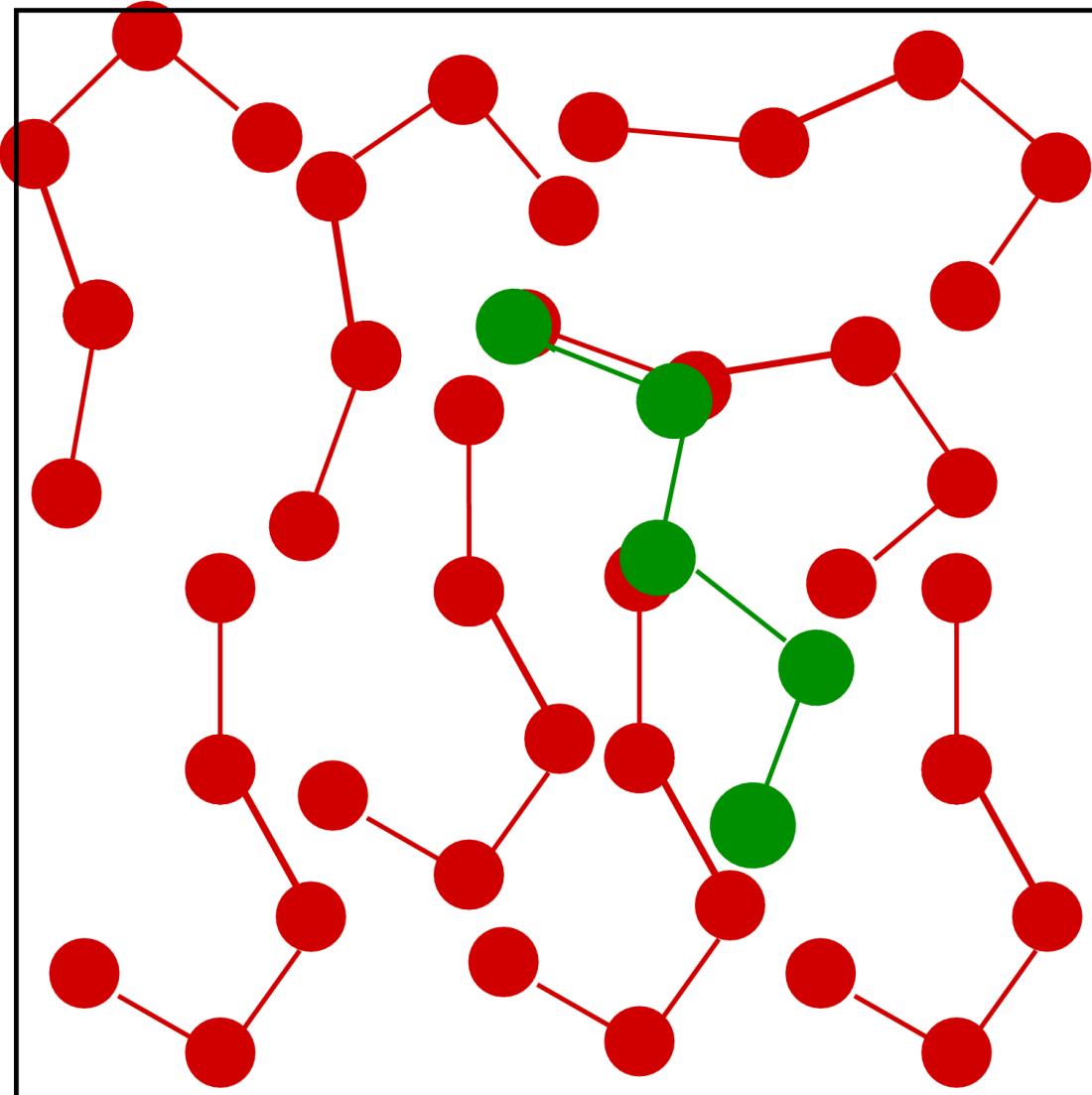
GEMC Bottleneck: Insertion Difficulty



GEMC Bottleneck: Insertion Difficulty



GEMC Bottleneck: Insertion Difficulty



Biased Growth: Rosenbluth Sampling

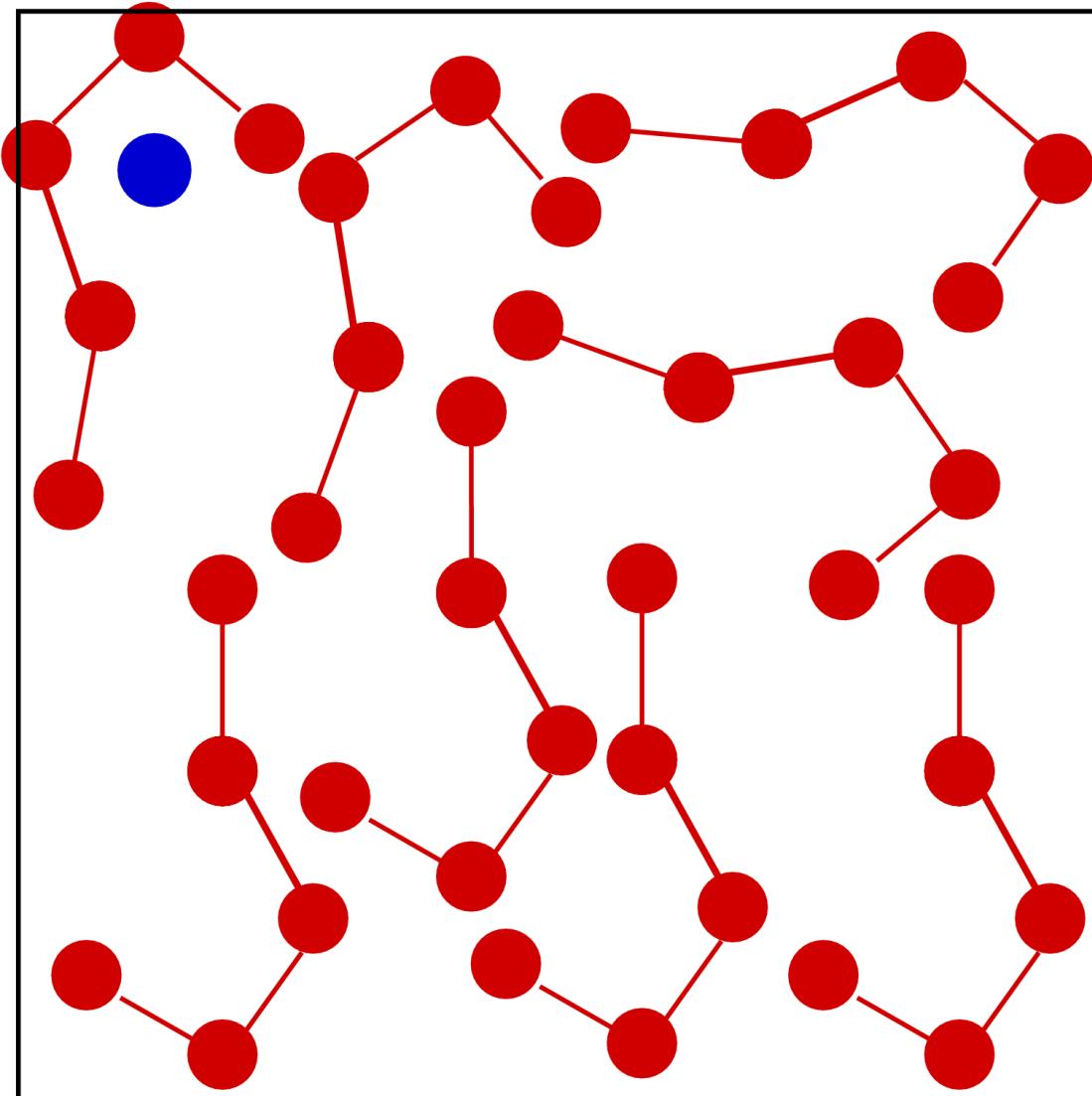


Grow chains by
looking around first!

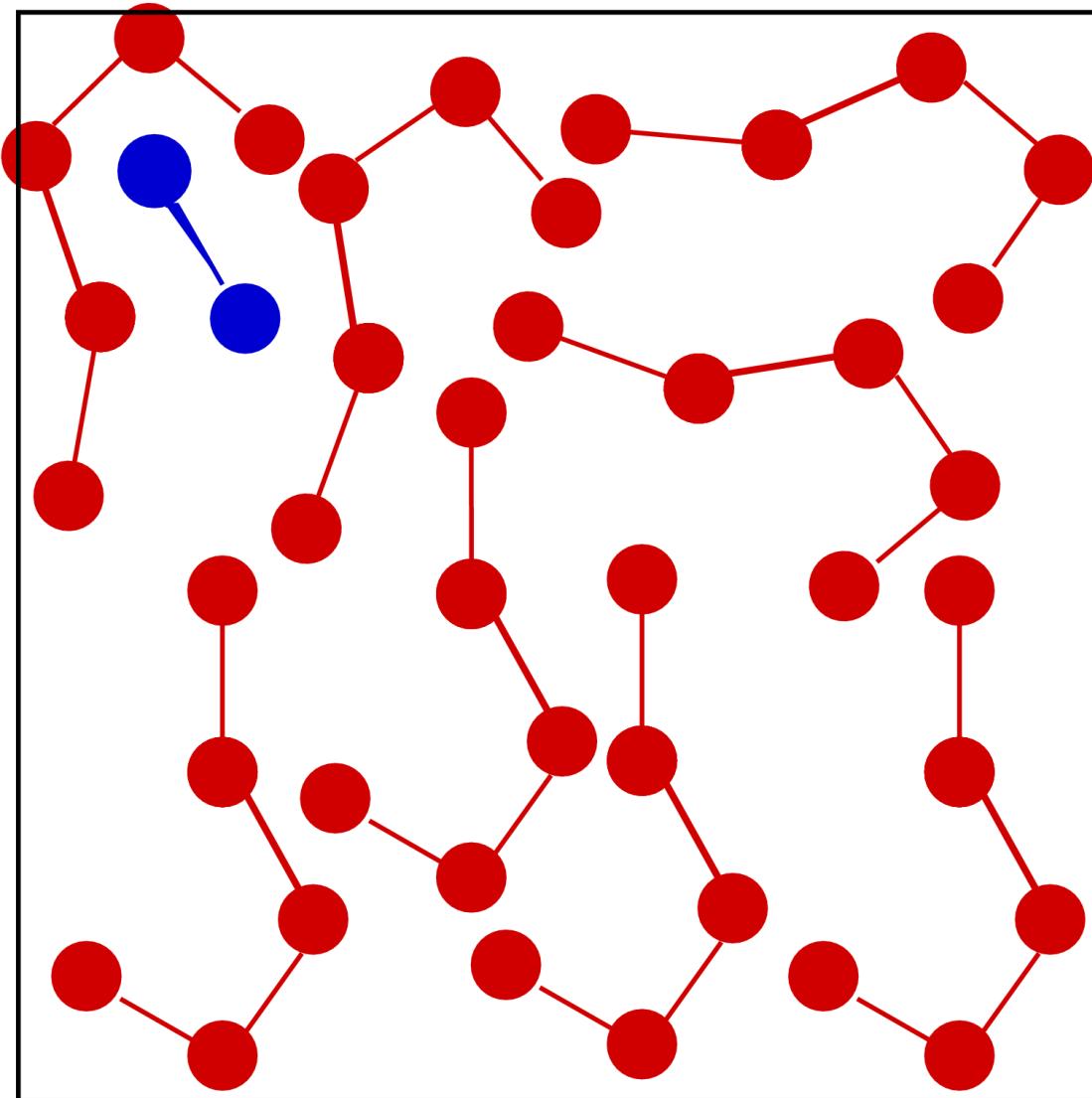
M. N. Rosenbluth and A. W. Rosenbluth *J. Chem Phys.* , **23**, 356 (1955)

Marshall N. Rosenbluth

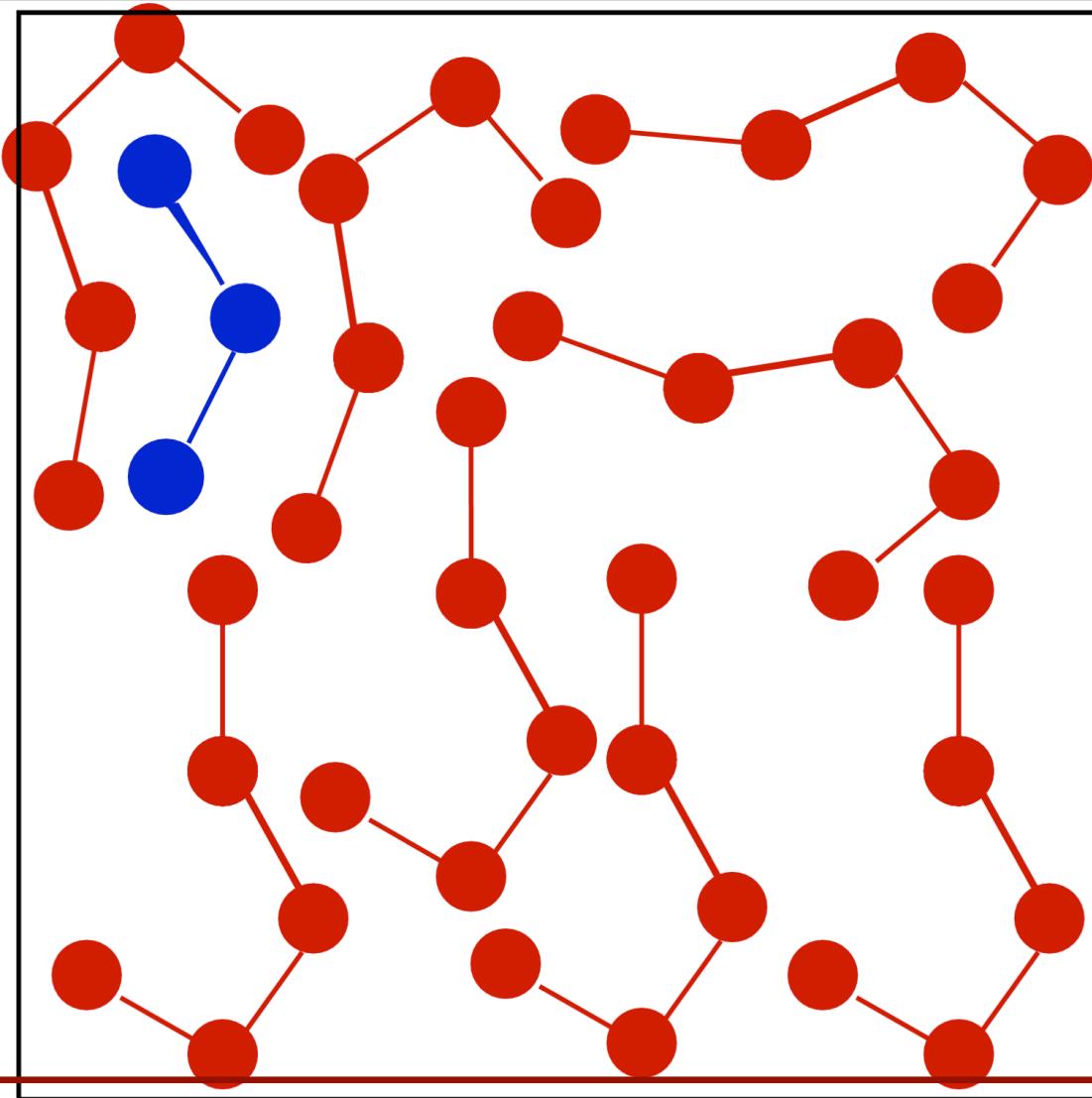
GEMC Bottleneck: Solution



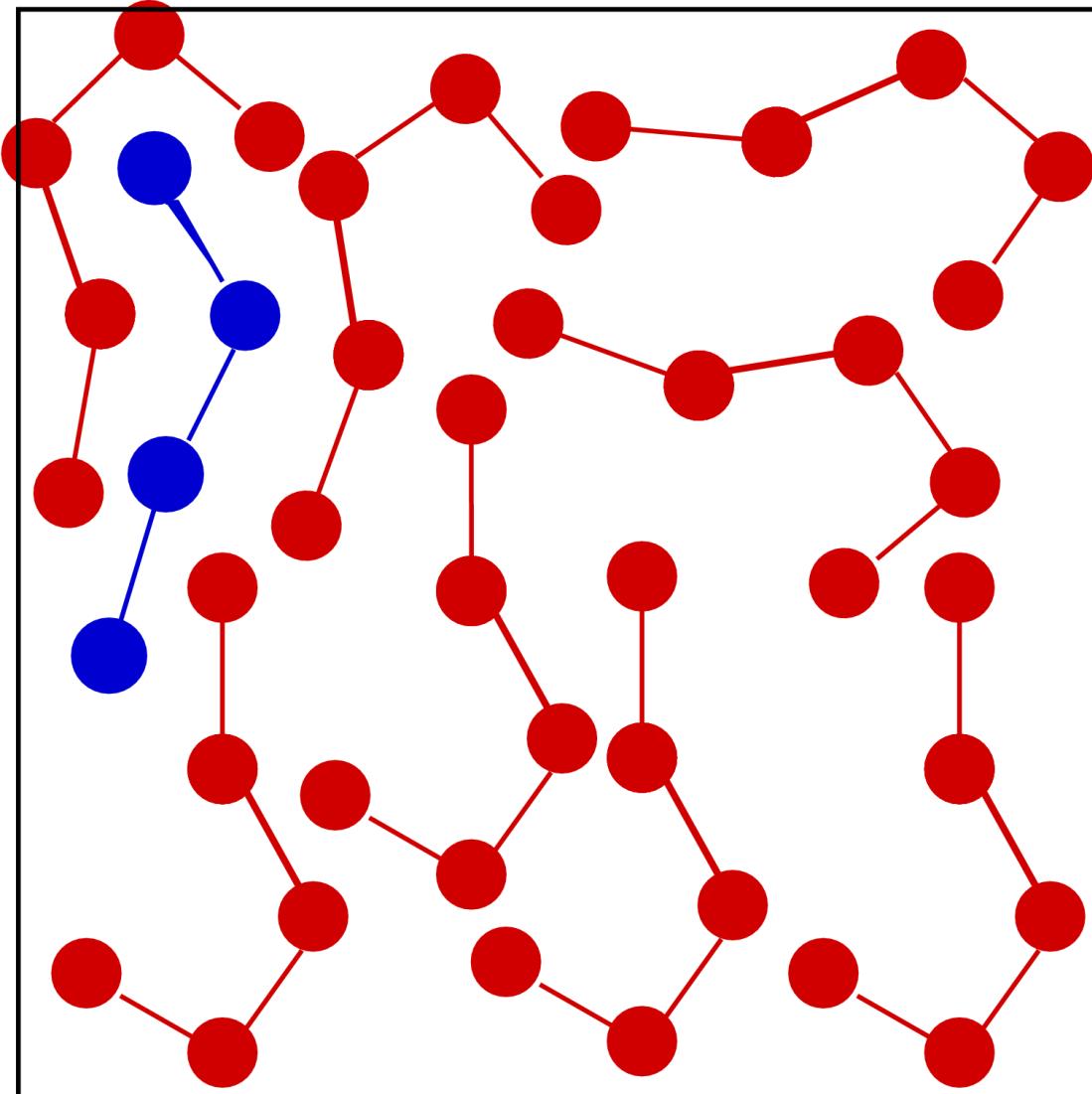
GEMC Bottleneck: Solution



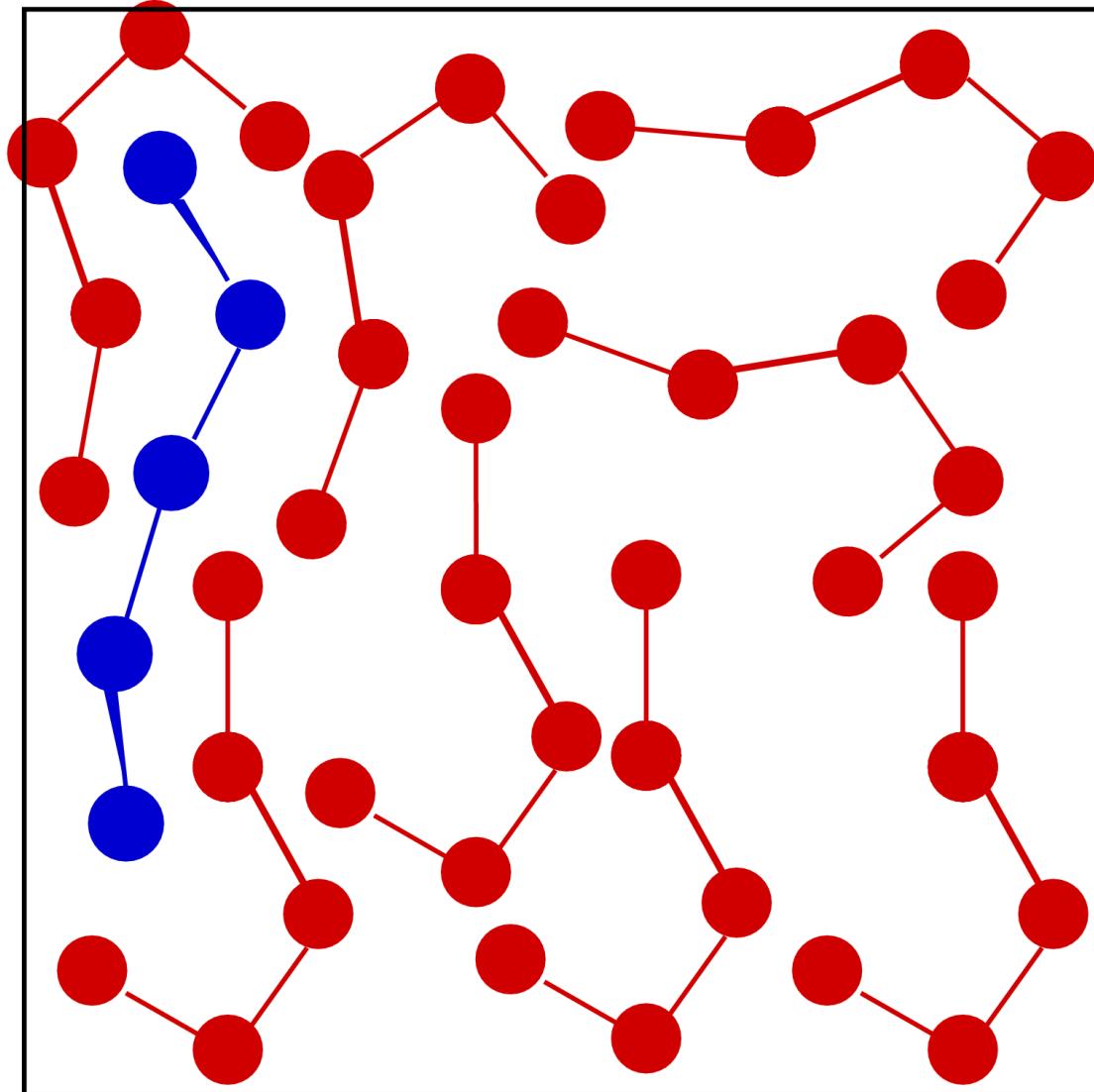
GEMC Bottleneck: Solution



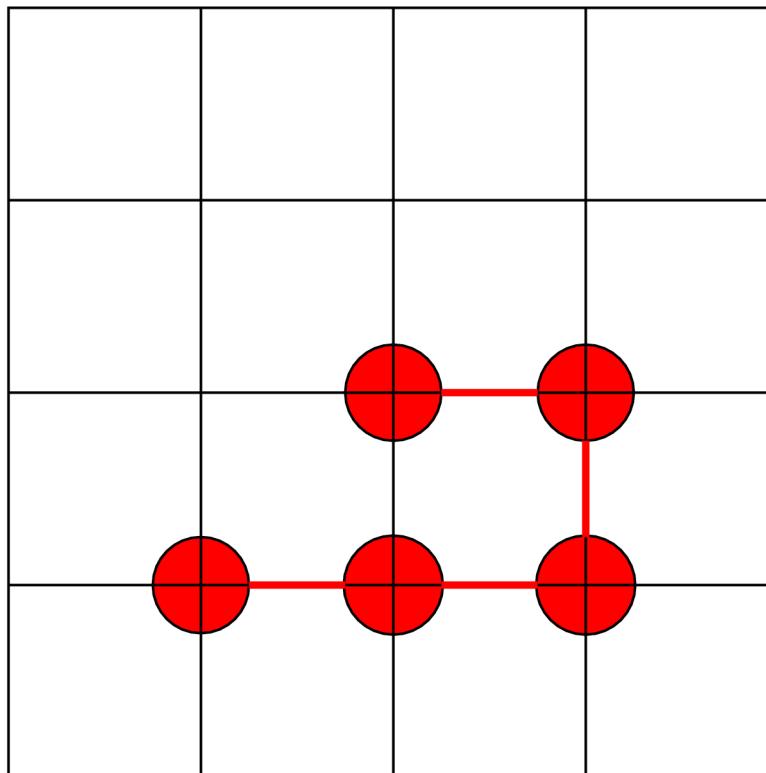
GEMC Bottleneck: Solution



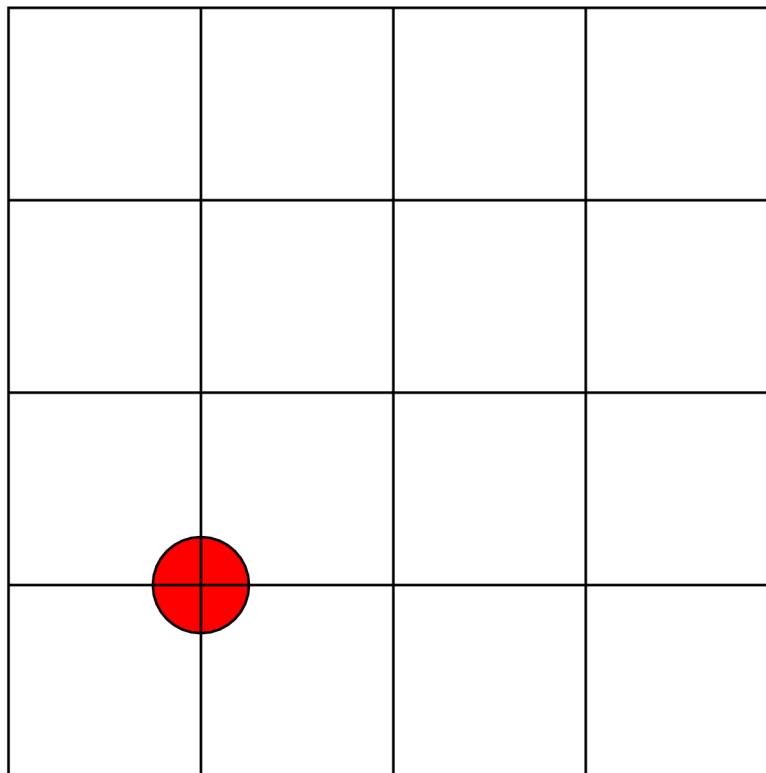
GEMC Bottleneck: Solution



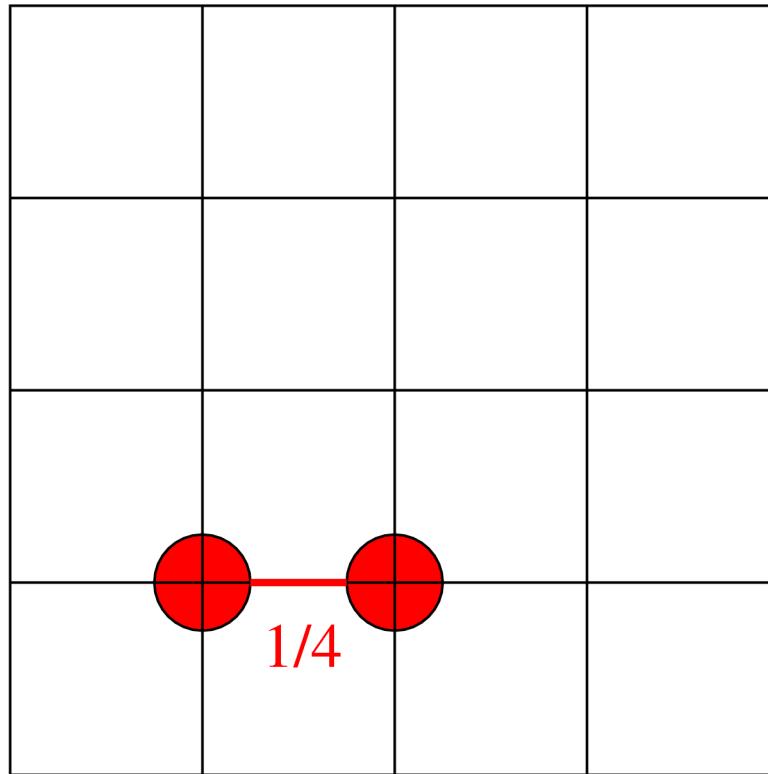
Biased Growth: Rosenbluth Sampling



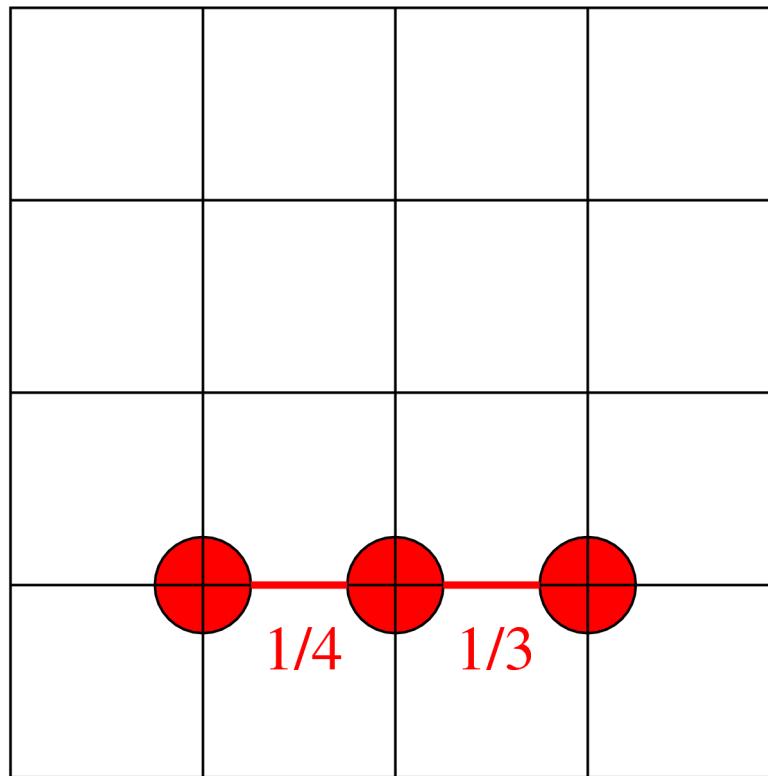
Biased Growth: Rosenbluth Sampling



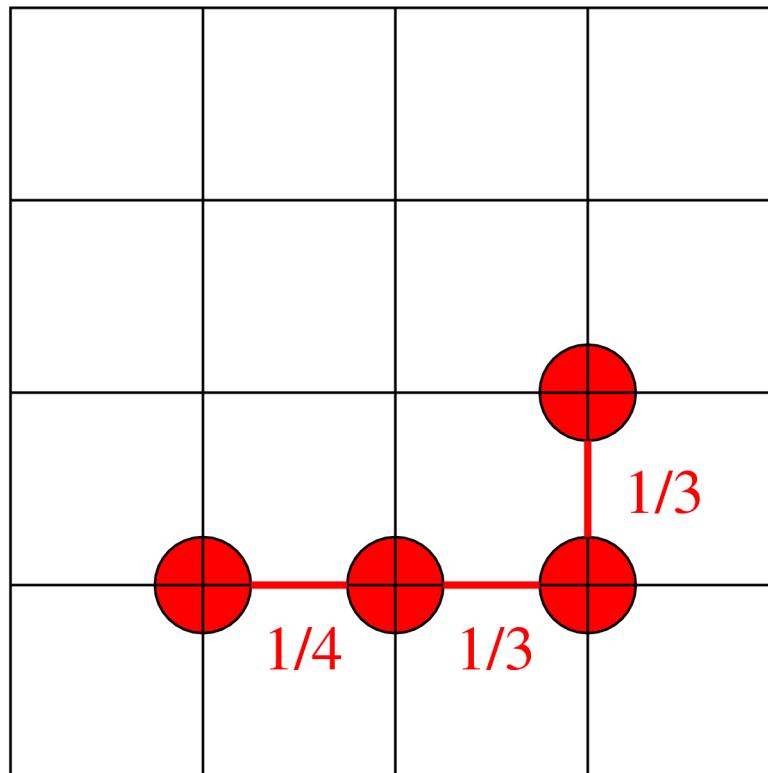
Biased Growth: Rosenbluth Sampling



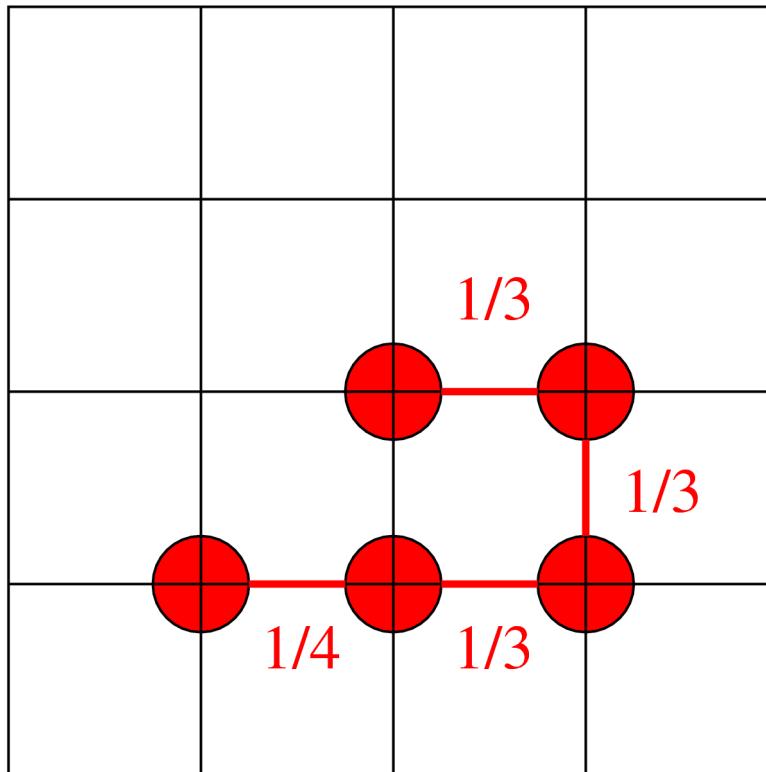
Biased Growth: Rosenbluth Sampling



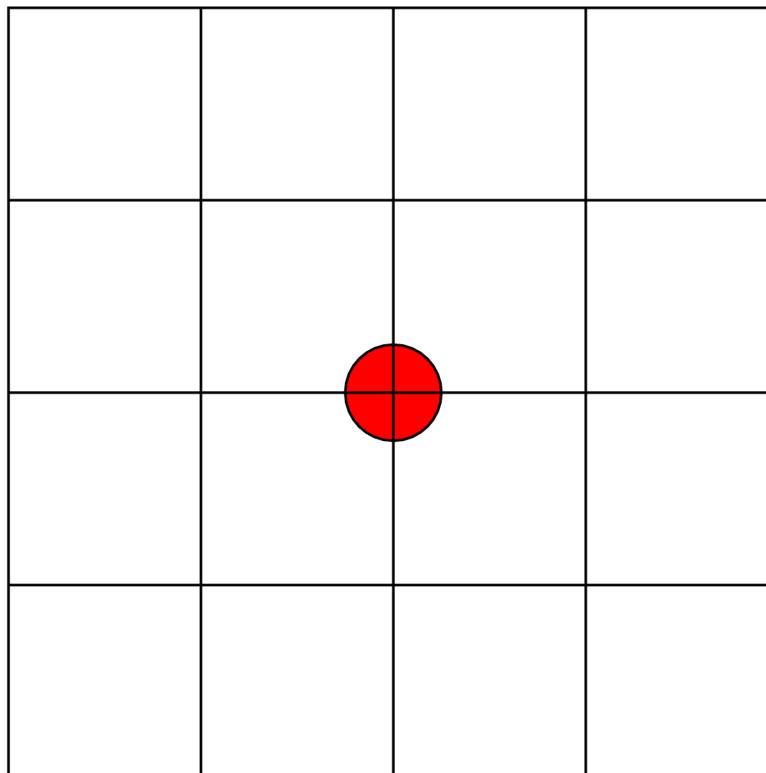
Biased Growth: Rosenbluth Sampling



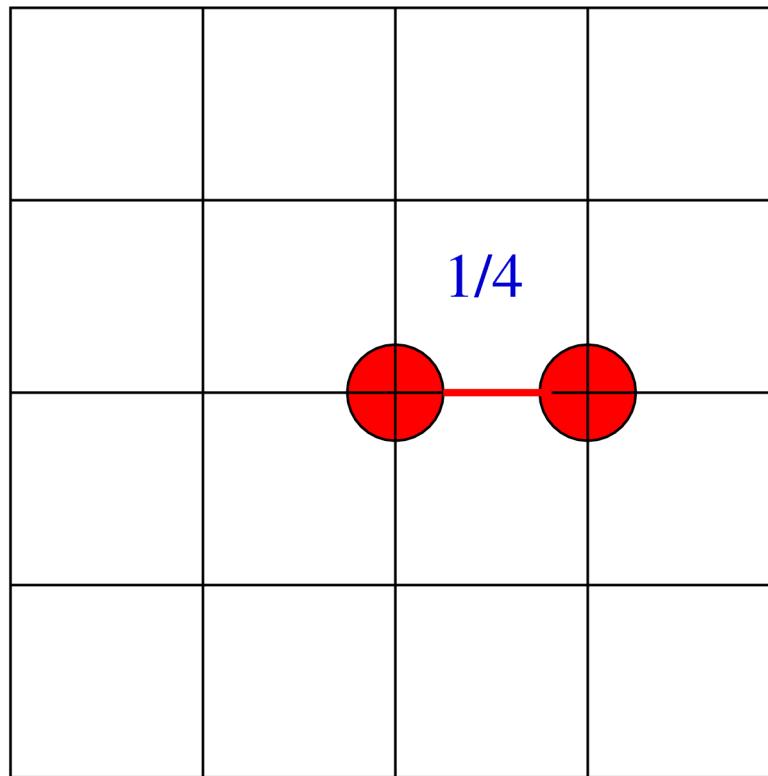
Biased Growth: Rosenbluth Sampling



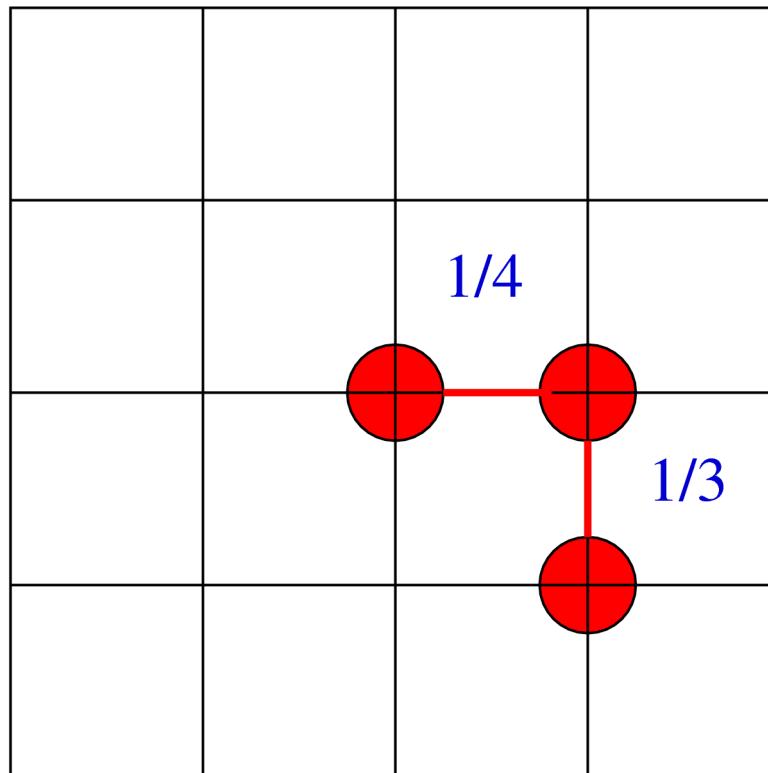
Biased Growth: Rosenbluth Sampling



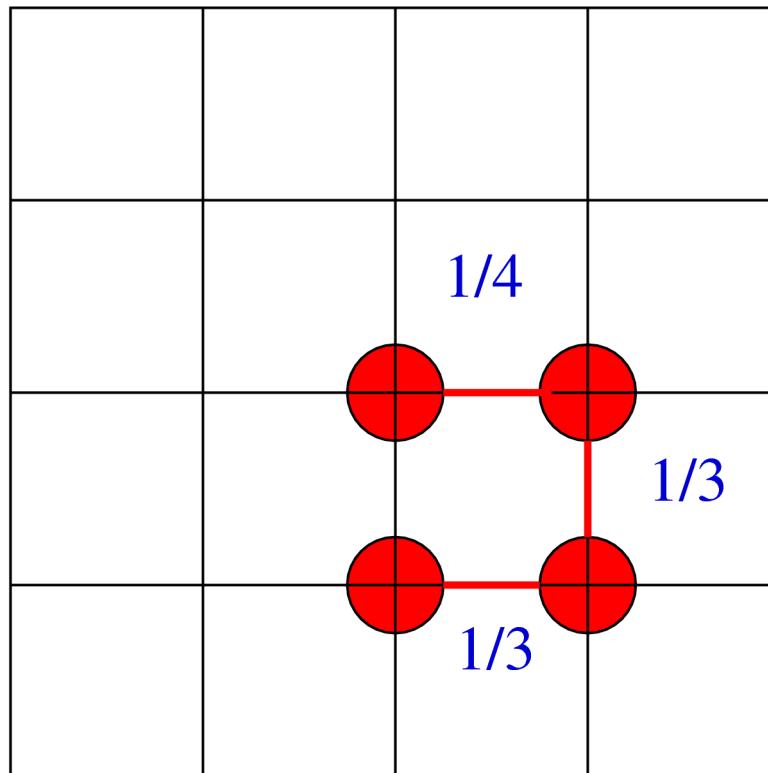
Biased Growth: Rosenbluth Sampling



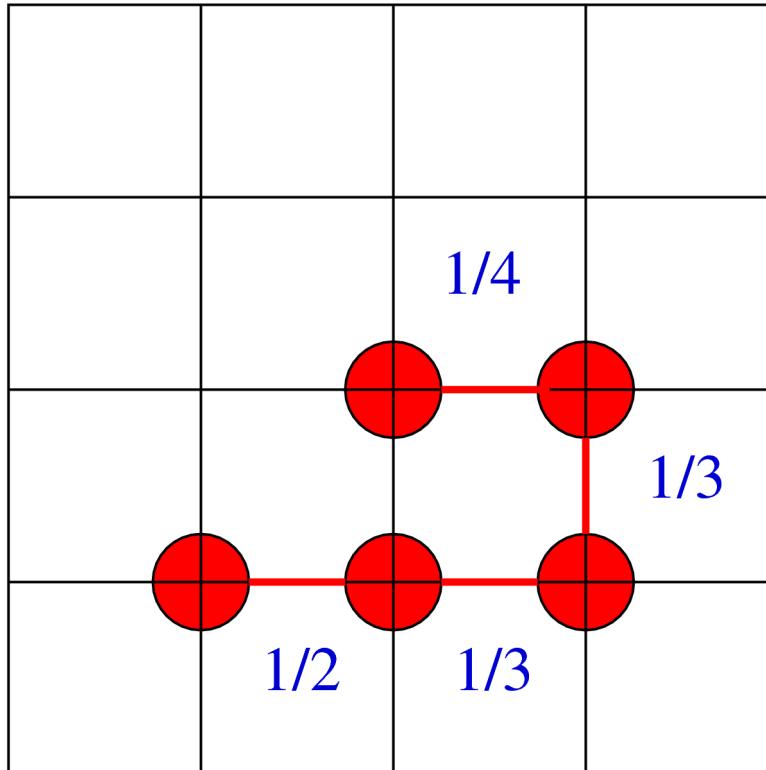
Biased Growth: Rosenbluth Sampling



Biased Growth: Rosenbluth Sampling

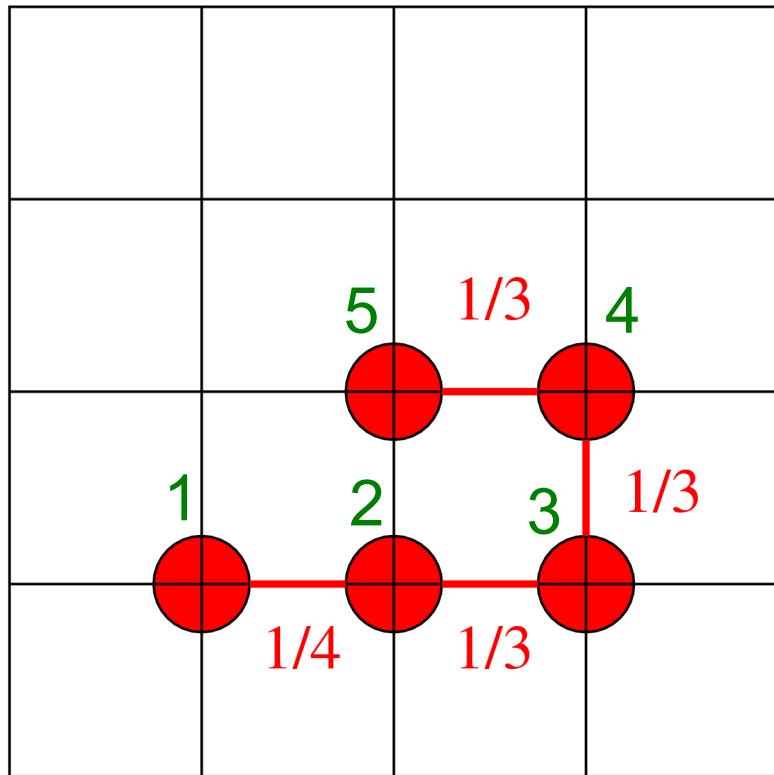


Biased Growth: Rosenbluth Sampling

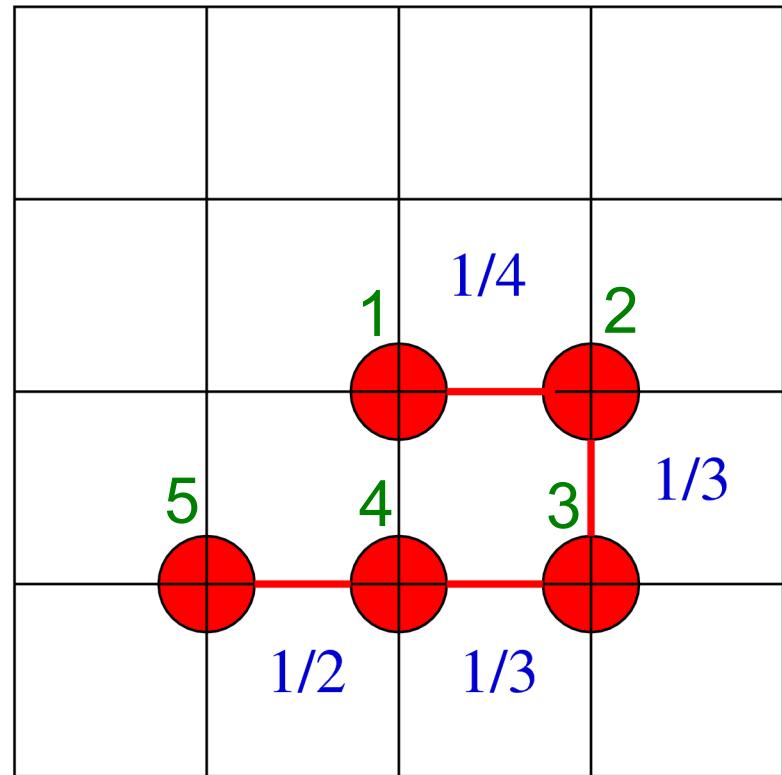


Biased Growth: Rosenbluth Sampling

Forward



Reverse



$$\text{forward : } P_5^{\text{RR}} = 1/4 \times 1/3 \times 1/3 \times 1/3 = 1/108$$

$$\text{reverse : } P_5^{\text{RR}} = 1/4 \times 1/3 \times 1/3 \times 1/2 = 1/72$$

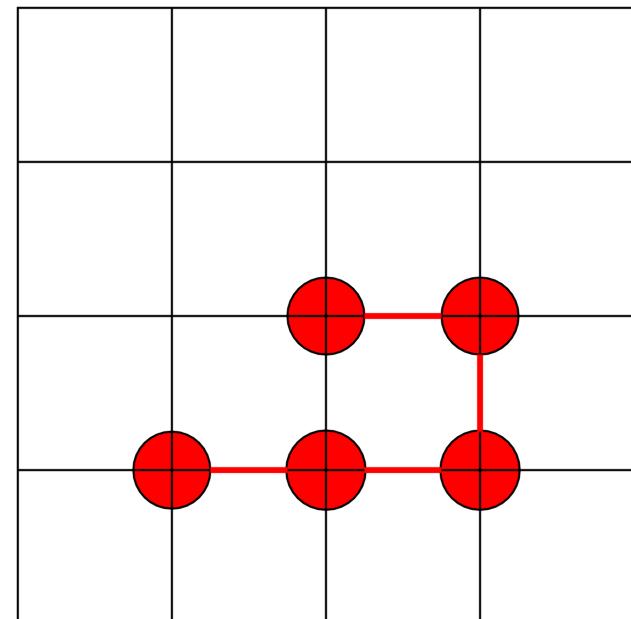
Rosenbluth Sampling: Correcting Bias

J. Batoulis and K. Kremer *J. Phys. A.: Math. Gen.*, 7, 259, 1988

$$P_N^{\text{RW}} = q_0^{-1} (q_0 - 1)^{-(N-1)} \quad q_0=4$$

$$P_N^{\text{RR}} = \prod_{i=1}^N (k_i)^{-1}$$

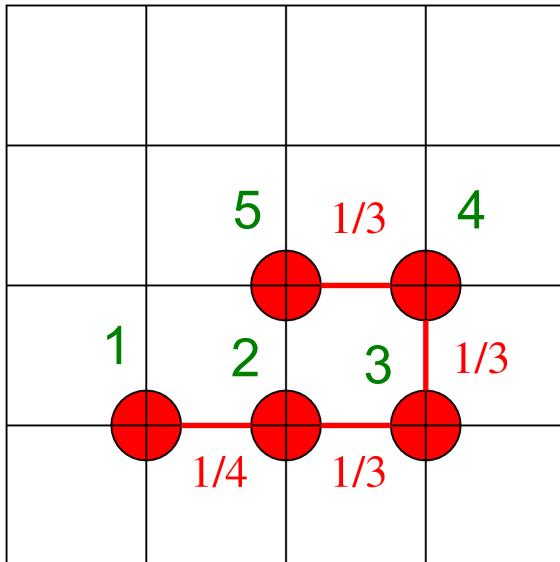
$$W_N = \prod_{i=1}^N \frac{k_i}{q_0 - 1}$$



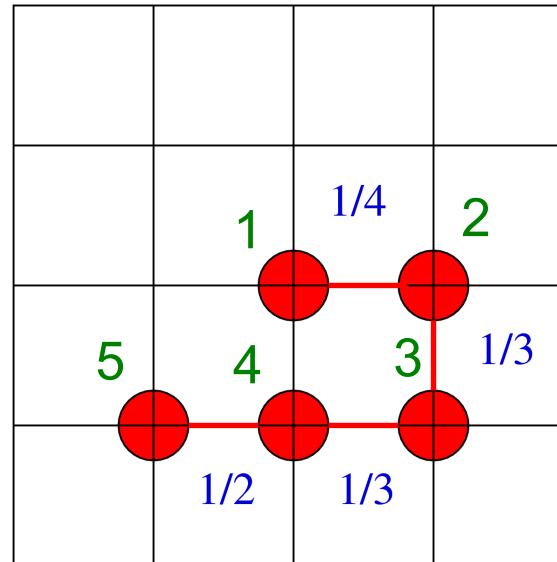
$$P_i^{\text{RR}} w_i = P_i^{\text{RW}}$$

Biased Growth: Rosenbluth Sampling

Forward



Reverse



$$\text{forward : } P_5^{\text{RR}} = 1/4 \times 1/3 \times 1/3 \times 1/3 = 1/108$$

$$W_5 = 4/4 \times 3/3 \times 3/3 \times 3/3 = 1$$

$$P_5 W_5 = 1/108 \times 1 = 1/108$$

$$\text{reverse : } P_5^{\text{RR}} = 1/4 \times 1/3 \times 1/3 \times 1/2 = 1/72$$

$$W_5 = 4/4 \times 3/3 \times 3/3 \times 2/3 = 2/3$$

$$P_5 W_5 = 1/72 \times 2/3 = 1/108$$

For random walk:

$$N=100 \quad q_0=12$$

Success rate: 0.022%

For Rosenbluth walk:

$$N=700 \quad q_0=12$$

Success rate: 99.3%

Configurational bias Monte Carlo: a new sampling scheme for flexible chains

By JÖRN ILJA SIEPMANN†

Department of Chemistry, University of Cambridge, Lensfield Road,
Cambridge CB2 1EW, UK

DAAN FRENKEL

FOM Institute for Atomic and Molecular Physics, Kruislaan 407,
NL-1098 SJ Amsterdam, The Netherlands

(Received 13 March 1991; accepted 31 May 1991)

We propose a novel approach that allows efficient numerical simulation of systems consisting of flexible chain molecules. The method is especially suitable for the numerical simulation of dense chain systems and monolayers. A new type of Monte Carlo move is introduced that makes it possible to carry out large scale conformational changes of the chain molecule in a single trial move. Our scheme is based on the selfavoiding random walk algorithm of Rosenbluth and Rosenbluth. As an illustration, we compare the results of a calculation of mean-square end to end lengths for single chains on a two-dimensional square lattice with corresponding data gained from other simulations.

Simulating the critical behaviour of complex fluids

J. Ilja Siepmann*, Sami Karaborni & Berend Smit†

Shell Research BV, Koninklijke/Shell-Laboratorium, Amsterdam,
PO Box 3003, 1003 AA Amsterdam, The Netherlands

