Monte Carlo
A general overview of methods, theory and practice

Alexandros Sopasakis
Department of Mathematics
UNC Charlotte

Lectures

- Lecture 1- Background and theoretical/analytical development of the Monte Carlo method
- **Lecture 2- Numerical simulation practices and common techniques used in modern modeling applications**
- Lecture 3- A research project perspective: application to traffic flow

Outline of lecture 1 from last week

We examined what it takes to create a Markov Chain Monte Carlo (MCMC)

The chain satisfies the *Markov property* (memoryless)

$$P(X(t_{n+1}) = \mu_{n+1} \mid X(t_n) = \mu_n, \ldots, X(t_0) = \mu_0) = P(X(t_{n+1}) = \mu_{n+1} \mid X(t_n) = \mu_n)$$

Task: sample from a given probability distribution π

and construct a discrete-time Markov Chain with a known probability matrix

P= $(P(\mu \to \nu))_{\mu,\nu \in \Sigma}$ having π as a stationary (invariant) distribution.

In that respect we:

- Define the transition probabilites $P(\mu \to \nu) = g(\mu \to \nu)\, A(\mu \to \nu)$
- Pick the selection probabilities $g(\mu \to \nu)$ for a proposed move from μ to ν
- Discuss optimal acceptance ratios $A(\mu \to \nu)$
- Use ergodicity to ensure that it is the invariant distribution π we sample
- Condition the Markov Chain to satisfy detailed balance.

$$\sum_{\nu} p_{\mu} P(\mu \to \nu) = \sum_{\nu} p_{\nu} P(\nu \to \mu)$$

The Ising Model and the Metropolis algorithm

An Ising model consists of micromagnets (dipoles) on a lattice. The magnetic spins are allowed to take two values -1,1. The energy of this system is given by its Hamiltonian,

$$E := H = -J \sum_{<i,j>} s_i s_j - h \sum_i s_i$$

where J is the interaction energy

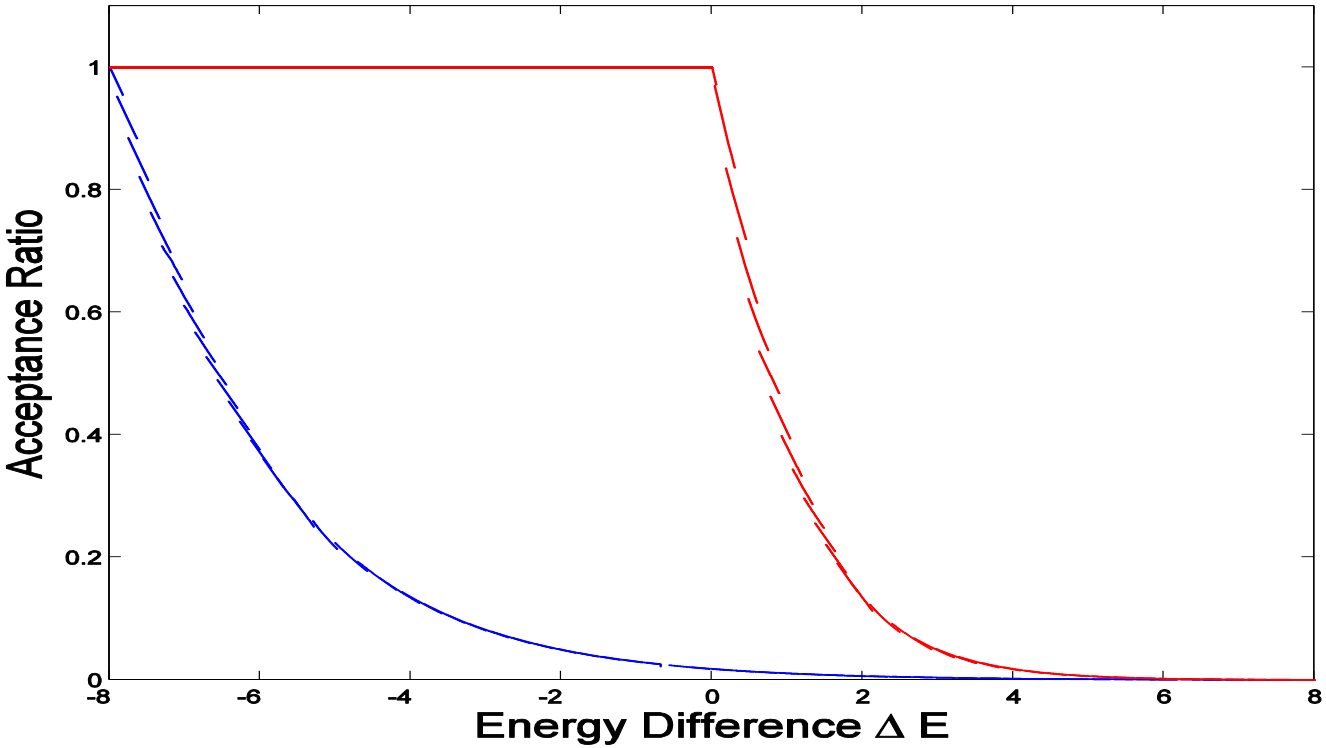
<i,j> represent nearest neighbor spins

with transition probabilities $P(\mu \to \nu) = A(\mu \to \nu)\, g(\mu \to \nu)$, as follows

$$g(\mu \to \nu) = 1/N$$

$$A(\mu \to \nu) = \begin{cases} e^{-\beta(E_\nu - E_\mu)} & \text{if } E_\nu > E_\mu \\ 1 & \text{otherwise} \end{cases}$$

This is the Metropolis algorithm!



Outline of what will follow

- Brief outline of overarching challenges in current computational problems
- Description of different types of interaction potentials and dynamics
- Description of different types of Monte Carlo algorithms according to application
- A typical Monte Carlo pseudo-code
- Collecting data
- Generating random sequences/numbers
- Complications with data collection
- Phase transitions

Year	Computer Name	Power (Watts)	Performance (adds/sec)	Memory (kByte)	Price (US dollars)
1951	UNIVAC I	124,500	1,900	48	$1,000,000
1964	IBM S360	10,000	500,000	64	$1,000,000
1965	PDP-8	500	330,000	4	$16,000
1976	Cray-1	60,000	166,000,000	32,768	$4,000,000
1981	IBM PC	150	240,000	256	$3,000
1991	HP 9000	500	50,000,000	16,384	$7,400
2005	IBM notebook	20	1,000,000,000	512,000	$1,900

Grand Challenges

Grand Challenges are the leading problems in science and engineering that can be solved only with the help of the fastest, most powerful computers.

They address issues of great societal impact,
such as
biomedicine,
the environment,
economic competitiveness, and
military.

Typical Equations for Scientific Computing

$$\vec{F} = m\,\vec{a} + \frac{dm}{dt}\,\vec{v}$$

Newton's equation (with damping)

$$\nabla^2 u = \frac{\partial u}{\partial t}$$

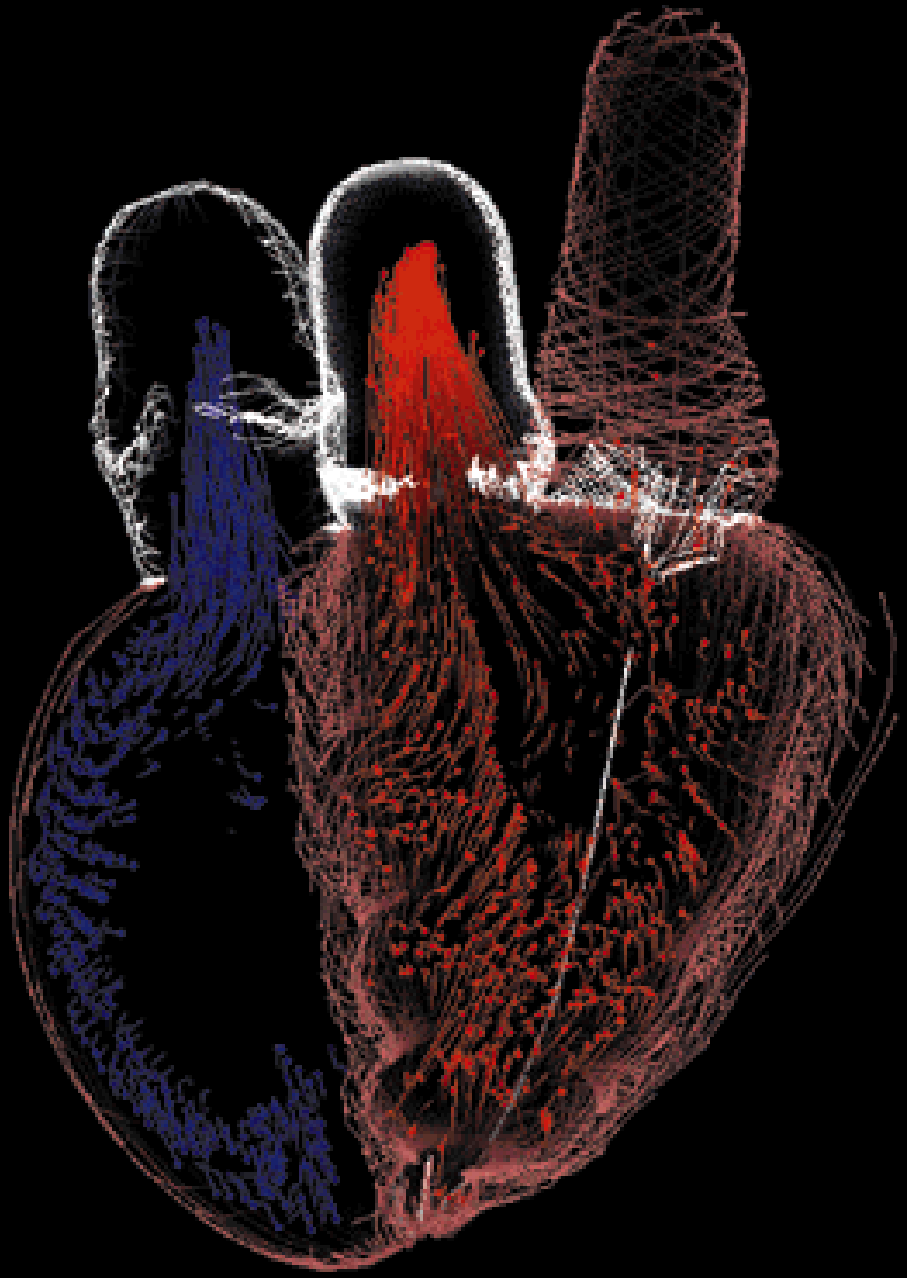
Diffusion or heat equation

$$F_j = \sum_{k=0}^{N-1} f_k e^{2\pi i jk/N}$$

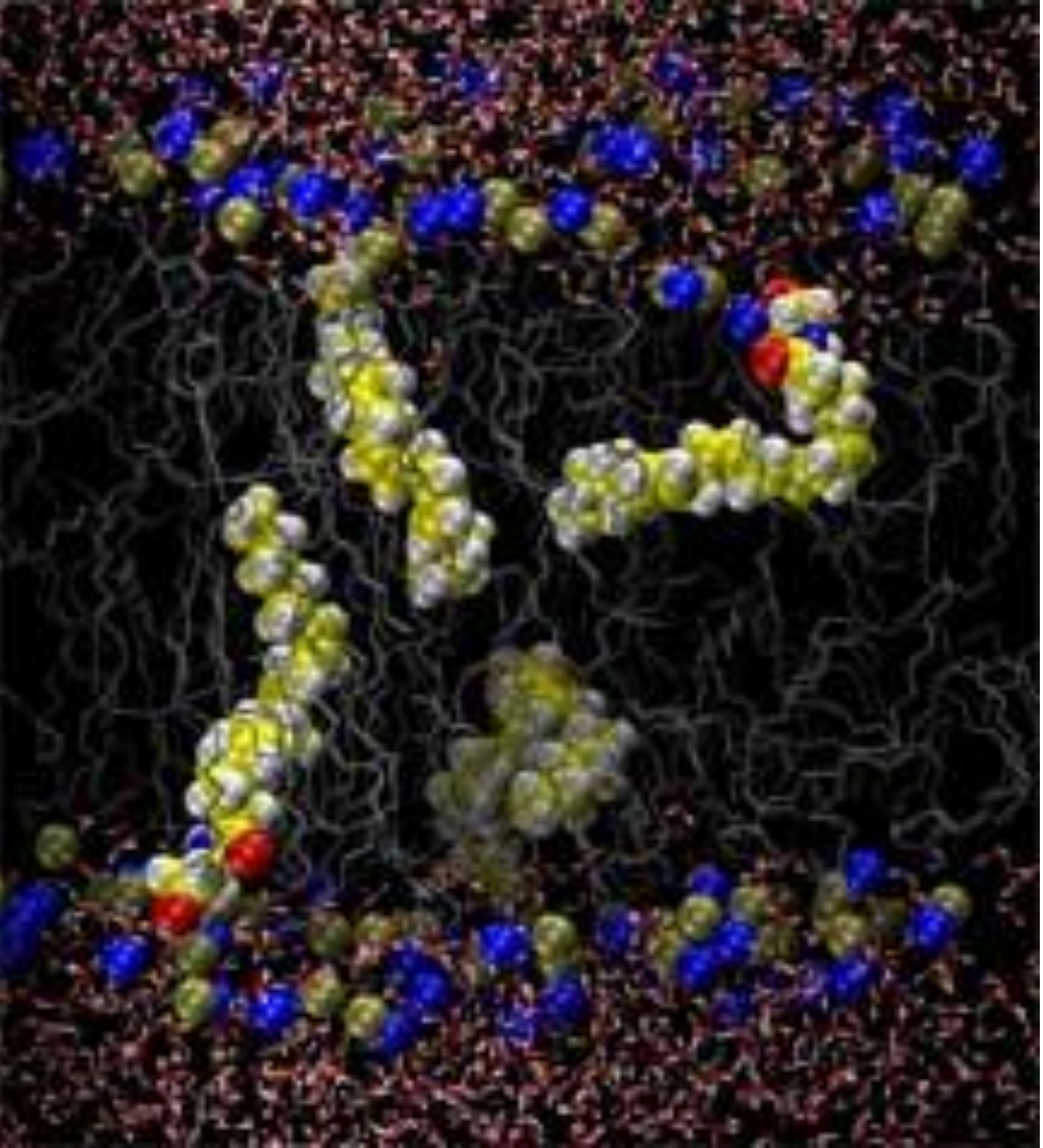
Fourier transform

$$-\frac{h^2}{8\pi^2 m}\nabla^2 \Psi(r,t) + V\,\Psi(r,t) = -\frac{h}{2\pi i}\frac{\partial \Psi(r,t)}{\partial t}$$

Schrödinger equation



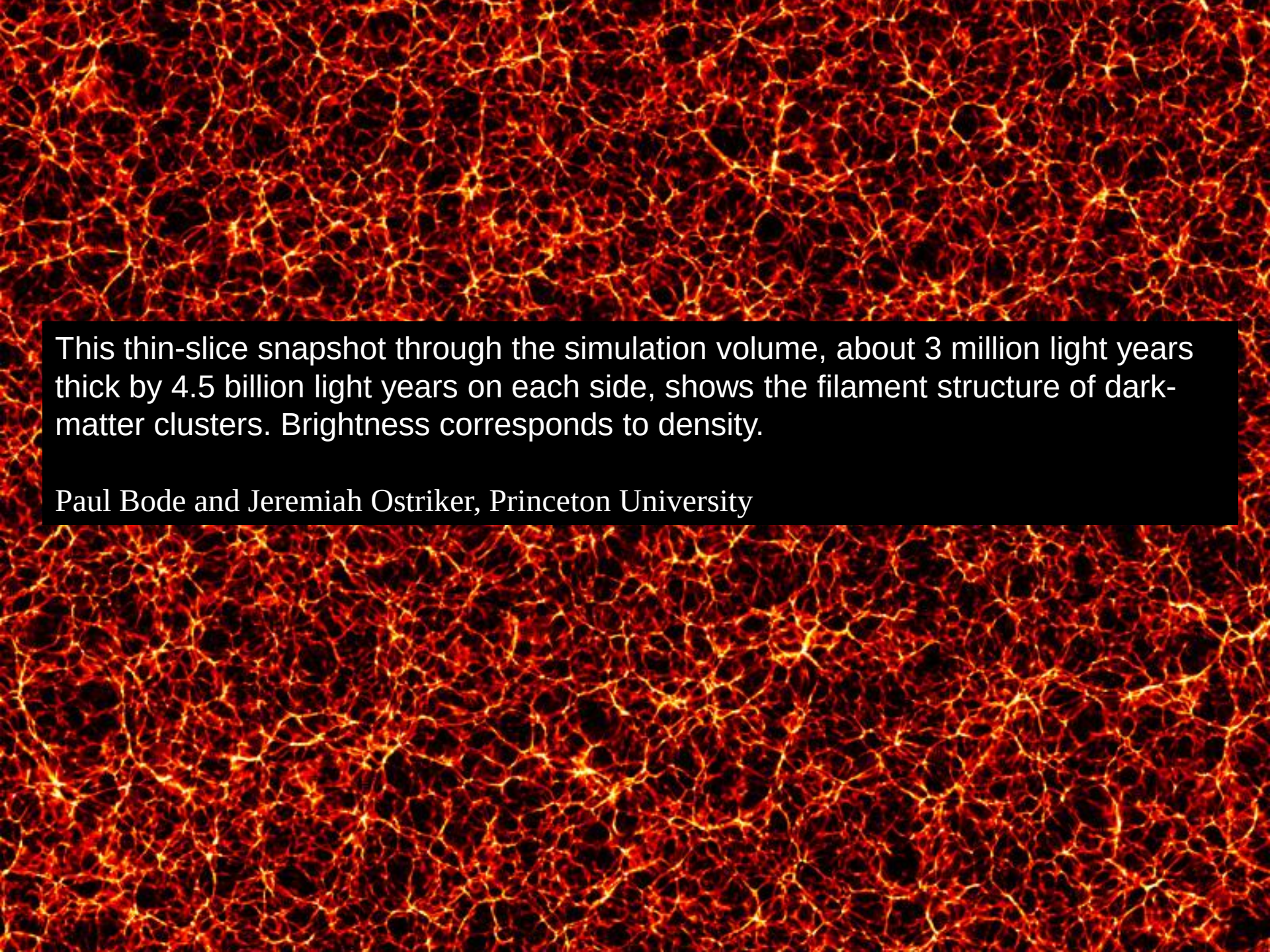
Blood flow
in heart
from
Navier-Stokes
equation, NIH



Brain Chemistry:
bi-layer sandwich of lipids

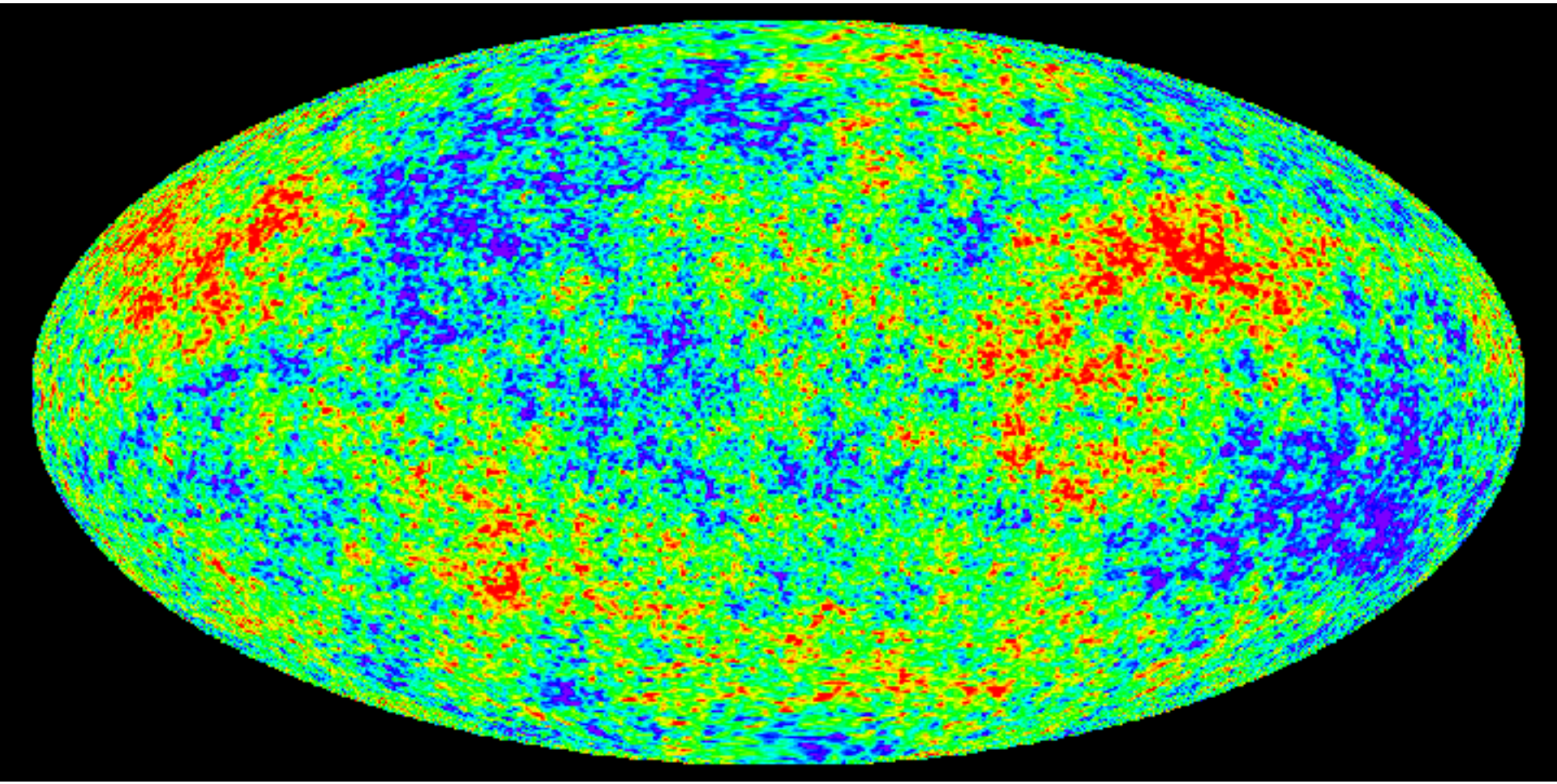
nitrogen (blue)
phosphorous (gold)
heads facing outward on both sides of filamentary tails (gray).

Patti Reggio,
UNC, Greensboro
Diane Lynch,
Kennesaw State Univ.

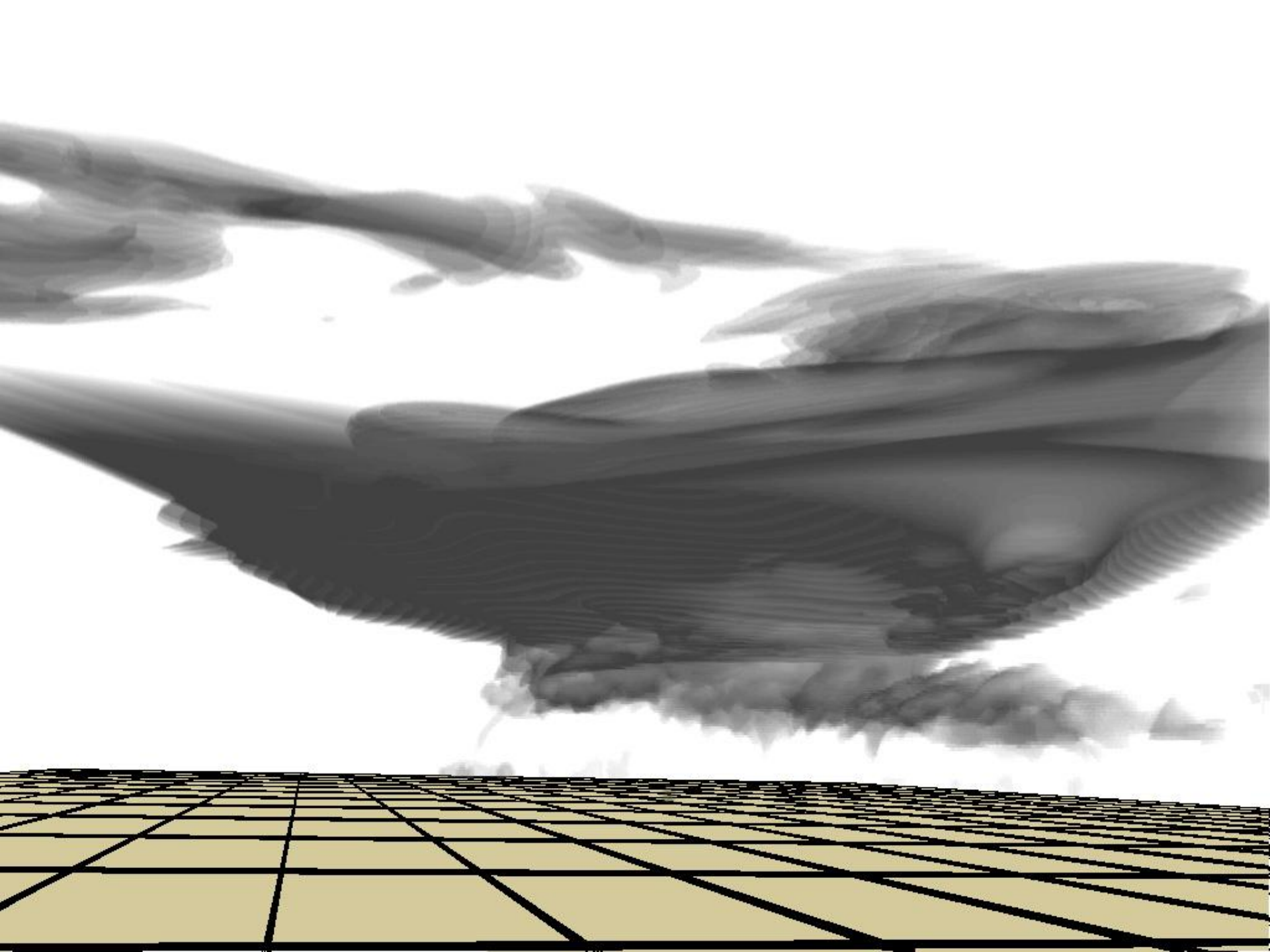


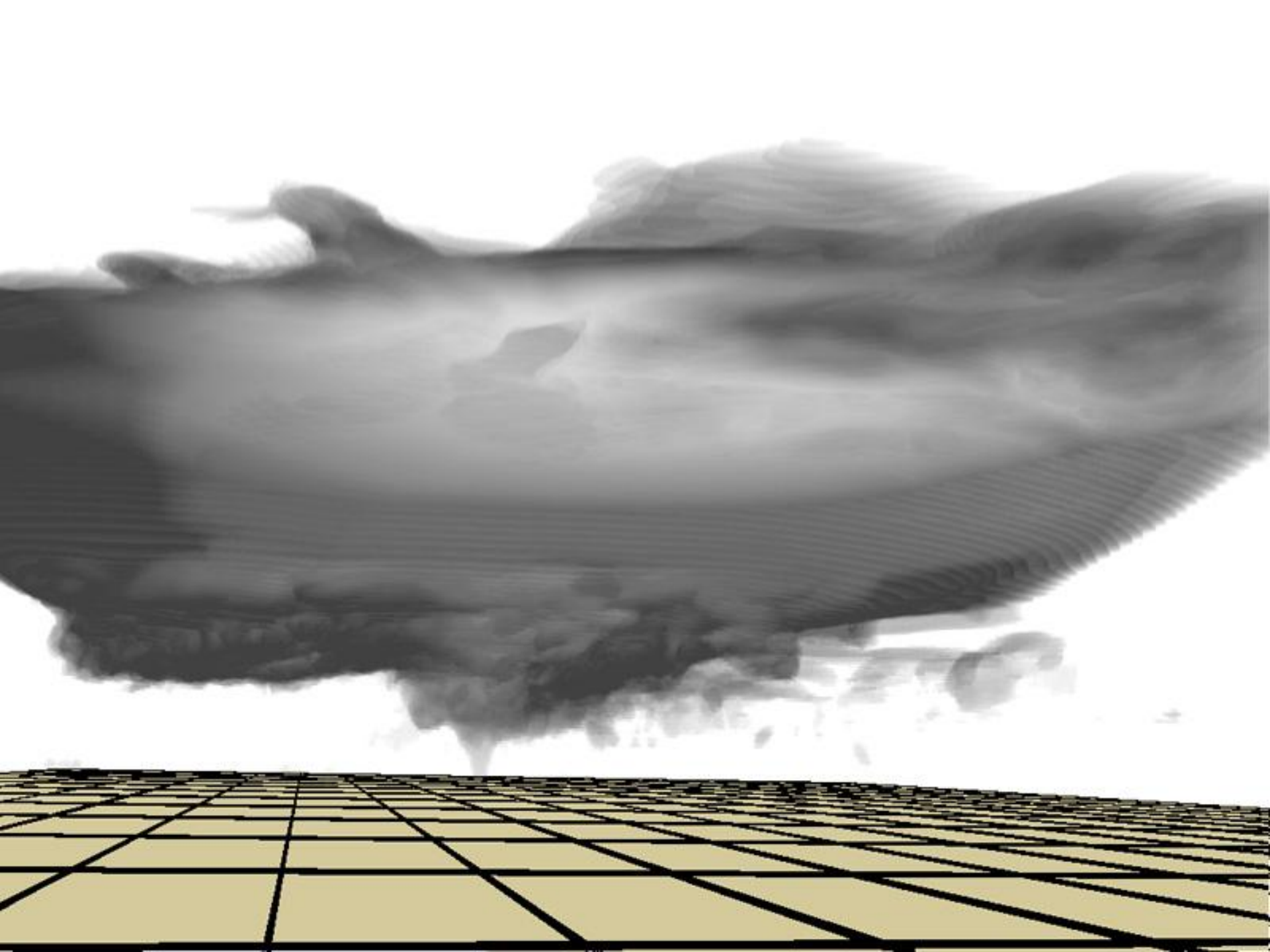
This thin-slice snapshot through the simulation volume, about 3 million light years thick by 4.5 billion light years on each side, shows the filament structure of dark-matter clusters. Brightness corresponds to density.

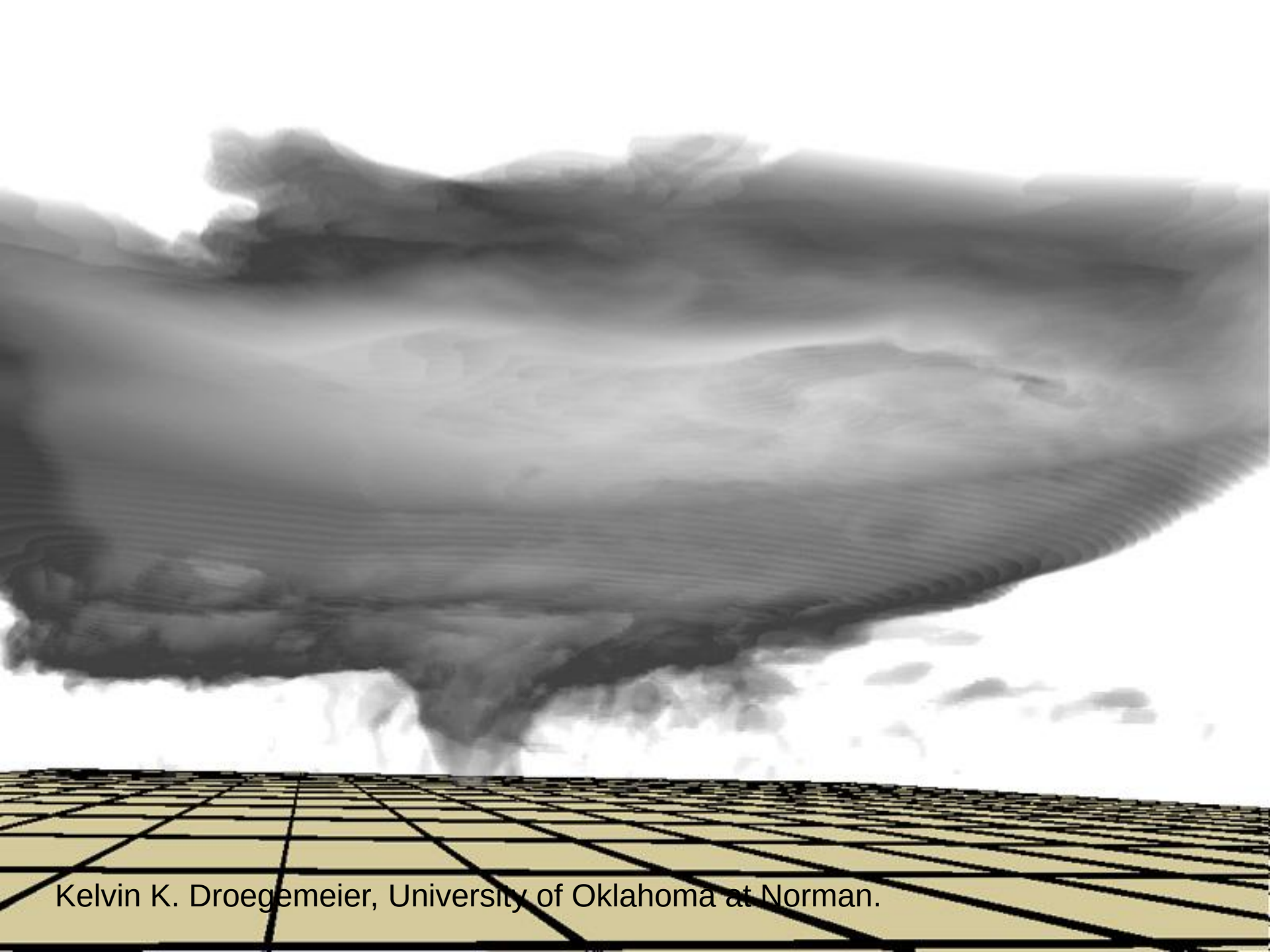
Paul Bode and Jeremiah Ostriker, Princeton University



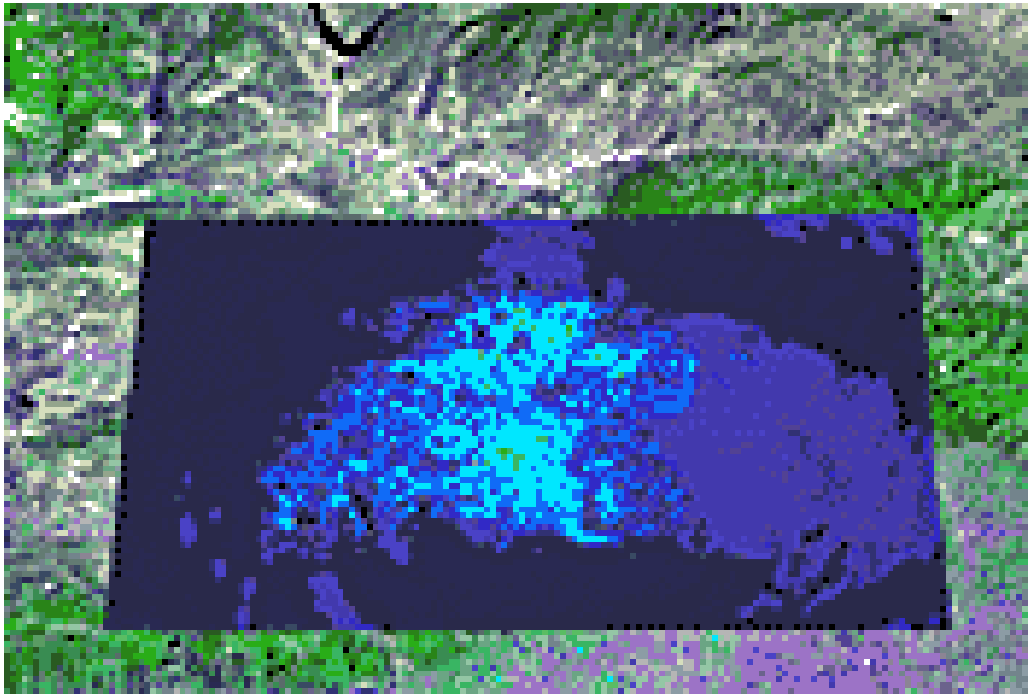
The infant universe hatching from its structureless shell.
This map represents Edmund Bertschinger's simulations on the CRAY T3D at Pittsburgh. This map shows negative (blue) and positive (red) fluctuations of 0.0002 degrees K. The simulation assumed a mixed hot and cold dark matter model with 5 eV neutrino mass.







Kelvin K. Droegemeier, University of Oklahoma at Norman.



Animation of a simulated earthquake in the San Fernando valley. Color depicts the peak magnitude of ground displacement. The simulation covered a 54 x 33 kilometer area, superimposed here on a satellite view of topography, to a depth of 15 km.

Carnegie Mellon University

Computational Needs- Biology & Bioinformatics

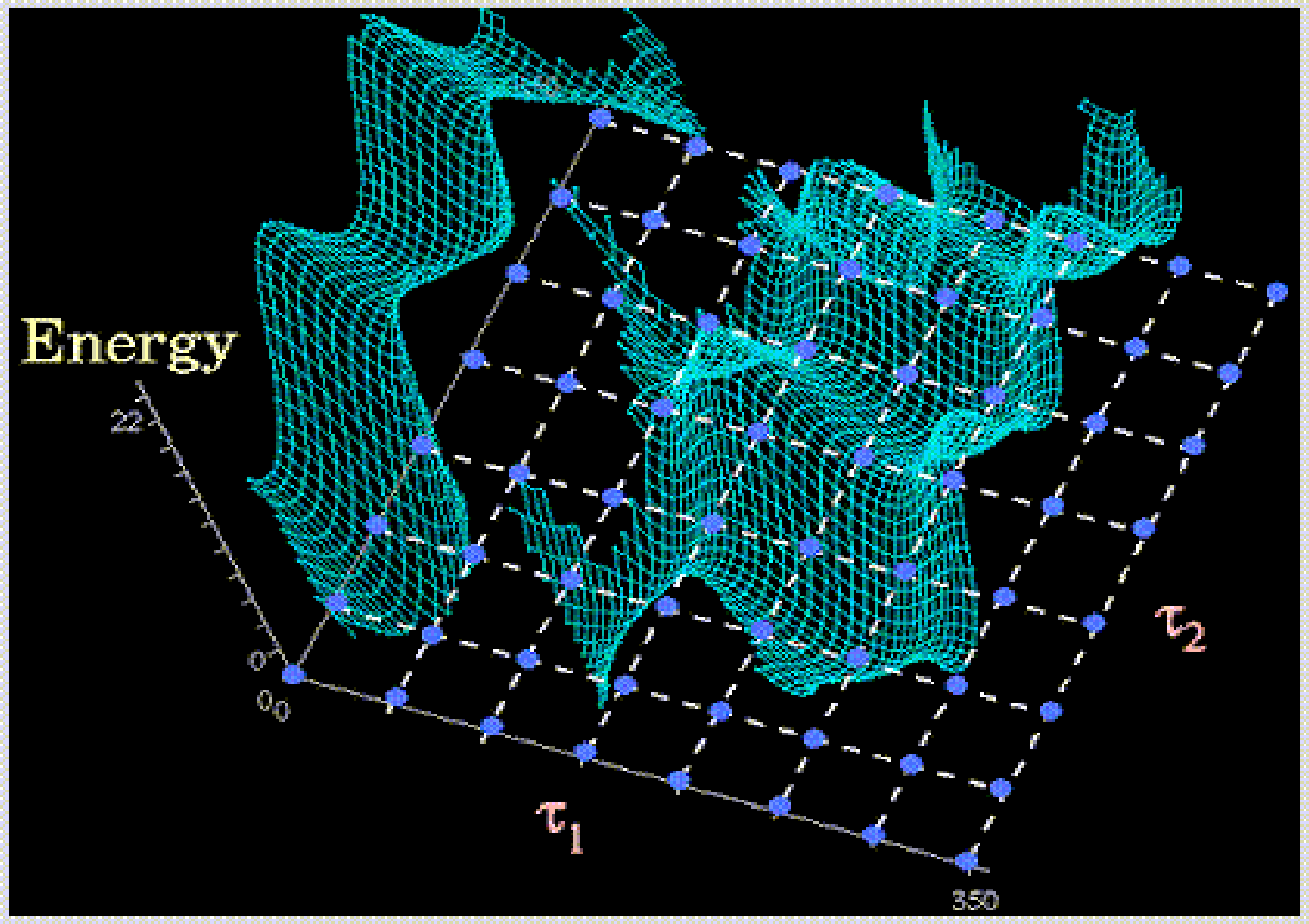
Problem Component	Computing Speed	Storage
Genome Assembly	>10 TeraFlops sustained to keep up with expected sequencing rates	300 TB of trace files per genome
Protein Structure Prediction	>100 TeraFlops per protein set in one microbial genome	Petabytes
Classical Molecular Dynamics	100 TeraFlops per DNA-protein interaction	10s of Petabytes
First Principles Molecular Dynamics	1 PetaFlops per reaction in enzyme active site	100s of Petabytes
Simulations of Biological Networks	>1 TeraFlops for simple correlation analyses of small biological networks	1000s of Petabytes

TOP 10
Sites for June 2006

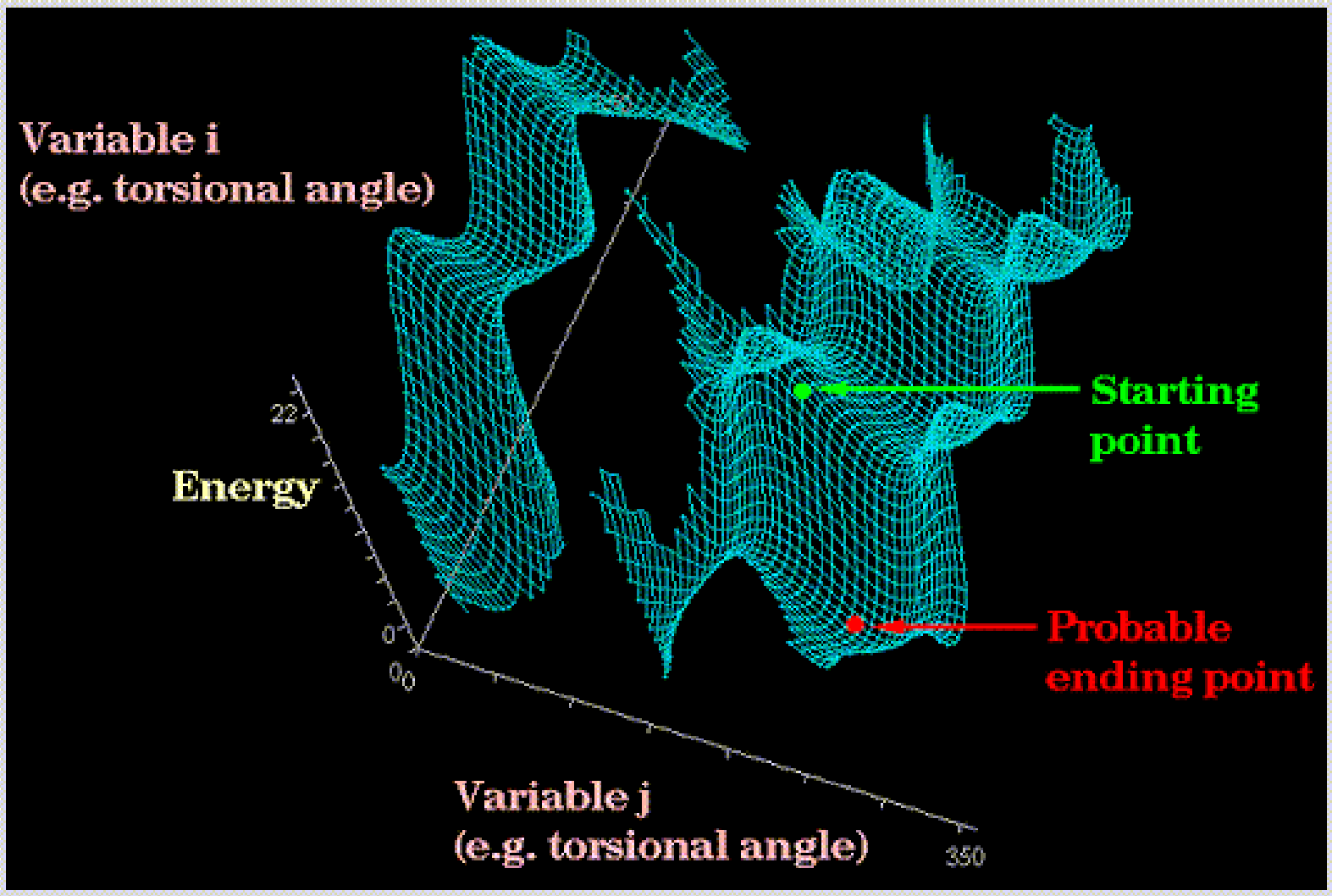
www.top500.org

Site	System Family	# Processors
DOE/NNSA/LLNL United States	Blue Gene L IBM	131 072
IBM Watson United States	Blue Gene L IBM	40 960
DOE/NNSA/LLNL United States	ASCPurple (p –series) IBM	12 208
NASA/Ames United States	Altix SGI	10 160
CEA France	Tera-10 (SMP cluster) Bull	8 704
Sandia Nat. Lab. United States	PowerEdge Dell	9 024
Tokyo Inst. Tech. Japan	Sun Fire NEC/Sun	10 368
FZ Juelich Germany	Blue Gene L IBM	16 384
Sandia Nat. Lab. United States	XT3 Cray	10 880
Earth Simulation Japan	NEC Vector NEC	5 120

An example of a typical energy landscape over a given lattice



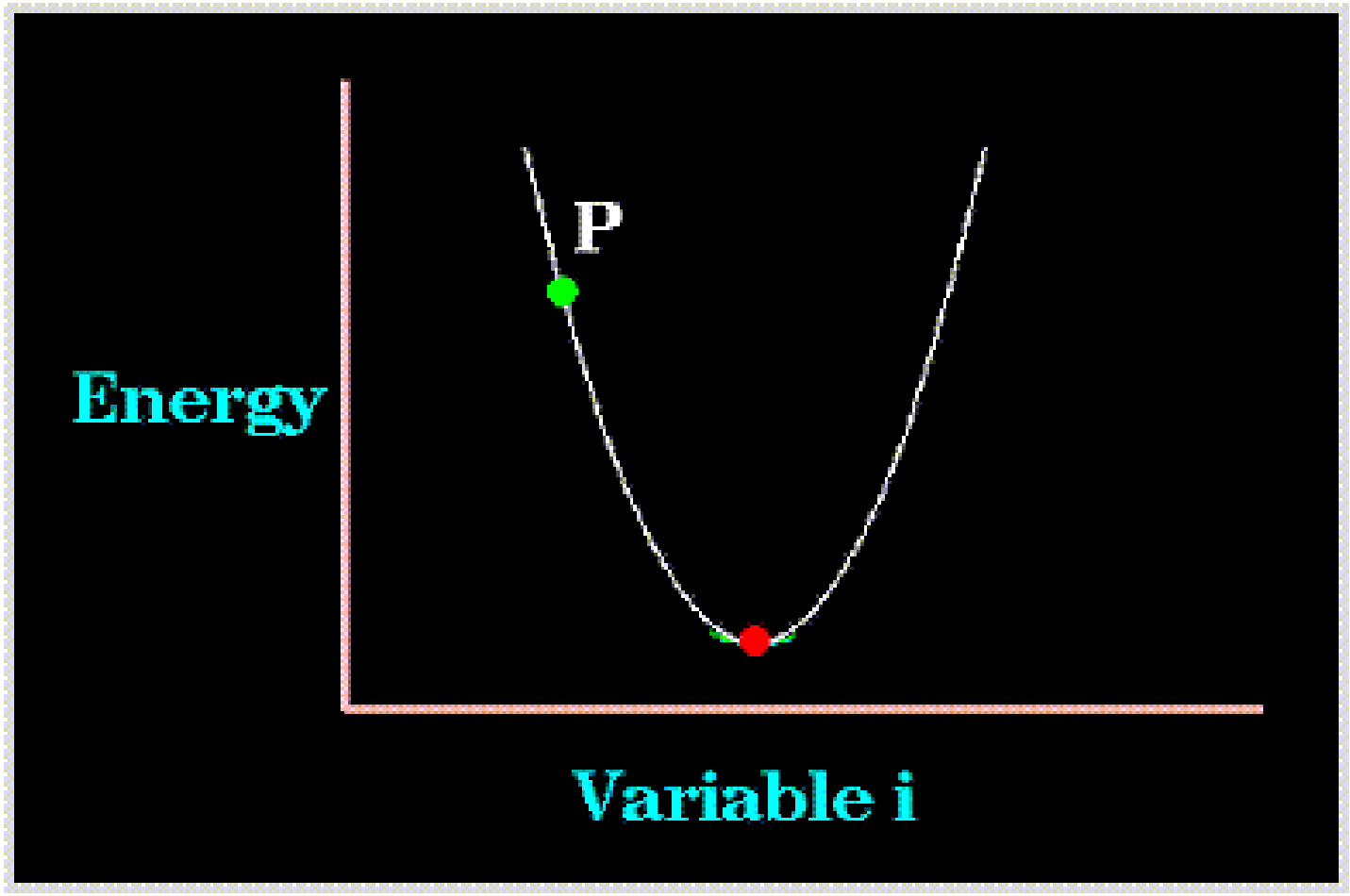
System energy tendencies



Implementation and Pseudo-code for Metropolis single spin-flip

1. Start simulation with an initial condition for all spins (typical T=0, T=infinity). Calculate the system energy $E\mu$ at this state μ
2. Choose a spin, with equal probability, from the lattice and calculate the new system energy $E\nu$ from flipping that spin.
3. Calculate difference in energy between old and new states $E\mu - E\nu$.
4. Pick a random number **r** between 0 and 1
5. Use the acceptance ratio to either move to the new state ν or not as follows:
   move to the new state ν if $E\mu - E\nu < 0$
   mote to the new state ν if **r** < $A(\mu \rightarrow \nu)$
6. Continue until statistics of interest have been collected

System energy tendencies



Obtaining statistical information for quantities of interest

In terms of measuring quantities of interest, say Q, we are interested at evaluating

$$\langle Q \rangle = \sum_{\mu} Q_{\mu} w_{\mu}(t)$$

Once we obtain the solution for $W_{\mu}(t)$ from

$$\lim_{t \to \infty} w_{\mu}(t) = p_{\mu} := \frac{1}{Z} e^{-\beta E_{\mu}}$$

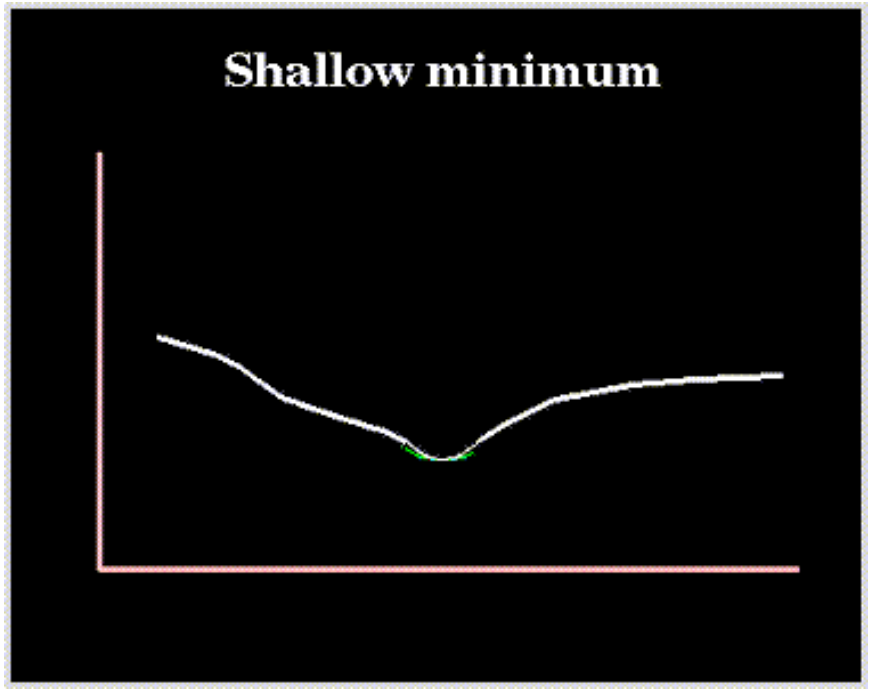
then $\langle Q \rangle = \dfrac{\sum_{\mu} Q_{\mu} e^{-\beta E_{\mu}}}{\sum_{\mu} e^{-\beta E_{\mu}}}$

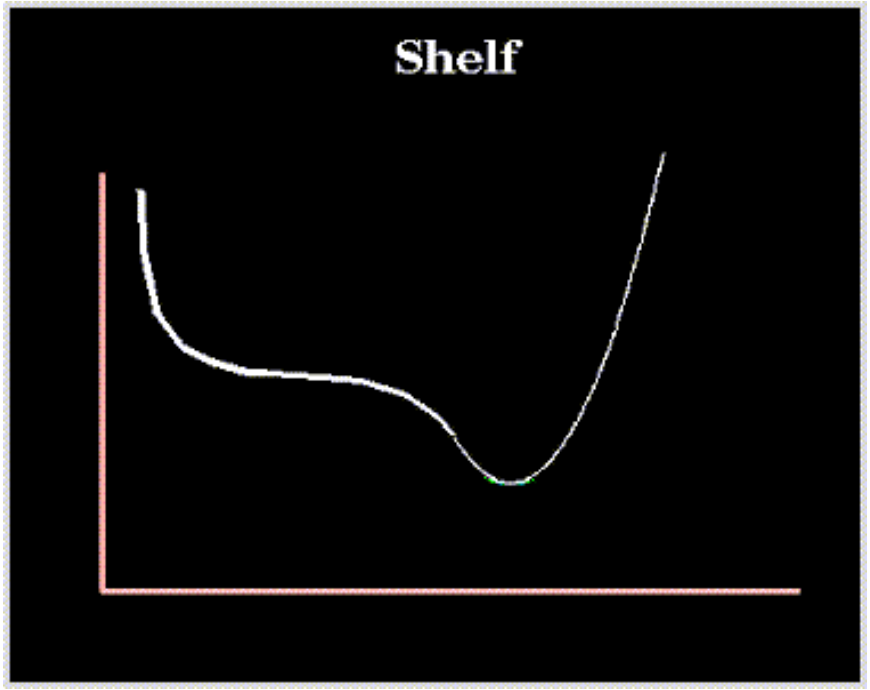
and we can use the following estimator for <Q>

$$Q_M = \frac{\sum_{i=1..M} Q_{\mu_i} p_{\mu_i}^{-1} e^{-\beta E_{\mu_i}}}{\sum_{j=1..M} p_{\mu_j}^{-1} e^{-\beta E_{\mu_j}}}$$

Thus $Q_M = \dfrac{1}{M} \sum_{i=1}^{M} Q_{\mu_i}$

Complications in collecting reliable statistics

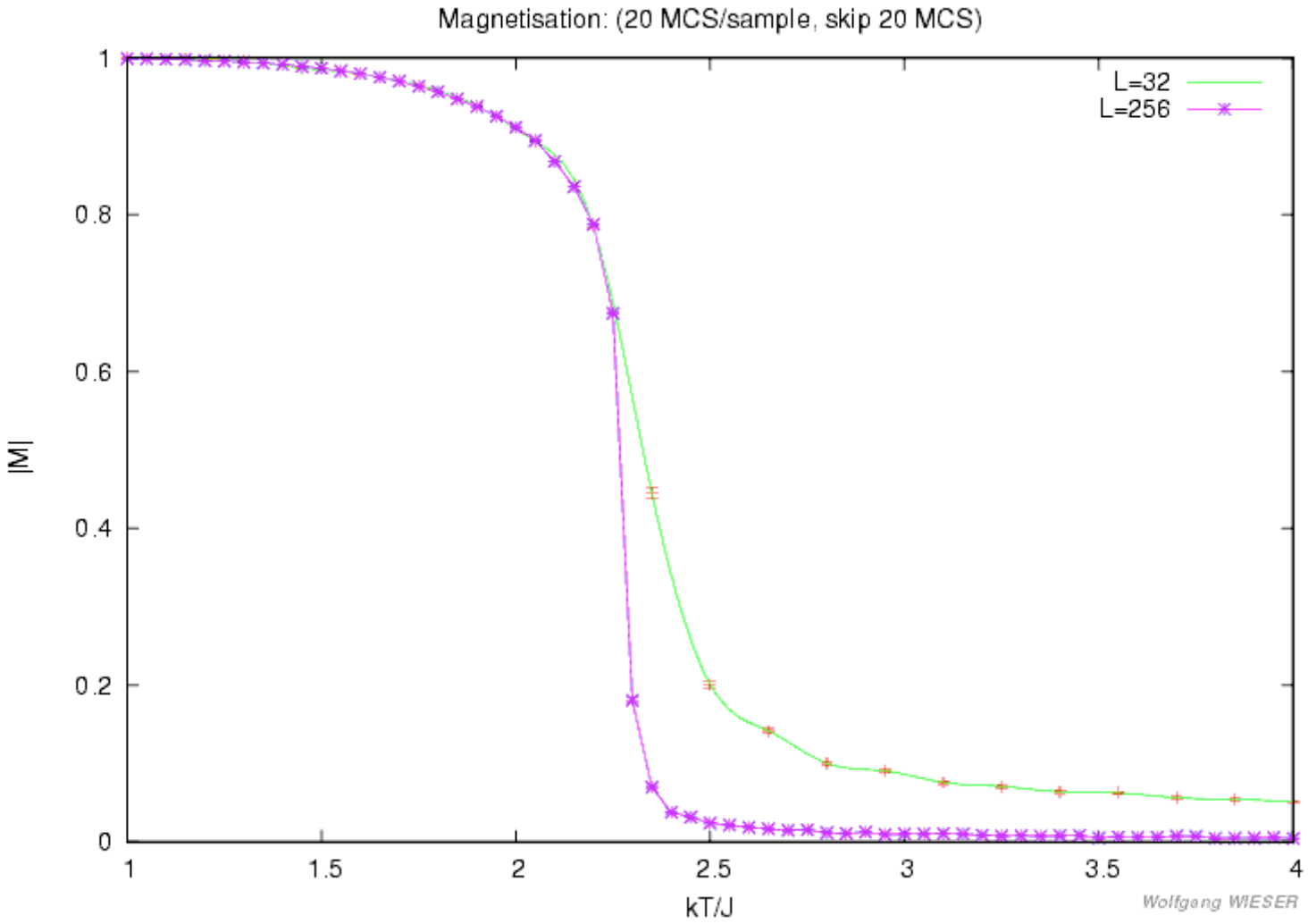




The phase transition phenomenon
Ising model

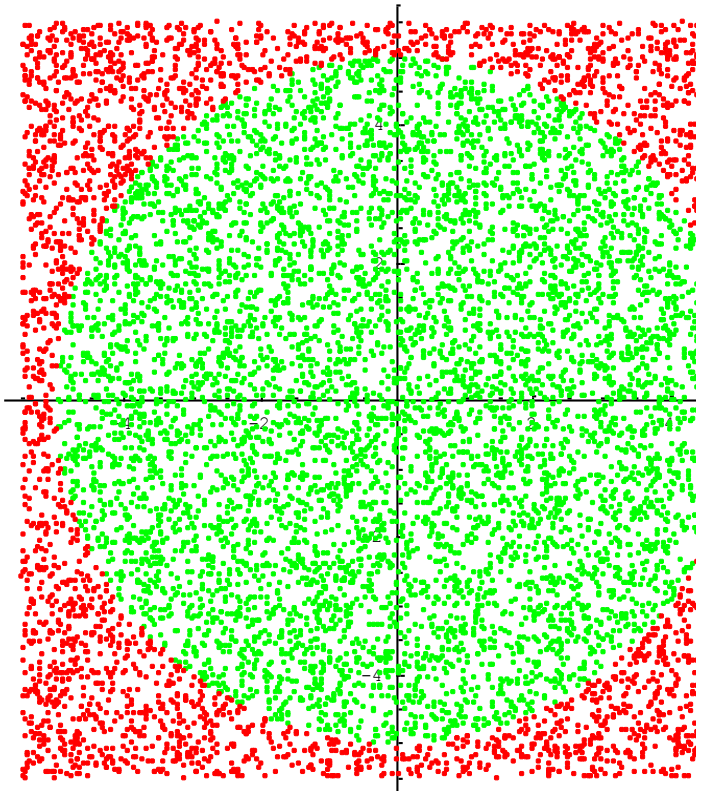
$$T_c = \frac{2J}{\log(1+\sqrt{2})} \approx 2.269J$$

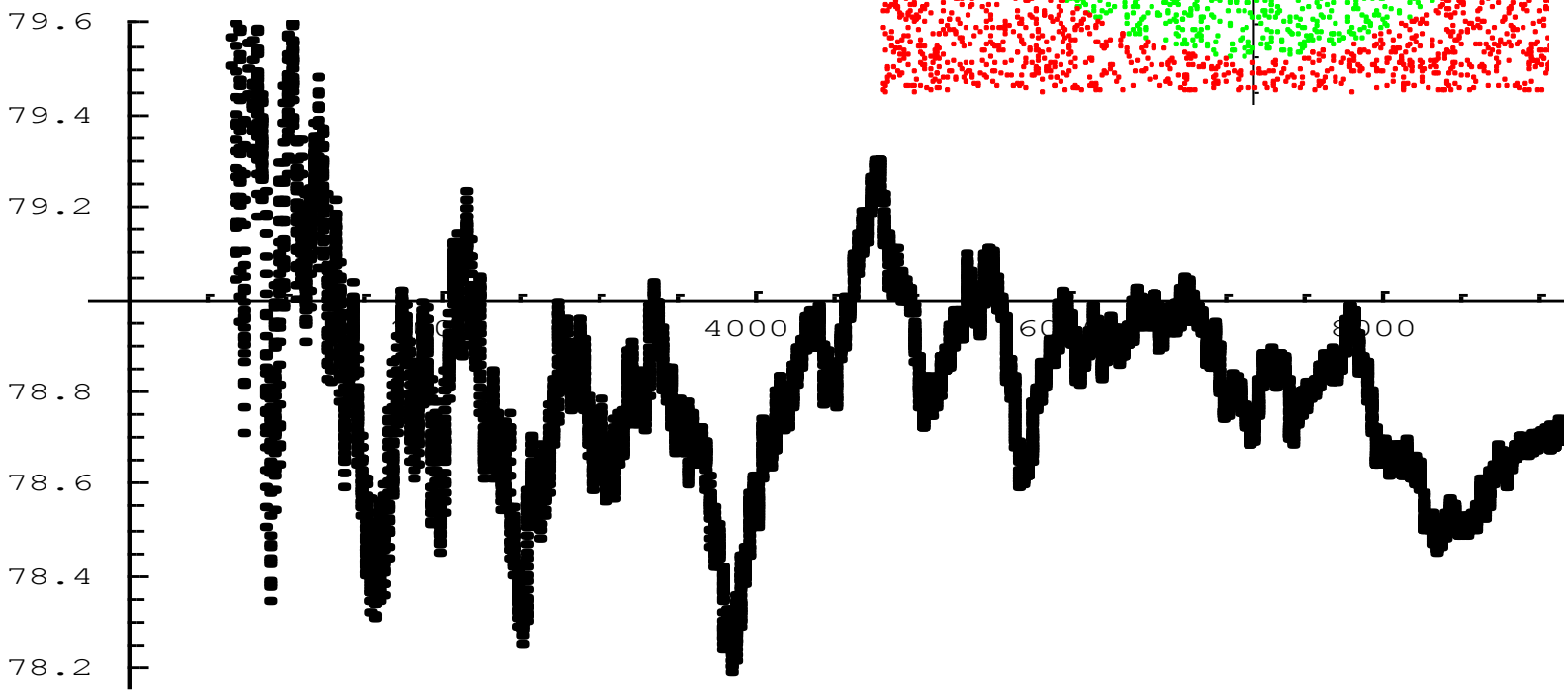
Onsager 1944, 1949, Yang 1952



Example: the area of a circle

- Sample points randomly from square surrounding circle of radius 5
- 10,000 sample points
- Acceptance criterion: inside circle
- Actual area: 78.54
- Calculated area: 78.66



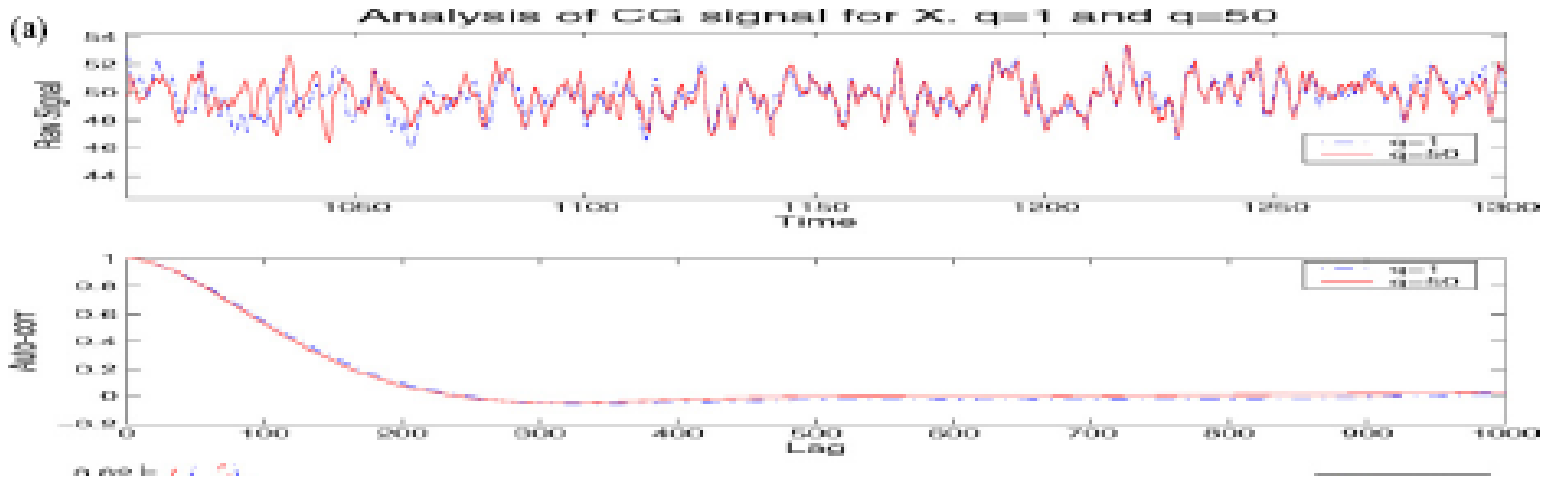


Equilibrium and Autocorrelation

One way to obtain reliable information, such as calculating an expected value, about quantities of interest is to obtain

$$Corr(Q)(t^*) = \int (Q(t+t^*) - < Q >)(Q(t) - < Q >) dt$$

This also reveals the correct MC **time τ necessary for equilibration**.



Usually we are supposed to re-run our simulation a number of times at the end of which we should collect our data.

As a rule of thumb we can use the resulting MC simulation and collect independent statistics for our parameters of interest every 2τ.

Generating Random Numbers

"Anyone who considers arithmetic methods of producing random digits is, of course, in a state of sin" John Von Neumann, 1951

RNG are designed to produce an unbiased sequence of numbers which appear to lack any pattern.

Reality: there are always patterns which somehow connect our random numbers

In practice: do not use the sequence past its cycle.

Ancient physical methods of generating random sequences: dice, flipping coins, wheels used as roulettes , Yi Jing (Classic of Changes).

Modern computational methods: pseudo-random number generators

RNG are initiated using a computer real time clock.

A congruential generator

As an example consider the following sequence which generates a sequence with maximum number m. We use a linear congruential generator [Knuth 1981]

$$I_k = (aI_{k-1} + b) \bmod m$$

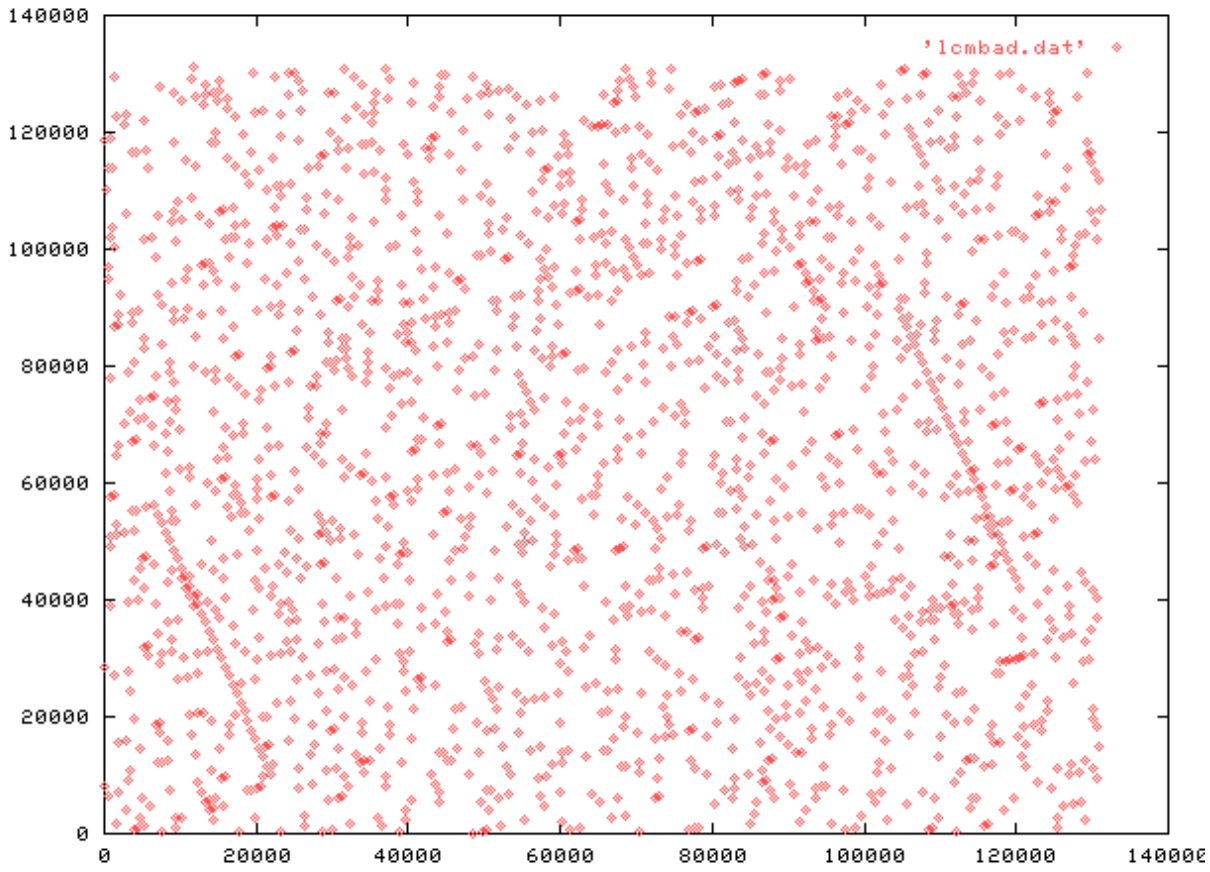
The values of the **multiplier a** the **increment b** and the **modulus m** are critical towards the quality of the random number produced.

Example 1. a = 1277, b=0, m=131072

Example 1.

a = 1277, b=0, m=131072

The following scatter plot for 2000 pairs from this generator shows linear bands.



A congruential generator

As an example consider the following generator which produces a sequence with maximum number m. We use a linear congruential generator [Knuth 1981]

$$I_k = (aI_{k-1} + b) \bmod m$$

The values of the **multiplier a** the **increment b** and the **modulus m** are critical towards the quality of the random number produced.

Example 1. a = 1277, b=0, m=131072 (naive example)

Example 2. a=16807, b=0, m=2147483647 [Park and Miller 1988]

Known as the **minimal standard RNG**. Mostly used for the RNG function in compilers

Example 3. Mersenne twister

Holds the record for longest period: 2^19937-1 (it is also very fast)

Major Algorithms

- Metropolis (1949): flipping single spins per time increment

- Wolff (1989): flipping a cluster instead of a single spin
- Swendsen-Wang (1987), Niedermayer's (1988): decide on flipping (or not) all clusters at once (with probability ½).

- Multigrid Methods (Kandel 1989) – reduce critical slowing down by grouping clusters of spins into single blocks and flipping them using Metropolis

- Kawasaki (1965) – used for conservative systems together with spin-exchange type dynamics
- Kinetic Monte Carlo (1979): variable time step within which a single spin is flipped

Algorithms according to type and number of spins flipped

- **Single spin-flip:** every step involves switching or not a single spin on the lattice
- **Cluster:** randomly choose a spin on the lattice and create a cluster of similar such spins around it. Probability to be added to cluster is given by

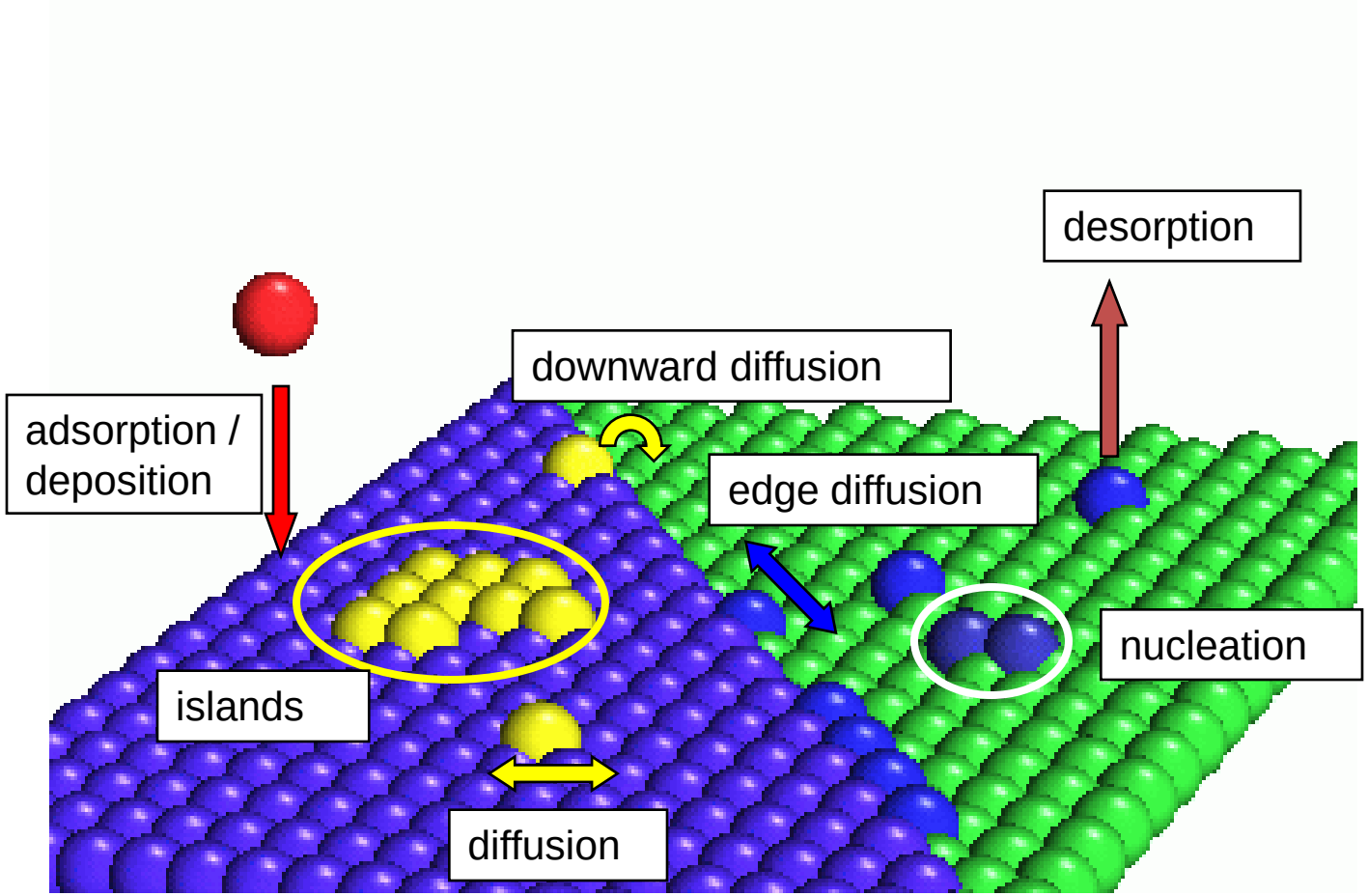
$$P = 1 - e^{-2\beta J}$$

- **Potts models:** allow spins to take values between 0 and q.
- **Heat Bath:** choose a random spin on lattice and regardless of its value change it to a new value, say l, where $0 \le l \le q$

  Then switch or not based on probability: $P_n = \dfrac{e^{-\beta E_n}}{\sum_{m=1}^{q} e^{-\beta E_m}}$

  This is much more efficient than Metropolis for large q values).
- **Cluster for Potts:** Use the probability $P_C = 1 - e^{-\beta J}$ for a spin to be added to the clusters and allow spins to take values between 0 and q.
- **Continuous spin - XY and Heisenberg:** allow the spins to take all the values possible in the xy plane or the sphere.

Examples of different types of microscopic processes on a growing surface



+more: incorporation, *knockout*
attachment to edges / islands
detachment processes, ...

Interaction Potential Dynamics

Suppose a general spin-flip rate

$$r(x, y, \sigma) = \Psi(-\beta \Delta_{x,y} H(\sigma))$$

Then the following is a short list of the possible choices for dynamics

Glauber	$\Psi(r) = (1+e^{r})^{-1}$
Metropolis/Arrhenius	$\Psi(r) = e^{-r}$
Kawasaki	$\Psi(r) = 2(1+e^{r})^{-1}$

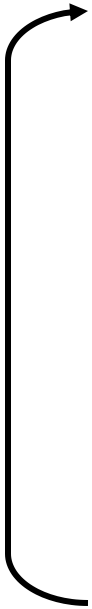
Kinetic Monte Carlo pseudocode (rejection free)

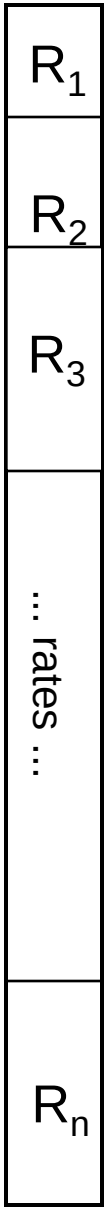
· provide initial configuration of the system
· catalogue all relevant processes i=1,2,...n
and corresponding Arrhenius rates

· pick one of the possible events
with probability $p_i \propto R_i$

· perform the selected event
(evaluate physical real time step)

· update the catalogue of processes
and associated energy barriers and rates





Advanced Topics

- Adaptive Monte Carlo
- Accelerating methods for MC and tau-leaping techniques
- Simulated annealing / tempering for global optimization
- Metropolis-Hastings algorithm – allows use of non-symmetric trial distributions in order to improve efficiency
- Hybrid systems of stochastic / deterministic systems

Recap and next talk

In this talk we presented the following topics:

- Acceptance ratios for Metropolis algorithm
- Different types of dynamics
- Different types of algorithms
- Outlined a general pseudo-code for Monte Carlo
- Examined random number generators

In the next talk we present hybrid systems consisting of deterministic stochastic components usually modeling different types of chemical or biological processes involving some kind of well described transport/diffusion coupled to noisy microscopic type boundary contributions. This multi-scale system will be examined in the microscopic as well as the macroscopic regimes for behavior and solutions during critical phenomena such as metastability, phase transitions, etc.

Books and other useful references

Books:

M. E. Kalos and P. A. Whitlock, 1986, *Monte Carlo Methods, Volume 1: Basics*, Wiley, New York.

K. Binder and D. W. Heermann, 1992, Monte Carlo simulation in statistical physics, Springer-Verlag, Berlin.

M. E. J. Newman and G. T. Bakerma, 1999, *Monte Carlo Methods in Statistical Physics*, Clarendon Press-Oxford.

D.E. Knuth, 1981, The art of computer programming, Vol 2, Seminumerical Algorithms, Addison Wesley, Reading Mass.

More General Books:

Gibbs, 1902, Elementary principles in statistical mechanics(reprinted: 1981, Ox. Bow Press, Woodridge)

H. Dorrie, 1965, *One hundred great problems of elementary mathematics*, Dover, New York.

E. Segre, 1980, *From X-Rays to Quarks*, Freeman, San Fransisco

Most Important Scientific Papers:

N. Metropolis and S. Ulam, 1949, J. Am. Stat. Assoc., 44, 335.

Bortz, Kalos and Lebowitch, 1975, J. Comput. Phys. 17, 10.

Swendsen and Wang, 1987, Phys. Rev. Lett., 58, 86.

Kawasaki, 1965, Phys. Rev. 145, 224.

Sweeny, 1983, Phys. Rev. B, 27, 4445.

Wolff, 1989, Phys. Rev. Lett. 62, 361.