An introduction to quantum Monte Carlo cluster algorithms

Matthias Troyer (ETH Zürich)

Classical cluster algorithms
Quantum Monte Carlo
- Path integrals
- Stochastis Series Expansion

Quantum Cluster algorithms
- Loop algorithm
- Worm algorithm

Wang-Landau sampling
- allows tunneling through free energy barrieres
- efficient for 1st order transitions and spin glasses

What can cluster algorithms do?

- Local updates (before 1994)
  - 200 spins
  - T=0.1
- Cluster algorithms (after 1995)
  - 2D quantum phase transition: 20'000 spins at T=0.005
  - 2D square lattice: 1'000'000 spins at T=0.2
  - 3D antiferromagnet 16'000'00 spins at T≈1 (Sandvik)
- Cluster algorithms allow to reach asymptotic scaling regime
  - Accurate estimates of critical exponents
  - Check of scaling predictions

Our applications

- Quasi-1D spin systems
  - Ladders, chains, spin-orbital models
- 2D+3D antiferromagnets
  - Quantum phase transition sin Heisenberg models
  - Magnetic field effects (quantum Hall bilayers)
  - Impurity doping
- 2D bosons
  - Solids, supersolids, smectics, nematics
  - Dirty hardcore bosons
  - Trapped bosons (BEC in optical lattice)
- Realistic modeling
  - 2D vanadates
  - Coupled cuprate and vanadate ladders
- Current and future
  - Search for spin liquids in SU(N) antiferromagnets
  - Dissipation-driven phase transitions
  - Quantum spin glasses
  - Trapped atomic condensates in optical lattices

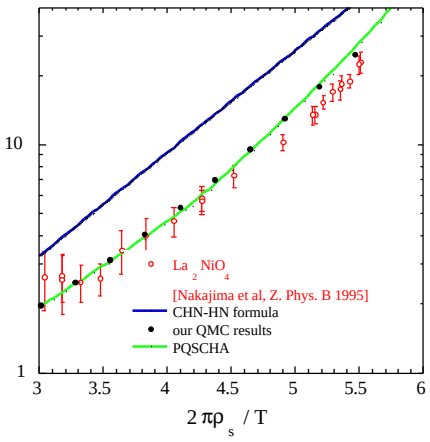
S=1 square lattice: nickelates

- Quantum field theory:

$$\frac{\xi}{a} = \frac{e}{8}\frac{\hbar c/a}{2\pi\rho_s}\exp\left(\frac{2\pi\rho_s}{k_B T}\right)\left[1-\frac{1}{2}\cdot\frac{k_B T}{2\pi\rho_s}+O\left(\frac{k_B^2 T^2}{4\pi^2\rho_s^2}\right)+\ldots\right]$$

- Semi-classical PQSCHA
  - Cuccoli et al. , PRL 1996, PRB 1997
  - quantum effects renormalize the temperature

$$\xi = \xi_{classical}\left(f_S(T)\right)$$



- Quantum effects weak for spin S>1/2, semi-classical description applicable

S=1/2 square lattice: cuprates

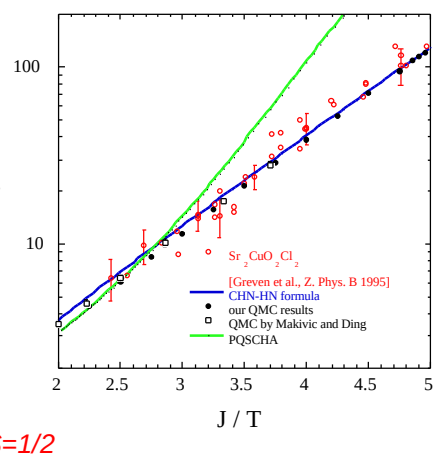
- Quantum field theory

$$\frac{\xi}{a} = \frac{e}{8}\frac{c/a}{2\pi\rho_s}\exp\left(\frac{2\pi\rho_s}{k_B T}\right)\left[1 - \frac{1}{2}\cdot\frac{k_B T}{2\pi\rho_s} + O\left(\frac{k_B^2 T^2}{4\pi^2\rho_s^2}\right) + \ldots\right]$$

- Semi-classical PQSCHA
  - Cuccoli et al. , PRL 1996, PRB 1997
  - quantum effects renormalize the temperature

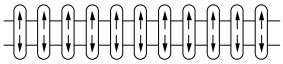
$$\xi = \xi_{classical}(f_S(T))$$

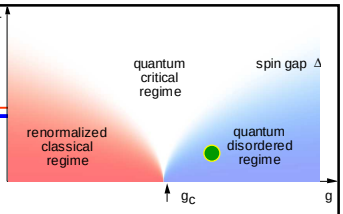
- *Quantum effects important for spin S=1/2*



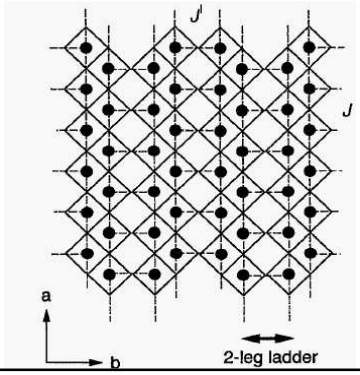
Spin ladder materials

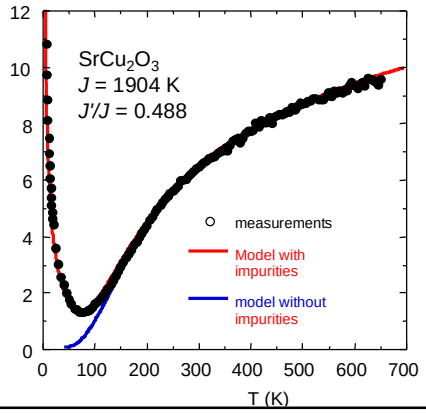
- are coupled dimers





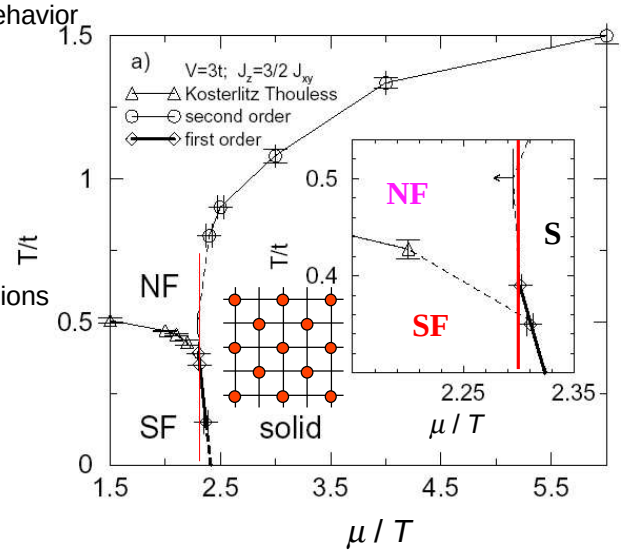
- Numerical simulations show that spin gap survives coupling, no ordering
- Quantitative fits of models to experimental measurements are possible





Phase diagram of 2D hard core bosons

- Unusual reentrant behavior
  - Normal fluid
  - Solid
  - Normal fluid
  - Superfluid
- Three melting transitions
  - Second order into normal fluid
  - First order into normal fluid
  - First order into superfluid



Classical Monte Carlo simulations

- We want to calculate a thermal average

$$\langle A \rangle = \sum_c A_c e^{-\beta E_c} / Z \quad \text{with} \quad Z = \sum_c e^{-\beta E_c}$$

- Exponentially large number of configurations
  $\Rightarrow$ draw a representative statistical sample by importance sampling
  - Pick M configurations c_i with probability $p_{c_i} = e^{-\beta E_{c_i}} / Z$
  - Calculate statistical average $\langle A \rangle \approx \overline{A} = \frac{1}{M} \sum_{i=1}^{M} A_{c_i}$
  - Within a statistical error $\Delta A = \sqrt{\frac{\mathrm{Var}\, A}{M}}$
- Problem: we cannot calculate $p_{c_i} = e^{-\beta E_{c_i}} / Z$ since we do not know Z

Markov chains and Metropolis algorithm

- Metropolis algorithm builds a Markov chain

$$c_1 \to c_2 \to \ldots \to c_i \to c_{i+1} \to \ldots$$

- Transition probabilities $W_{x,y}$ for transition $x \to y$ need to fulfill
  - Ergodicity: any configuration reachable from any other

$$\forall x,y \; \exists n \; : \; \left(W^n\right)_{x,y} \neq 0$$

  - Detailed balance:

$$\frac{W_{x,y}}{W_{y,x}} = \frac{p_y}{p_x}$$

- Simplest algorithm due to Metropolis (1953): $W_{x,y} = \min[1, p_y / p_x]$
- Needs only relative probabilities (energy differences) $p_y / p_x = e^{-\beta(E_y - E_x)}$

Critical slowing down of local updates

- Autocorrelations in Markov chain $c_1 \to c_2 \to \ldots \to c_i \to c_{i+1} \to \ldots$ change error estimate:

$$\Delta A = \sqrt{\left\langle \left(\overline{A} - \langle A \rangle\right)^2 \right\rangle} = \sqrt{\frac{\mathrm{Var}\, A}{M}(1 + 2\tau_A)}$$

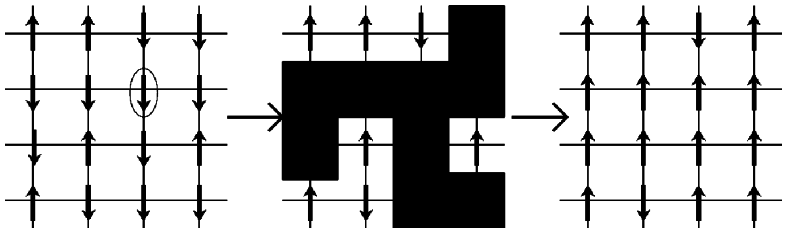
- Autocorrelation time τ_A diverges at criticality (critical slowing down)

$$\tau \propto \min[\xi, L]^z$$

- Local spin flips: $z \approx 2 \Rightarrow$ effort increased by L^2
- Advantage of Monte Carlo simulations
  - Can change the dynamics!
  - Dynamical exponent is non-universal
  - Need improved dynamics that change the system on length scale ξ

Cluster Algorithms: the simple explanation

- For each spin we ask:
  "do we want to change its alignment with its neigbors?"
  - Antiparallel spins gain energy ⇒ can change it
  - Parallel spins lose energy
    - Accepted only with probability $\exp(-2\beta J)$ (introduce a domain wall)
    - Otherwise we also have to flip its neighbor
      ⇒ we add the neighbor to the cluster with probability $1-\exp(-2\beta J)$



- Can be extended to cluster representation of the partition function

Cluster algorithms: the formal explanation

- We extend the phase space
  - From configurations C to configurations + graphs (C,G)

$$Z=\sum_C W(C)=\sum_C\sum_G W(C,G) \text{ with } W(C)=\sum_G W(C,G)$$

- Choose graph weights independent of configuration

$$W(C,G)=\Delta(C,G)V(G) \text{ where } \Delta(C,G)=\begin{cases}1 & \text{graph } G \text{ allowed for } C\\ 0 & \text{otherwise}\end{cases}$$

- Perform updates

2. Discard configuration — 4. Discard graph

$$C_i \rightarrow (C_i,G) \rightarrow G \rightarrow (C_{i+1},G) \rightarrow C_{i+1}$$

1. Pick a graph G $P[G]=\frac{V(G)}{W(C)}$ — 3. Pick any allowed new configuration

- Detailed balance

$$\frac{P[(C_i,G)\rightarrow(C_{i+1},G)]}{P[(C_{i+1},G)\rightarrow(C_i,G)]}\frac{1/\mathrm{N_C}}{1/\mathrm{N_C}}=1=\frac{\Delta(C_{i+1},G)V(G)}{\Delta(C_i,G)V(G)}=\frac{P[(C_{i+1},G)]}{P[(C_i,G)]}$$

Cluster algorithms: Ising model

- We need to find $\Delta(C,G)$ and $V(G)$ to fulfill $W(C) = \sum_G W(C,G) = \sum_G \Delta(C,G) V(G)$

$\Delta(C,G)$	[redacted]	[redacted]	$W(C)$
$\uparrow\uparrow, \downarrow\downarrow$	1	1	$e^{+\beta J}$
$\uparrow\downarrow, \downarrow\uparrow$	0	1	$e^{-\beta J}$
$W(G)$	$e^{+\beta J} - e^{-\beta J}$	$e^{-\beta J}$	

- This means for: $C_i \rightarrow (C_i, G) \rightarrow G$
  - Parallel spins: pick connected graph [redacted] with $P([redacted]) = \frac{e^{+\beta J} + e^{-\beta J}}{e^{+\beta J}} = 1 - e^{-2\beta J}$
  - Antiparallel spins: always pick open graph [redacted]
- And for: $G \rightarrow (C_{i+1}, G) \rightarrow C_{i+1}$
  - Configuration must be allowed $\Rightarrow$ connected spins must pe parallel
  - $\Rightarrow$ connected spins flipped as one cluster

Quantum Monte Carlo

- Not as easy as classical Monte Carlo

$$Z = \sum_c e^{-E_c / k_B T}$$

- Calculating the energy eigenvalue E_c is equivalent to solving the problem
- Need to find a mapping of the quantum partition function to a classical problem

$$Z = \mathrm{Tr}\, e^{-\beta H} \equiv \sum_c p_c$$

- Two approaches
  - Path integrals (time-dependent perturbation theory in imaginary time)
  - Series Expansion
- Sign problem if some $p_c < 0$

Path Integral Representation

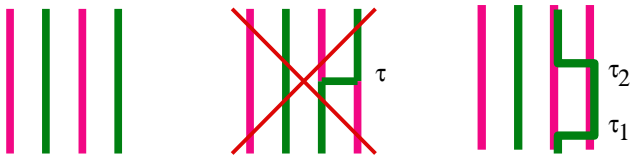
- interaction representation

$$H = H_0 + V, \quad H_0 = \sum_{<i,j>} J_{ij}^z S_i^z S_j^z - \sum_i h S_i^z, \quad V = \sum_{<i,j>} J_{ij}^{xy} (S_i^x S_j^x + S_i^y S_j^y)$$

$$Z = \mathrm{Tr}(e^{-\beta H}) = \mathrm{Tr}(e^{-\beta H_0} \mathcal{T} e^{-\int_0^\beta d\tau V(\tau)})$$

$$Z = \mathrm{Tr}(e^{-\beta H_0} (1 - \int_0^\beta d\tau V(\tau) + \int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 V(\tau_1) V(\tau_2) + \ldots))$$

- each term is represented by a world line configuration



- Advantage: diagonal terms treated exactly
- Disadvantage: keeping tracks of times can be computationally expensive

Series Expansion Representation (SSE)

- based on high temperature expansion, developed by A. Sandvik

$$Z = \mathrm{Tr}(e^{-\beta H}) = \sum_{n=0}^{\infty} (-\beta)^n \mathrm{Tr}(H^n)$$

$$= \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{|\alpha\rangle} \sum_{(b_1,\ldots,b_n)} \langle \alpha | \prod_{i=1}^{n} (-H_{b_i}) | \alpha \rangle$$

$$\text{with } H = \sum_i H_i$$

- Similar world line representation but without times assigned

- Advantage: easier calculations, no imaginary times needed
- Disadvantage: diagonal terms treated perturbatively

More technical details

- Example: 1D Heisenberg AFM $H = J\sum_i \vec{S}_i \vec{S}_{i+1} = \sum_i \left[J S_i^z S_{i+1}^z + \frac{J}{2}\left(S_i^+ S_{i+1}^- - S_i^+ S_{i+1}^-\right)\right]$
- Make all matrix elements of $-H$ positive
  - Add offset to diagonal elements
    $-H_{(i,d)} = -J S_i^z S_{i+1}^z + \frac{J}{4}$
  - Gauge transformation on one sublattice for offdiagonal operators
    $-H_{(i,o)} = -J\left(S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+\right) \xrightarrow{S_i^\pm \to (-1)^i S_i^\pm} +J\left(S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+\right)$
- Extend operator string to fixed length Λ by adding extra unit operators:
  $-H_0 = 1$ n … number of non-unit operators

$$Z = \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{|\alpha\rangle} \sum_{(b_1,\ldots,b_n)} \langle\alpha| \prod_{i=1}^{n} (-H_{b_i}) |\alpha\rangle \approx \sum_{n=0}^{\Lambda} \sum_{|\alpha\rangle} \sum_{(b_1,\ldots,b_\Lambda)} \frac{(\Lambda-n)!\beta^n}{\Lambda!} \langle\alpha| \prod_{i=1}^{\Lambda} (-H_{b_i}) |\alpha\rangle$$

Negative Sign Problem

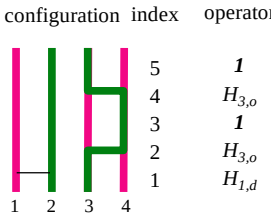
- All the stochastic methods map quantum to classical system
  $\langle A \rangle = \mathrm{Tr}[A \exp(-\beta H)] / \mathrm{Tr}[\exp(-\beta H)] = \sum_i A_i p_i \Big/ \sum_i p_i$
- Sign problem if one of the $p_i < 0$ and cannot be gauged away
  - Occurs in fermionic and frustrated problems

$$\langle A \rangle = \sum_i A_i p_i \Big/ \sum_i p_i = \frac{\sum_i A_i \,\mathrm{sgn}\, p_i |p_i| \Big/ \sum_i |p_i|}{\sum_i \mathrm{sgn}\, p_i |p_i| \Big/ \sum_i |p_i|} \equiv \frac{\langle A \cdot sign \rangle_{|p|}}{\langle sign \rangle_{|p|}}$$

- Exponentially increasing cancellation problem in sign
  $\langle A \cdot sign \rangle_{|p|} \approx \langle sign \rangle_{|p|} \approx \exp(-c\beta N) \Rightarrow \Delta A \approx \exp(+c\beta N)$
- In general harder than NP [MT and U.-J. Wiese]
  - But meron algorithm can solve sign problem in some cases

Step 1: Local diagonal updates

- Diagonal operators can be inserted and removed anywhere
  - In SSE:



$$P[1 \to H_{i,d}] = \min\left(1, \frac{\beta N_{bonds} \langle \alpha | H_{i,d} | \alpha \rangle}{L-n}\right)$$

$$P[H_{i,d} \to 1] = \min\left(1, \frac{L-n+1}{\beta N_{bonds} \langle \alpha | H_{i,d} | \alpha \rangle}\right)$$

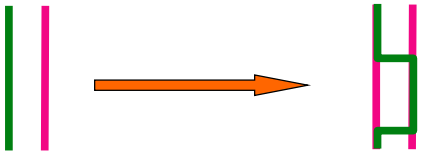
  - In path integrals:
    infinitesimal probabilities per time step $d\tau$:

$$P_{insertion} = H_{i,d} d\tau$$

    like radioactive decay process, we determine a decay time where the diagonal operator is inserted

Step 2a: Offdiagonal updates (local)

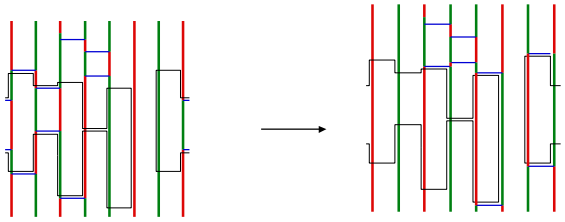
- Are very easy, can be done in any representation



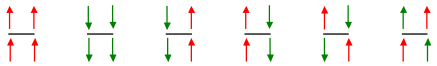
- Problem 1: only local changes
  - **Nonergodic**
  - No change of magnetization, particle number, winding number
- Problem 2:
  - **Critical slowing down**
- Solution: cluster algorithms

Offdiagonal loop-cluster updates

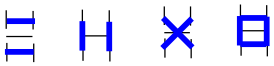
- Updates form closed loops since world lines may not be broken



- Introduce graph-representation
  - Connected spins must be flipped together
  - Local configurations:



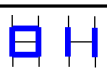
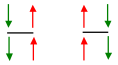
  - Local graphs:



Cluster building rules: Heisenberg antiferromagnet

- Example: XY-like AFM:

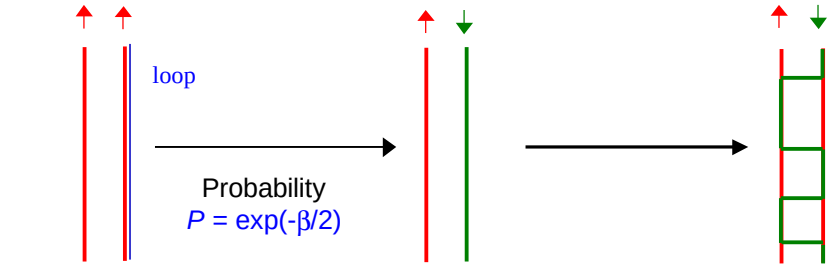
$$W(C) = \sum_G W(C,G) = \sum_G \Delta(C,G) V(G)$$

$\Delta(C,G)$				$W(C)$
	1	0	0	$J_z/2$
	0	0	0	0
	1	1	0	$J_{xy}/2$
$W(G)$	$J_z/2$	$(J_{xy}-J_z)/2$	0	

- Connected spins form a cluster and have to be flipped together
- Very simple and deterministic for Heisenberg model

Loop algorithm in a magnetic field

- Loop algorithm requires spin inversion symmetry
  - Magnetic field implemented by a-posteriori acceptance rate
- Example: spin dimer at $J = h = 1$



Triplet
$E = J/4 - h = -3/4$

$E = -J/4 = -1/4$

Singlet
$E = -3J/4 = -3/4$

Loop algorithm must go through high energy intermediate state

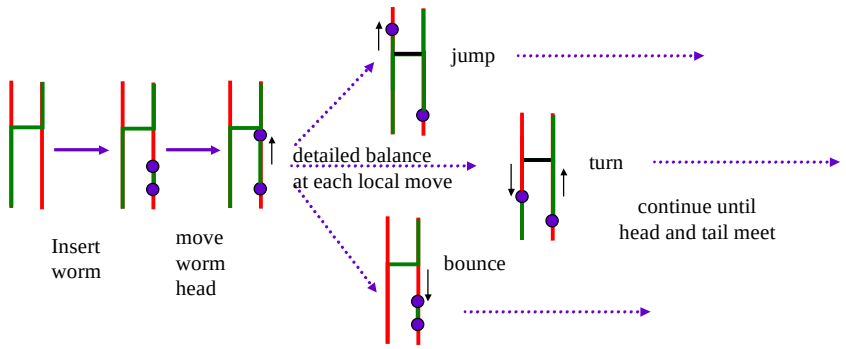
Exponential slowdown

Worm updates

- Prokof'ev *et al* 1997 (path integrals), Sandvik 1999 (SSE)
- Insert pair of creation/annihilation operators

$$H \leftarrow H + \eta(S^+ + S^-)$$

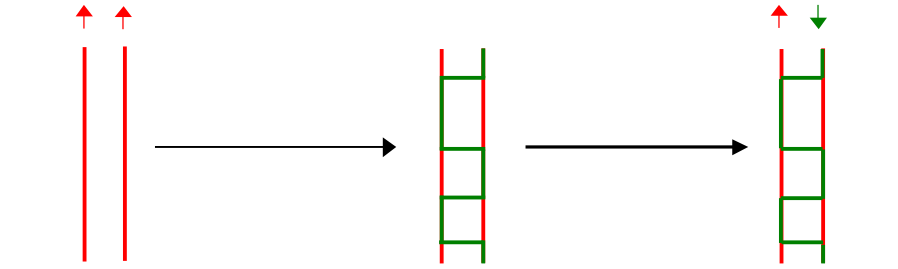
  - move these operators (worm head/tail) using local moves
  - when head and tail meet -> have created a loop, update is finished



- If bounce path can be eliminated ⇒ loop algorithm with pre-chosen paths

Worm algorithm in a magnetic field

- Worm algorithm performs a random walk
  - Change of configuration done in small steps
- Example: spin dimer at $J = h = 1$



Triplet
$E = J/4 - h = -3/4$

Singlet
$E = -3J/4 = -3/4$

No high energy intermediate state

Efficient update in presence of a magnetic field

When to use which ?

- All algorithms possible in both representations
- SSE usually preferred (computationally simpler)
- Use path integrals only if
  - Large diagonal matrix elements
  - Time-dependent interactions (dissipation or DMFT)
- Use loops if possible
- Use worms if
  - Loops not possible
  - Loops have small acceptance rates
- Examples:

	SSE	Path Integrals
Loops	Spin models	Spin models + dissipation
Worms	Spin models + magnetic field	Bose-Hubbard models

Further problems for Monte Carlo Methods

1. *Critical slowing down* at second order phase transitions
   Solved by Swendsen-Wang cluster algorithm

2. *Tunneling* through *free energy barriers* at first order transitions and for disordered systems

3. Calculation of the *free energy* F and entropy S
   Metropolis algorithm only give transition probabilities

$$W[i \to j] = \min\left[1, \frac{p_j}{p_i}\right]$$

   Normalization factor (partition function Z) unknown

$$p_i = \exp(-E/k_B T)/Z \qquad F = -k_B T \ln Z$$

Wang-Landau sampling solves all three problems

The Wang-Landau method

- Directly calculates density of states $\rho(E)$ instead of canonical average
  - Acceptance rate proportional to inverse density of states

$$p_i = \frac{1}{\rho(E_i)} \text{ instead of } p_i = e^{-E_i/k_B T}$$

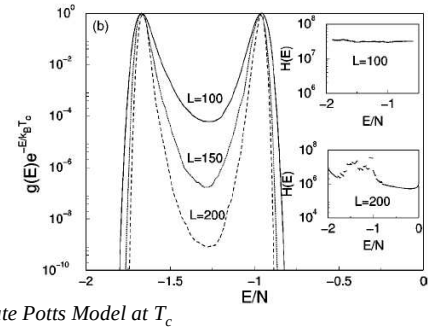
- Random walk in energy space
  - Flat histogram in energy space => *no tunneling problem*

- *Free energy accessible*

$$F = -k_B T \ln \sum_E \rho(E) e^{-E/k_B T}$$

- One simulation gives results for all T

- F. Wang & D.P. Landau
  PRL **86**, 2050 (2001),
  PRE **64**, 056101 (2001)



10-state Potts Model at T_c

Details of the Wang-Landau method

- Initially $\rho(E)$ is unknown
  - Start with $\rho(E)$=1 and adjust iteratively
- Only a few modifications to usual sampling needed

  Start with modification factor f=1
  do {
    do {
      Metropolis updates with transition probability W[i->j] = min[1,$\rho(E_i)$ / $\rho(E_j)$]
      Adjust $\rho(E)$ at each step: $\rho(E)$ <– $\rho(E)$ x exp(f)
    } until histogram $H(E)$ is "flat"
    decrease f <– f / 2
  } until $f \lesssim 10^{-8}$
- Comments
  - Initially: multiplicative changes with large f allow rapid crude convergence
  - Finally: small f means no systematic errors, detailed balance

Generalizations and Applications

- 1st and 2nd order transitions and disorder
  - F. Wang and D.P. Landau, PRL **86**, 2050 (2001); PRE **64**, 056101 (2001)
- Improved sampling
  - B.J. Schulz *et al.*, Int. J. Mod. Phys. C **13**, 477 (2002).
  - C. Yamaguchi and Naoki Kawashima, PRE **65,** 056710 (2002).
- Proteins
  - N. Rathore *et al.*, J. Chem. Phys. **116**, 7225 (2002).
- Polymer films
  - T.S. Jain *et al.*, J. Chem. Phys. **116**, 7238 (2002).
- Continuum simulations
  - Q. Yan *et al.* J. Chem. Phys. **116**, 8745 (2002).
  - M.S. Shell *et al.*, cond-mat/0206461
- Potts model
  - J. Phys. A**34**, 8781 (2001)
- Reaction coordinates
  - F. Calvo, Report cond-mat/0205428.
- Quantum problems
  - M. Troyer, S. Wessel and F. Alet, cond-mat/0207138

Quantum version - temperature expansion

- Density of states not accessible directly in quantum case
  - Classical $Z = \sum_c e^{-E_c/k_B T} = \sum_E \rho(E) e^{-E/k_B T}$
  - Quantum $Z = \mathrm{Tr}\, e^{-\beta H}$
- Our method based on high temperature expansion (SSE method by Sandvik)

$$Z = \mathrm{Tr}(e^{-\beta H}) = \sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} \mathrm{Tr}(H^n) = \sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} \sum_{|\alpha\rangle} \sum_{(b_1,\ldots,b_n)} \langle \alpha | \prod_{i=1}^{n} H_{b_i} | \alpha \rangle \qquad \text{with } H = \sum_i H_i$$

$$= \sum_{n=0}^{\infty} \frac{\beta^n}{n!} g(n) \approx \sum_{n=0}^{\Lambda} \frac{\beta^n}{n!} g(n)$$

- We ***sample*** all *orders*, *basis states* and *operator sequences*
  - Acceptance rates proportional to $\langle \alpha | \prod_{i=1}^{n} H_{b_i} | \alpha \rangle \Big/ g(n)$
  - Series coefficients to arbitrary order Λ
  - Flat histogram in order $n \propto E\, k_B T$

Quantum version – perturbation expansion

- Instead of temperature a coupling constant can be changed
- Based on finite temperature perturbation expansion

$$Z = \mathrm{Tr}(e^{-\beta H}) = \sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} \sum_{|\alpha\rangle} \sum_{(b_1,\ldots,b_n)} \langle \alpha | \prod_{i=1}^{n} H_{b_i} | \alpha \rangle \lambda^{n_\lambda(b_1,\ldots,b_n)}$$

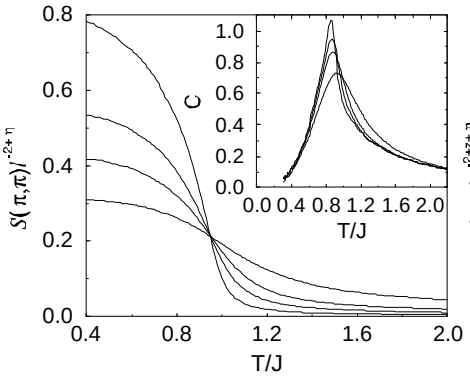
$$= \sum_{n_\lambda=0}^{\infty} \lambda^{n_\lambda} g(n_\lambda) \qquad \text{with } H = H_0 + \lambda V = \sum_i \lambda^{n_\lambda(i)} H_i$$

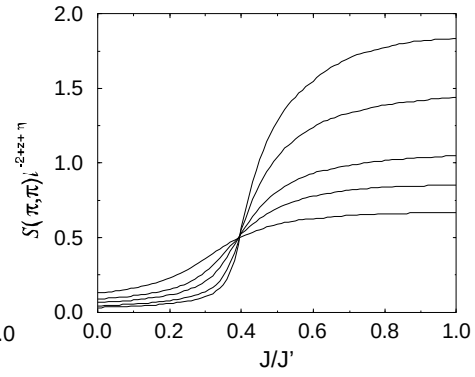
$n_\lambda(b_1,\ldots,b_n)$ counts # of λ terms

- Update probabilities proportional to $\frac{(-\beta)^n}{n!} \langle \alpha | \prod_{i=1}^{n} H_{b_i} | \alpha \rangle \Big/ g(n_\lambda)$
- Series coefficients to arbitrary order Λ
- Flat histogram in order n_λ of perturbation expansion

Thermal and quantum phase transition

- Thermal phase transition in 3D quantum Heisenberg antiferromagnet
- Quantum phase transition in bilayer quantum Heisenberg antiferromagnet





- critical temperature determined to three significant digits in one weekend on a PC

Summary

- Classical cluster algorithms
  - Have enabled high accuracy simulations of classical critical systems
- Quantum cluster algorithms
  - Enable similar accuracy for quantum systems
  - Quantum phase transitions
  - Finite temperature phase transitions
- Sign problem
  - No problem for nonfrustrated bosons and spins
  - Prohibits simulations for fermions and frustrated systems
  - In general harder than NP
- Wang Landau sampling
  - Allows calculation of free energy
  - Tunneling through free energy barriers
  - Effective for 1st order transitions and disordered quantum systems