

Introduction to Stochastic Series Expansion (SSE) Quantum Monte Carlo (QMC)

Stephan Humeniuk
ICFO - The Institute of Photonic Sciences

UPC, 24th of January 2013

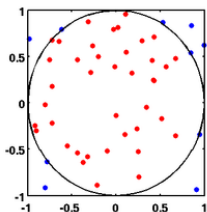
Introduction to Stochastic Series Expansion (SSE)

Quantum Monte Carlo (QMC)

- 1 Classical Monte Carlo: Ergodicity and detailed balance
- 2 Quantum-to-classical mapping:
 - 1.1 SSE representation of the partition sum
 - 2.2 Comparison between SSE and Path Integral representation
- 3 Update schemes
 - 1.1 Diagonal and off-diagonal update
 - 2.2 Directed loop algorithm: detailed balance equations
- 4 SSE estimators for
 - 1.1 Energy, magnetization, ...
 - 2.2 Superfluid fraction
 - 3.3 Extended ensembles and off-diagonal correlation functions
- 5 SSE for long-range transverse Ising systems
- 6 The sign problem

What is Monte Carlo ?

- ▶ An efficient method for calculating **high-dim. integrals** ...



Number of sampling points M ,
systematic error ϵ

- ▶ Riemann integration: $\epsilon \approx M^{-k/d}$
- ▶ MC sampling: $\epsilon \approx \frac{1}{\sqrt{M}}$ independent of spatial dimension (CLT)

- ▶ ... in particular expectation values in statistical physics:

$$\langle f \rangle_P = \int d^{3N} \vec{x} \int d^{3N} \vec{p} \quad P(\vec{x}, \vec{p}) f(\vec{x}, \vec{p}).$$

- ▶ In quantum statistical physics there are **many variants**:

- ▶ Path Integral MC, Determinantal MC, **Stochastic Series Expansion MC**

→ stochastic sampling of the **partition function**

- ▶ (fixed node) diffusion MC, projector MC, variational MC

→ based on the **wave function**

- ▶ diagrammatic MC (for fermions!) → stochastic sampling of Feynman diagrams

The Metropolis Algorithm (1953)

THE JOURNAL OF CHEMICAL PHYSICS

VOLUME 21, NUMBER 6

JUNE, 1953

Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

AND

EDWARD TELLER,* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.

I. INTRODUCTION

THE purpose of this paper is to describe a general method, suitable for fast electronic computing machines, of calculating the properties of any substance which may be considered as composed of interacting individual molecules. Classical statistics is assumed,

II. THE GENERAL METHOD FOR AN ARBITRARY POTENTIAL BETWEEN THE PARTICLES

In order to reduce the problem to a feasible size for numerical work, we can, of course, consider only a finite number of particles. This number N may be as high as several hundred. Our system consists of a square† con-

Classical Monte Carlo

$$\langle A \rangle = \frac{\sum_{\mu} A_{\mu} e^{-\beta E_{\mu}}}{\sum_{\mu} e^{-\beta E_{\mu}}}$$

Naive approach: Generate configurations μ randomly, compute A_{μ} and its Boltzmann weight, and then sum.

Problem: Most states will have vanishing weight.

► Importance sampling

Do not pick states from a uniform distribution, but instead perform a **guided random walk** in configuration space **that visits each state as often as corresponds to its weight**, i.e. $p_{\nu} = Z^{-1} e^{-\beta E_{\nu}}$.

Then the expectation values are simple averages:

$$A_M = \frac{1}{M} \sum_{i=1}^M A_{\mu_i} \rightarrow \langle A \rangle, M \rightarrow \infty$$

The random walk with transition probability $P(\nu \rightarrow \mu)$ must obey

- 1 **ergodicity**: Any state can be reached from any other state with non-vanishing probability.
- 2 **detailed balance** w.r.t. the desired probability distribution $\{p_\mu\}$: balance of fluxes:

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

detailed balance (stricter):

$$p_{\mu} P(\mu \rightarrow \nu) = p_{\nu} P(\nu \rightarrow \mu)$$

Quantum-to-classical mapping

Every D-dimensional quantum systems corresponds to a (D+1)-dimensional effective classical system.

$$\langle \hat{A} \rangle = \text{Tr}[\rho \hat{A}] = \frac{1}{Z} \sum_{|\alpha\rangle} \langle \alpha | e^{-\beta \hat{H}} \hat{A} | \alpha \rangle$$

Note: The eigenenergies are not known and one needs to expand the expression in a suitable way.

SSE representation (Taylor exp.)

$$\langle \hat{A} \rangle = \frac{1}{Z} \sum_{|\alpha\rangle} \langle \alpha | \sum_{n=0}^{\infty} \frac{(-\beta \hat{H})^n}{n!} \hat{A} | \alpha \rangle$$

Determinantal QMC
(Hubbard-Stratonovich)

$$Z_{HS} = \sum_{\substack{i,r \\ \text{auxiliary} \\ \text{Ising field}}} \det M_{\uparrow} \cdot \det M_{\downarrow}$$

Path integral representation (Trotter-Suzuki)

$$Z_{TS} = \sum_{m_1 \dots m_{2L}} \langle m_1 | e^{-\Delta\tau H_{\text{odd}}} | m_{2L} \rangle \langle m_{2L} | e^{-\Delta\tau H_{\text{even}}} | m_{2L-1} \rangle \\ \dots \langle m_3 | e^{-\Delta\tau H_{\text{odd}}} | m_2 \rangle \langle m_2 | e^{-\Delta\tau H_{\text{even}}} | m_1 \rangle$$

$$\langle A \rangle = \sum_C w(C) A(C)$$

classical weights

SSE representation for the spin $\frac{1}{2}$ XXZ model

$$\langle A \rangle = \frac{1}{Z} \text{Tr}[e^{-\beta H} A] = \frac{1}{Z} \sum_{|\alpha\rangle} \langle \alpha | \sum_{n=0}^{\infty} \frac{(-\beta H)^n}{n!} A | \alpha \rangle$$

$$H_{\text{XXZ}} = -J \sum_{\langle ij \rangle} \left\{ \frac{1}{2} (S_i^+ S_j^- + S_j^+ S_i^-) + \Delta S_i^z S_j^z \right\} - h \sum_i S_i^z$$

Decompose into diagonal (D) and off-diagonal (oD) bond operators:

$$H = - \sum_{b=1}^{N_{\text{bonds}}} H_b = +J \sum_{b=1}^{N_{\text{bonds}}} (H_{D,b} + H_{oD,b})$$

Multiplying out the n th power, we obtain

$$H^n = \sum_{\{S_n\}} \prod_{i=1}^n H_{t_i, b_i}$$

where the indices t_i (=operator type) and b_i (=bond index) are drawn from an **operator string** $S_n = \{[t_1, b_1], [t_2, b_2], \dots, [t_n, b_n]\}$. Then:

$$\langle A \rangle = \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{|\alpha\rangle} \sum_{\{S_n\}} \langle \alpha | \prod_{i=1}^n H_{t_i, b_i} A | \alpha \rangle = \sum_{n=0}^{\infty} \sum_{|\alpha\rangle} \sum_{\{S_n\}} w(\alpha, S_n) A(\alpha, S_n),$$

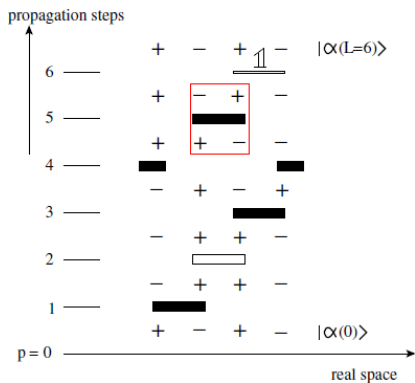
which is a sum over classical weights.

SSE simulation cell for the spin $\frac{1}{2}$ XXZ model

Note that there is no branching:

$$H_{t_i, b_i} |\alpha(p)\rangle = |\alpha(p+1)\rangle,$$

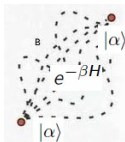
i.e. all $|\alpha(p)\rangle$ are basis states and no superpositions are created.



- ▶ One SSE configuration is specified by an initial state and an operator string $(|\alpha\rangle, S_n)$.
- ▶ Periodic boundary conditions in imaginary time due to the trace structure of the partition sum.
- ▶ MC update consists in exchanging operators: **diagonal** and **off-diagonal update**.
- ▶ For convenience we truncate the expansion order to $n_{max} = M$ and fill smaller expansion orders up with identity operators.

Path integral formulation of Z

- ▶ is an integral over different closed propagation paths in imaginary time.
- ▶ The quantum operator driving the propagation is always the **same**, $e^{-\beta H}$, so the integration runs over **initial** and **intermediate** states.
- ▶ “Schrödinger picture” of QM
- ▶ Trotter error



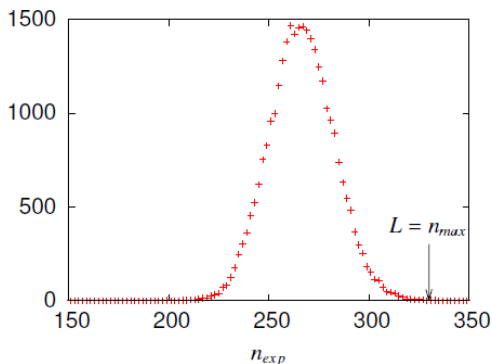
SSE formulation of Z

- ▶ is an integral over different closed propagation routes uniquely specified by an **operator string**. The integral runs over **initial/final** states of the propagation, and over the **operator string** driving the propagation.
- ▶ “Heisenberg picture” of QM
- ▶ No intrinsic approximation error

$$|\alpha\rangle = |\alpha(p=0)\rangle \rightarrow |\alpha(1)\rangle \rightarrow \dots \rightarrow |\alpha(p=L) = \alpha(0)\rangle$$

Comparison between SSE and PI

Figure 5.1: Histogram of the expansion orders n for a 1-dim XX model with 20 sites at temperature $T=0.05$. With the chosen cut-off L , which is determined during thermalization, the fixed length scheme does not lead to any approximation.



- ▶ The distribution of expansion orders shows that there is no intrinsic approximation involved in SSE.
- ▶ There is a statistical correspondence between worldline configurations within SSE formulation and worldline configurations in PI. Imaginary-time intervals and propagation intervals tend to coincide for $\beta \rightarrow \infty$.

Diagonal update: $id \leftrightarrow D$

Exchange identity and diagonal operators with Metropolis acceptance probabilities

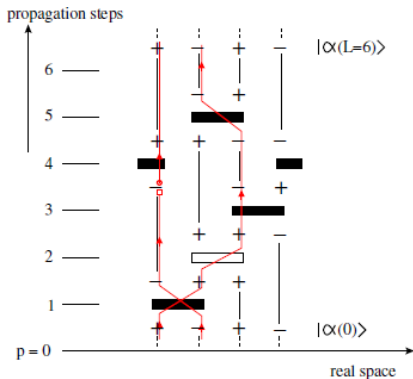
$$\begin{aligned} P_{\text{add}} = P([I, b]_p \rightarrow [D, b]_p) &= \min \left(1, \frac{N_{\text{bonds}} W(\dots [D, b]_p \dots; \alpha)}{W(\dots [I, b]_p \dots; \alpha)} \right) \\ &= \min \left(1, \frac{\beta N_{\text{bonds}}}{(M - n)} \cdot \langle \alpha(p) | H_{D,b} | \alpha(p) \rangle \right), \end{aligned}$$

$$\begin{aligned} P_{\text{remove}} = P([D, b]_p \rightarrow [I, b]_p) &= \min \left(1, \frac{W(\dots [I, b]_p \dots; \alpha)}{N_{\text{bonds}} W(\dots [D, b]_p \dots; \alpha)} \right) \\ &= \min \left(1, \frac{M - (n - 1)}{\beta N_{\text{bonds}}} \cdot \frac{1}{\langle \alpha(p) | H_{D,b} | \alpha(p) \rangle} \right) \end{aligned}$$

This changes the expansion order, which is related to the energy $\langle E \rangle = -\frac{\langle n \rangle}{\beta}$. The magnetization is not changed so that the diagonal update needs to be complemented by the off-diagonal update to satisfy ergodicity.

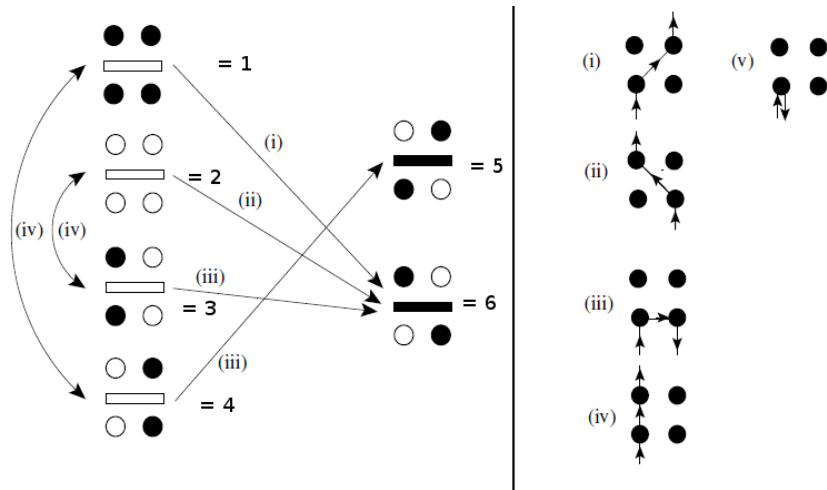
Off-diagonal update (“worm” or loop update): $D \leftrightarrow oD$

- ▶ Energy remains fixed, grand-canonical moves



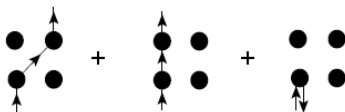
- ▶ The worm travels on the linked list flipping spins as it goes and thereby converting diagonal into off-diagonal operators and vice versa ($D \leftrightarrow oD$).
- ▶ It has to close on itself. This ensures that the replacements $D \leftrightarrow oD$ occur an **even number** of times which implies that the **periodic boundary conditions in imaginary time** are preserved, i.e. a new configuration with non-vanishing weight is generated during the update.

Off-diagonal update (“worm” or loop update)



Directed loop equations

- ▶ Transition probabilities for the worm must sum to unity



$$p(1 \rightarrow 6) + p(1 \rightarrow 2) + \underbrace{p(1, b)}_{\text{bounce probability}} = 1$$

etc. for all independent transition processes

- ▶ Multiply with the weight of the initial vertex and introduce the notation

$$w(i \rightarrow j) = w(i)p(i \rightarrow j)$$

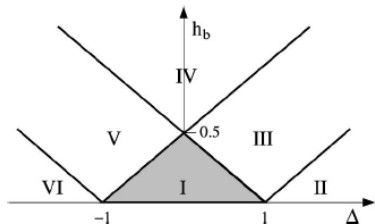
The detailed balance conditions take the simple form

$$w(i \rightarrow j) = w(j \rightarrow i)$$

and allow to identify different coefficients with each other.

- ▶ \Rightarrow (under-determined) set of equations which have to be solved
 - ▶ minimizing the bounce probabilities while
 - ▶ keeping all transition rates $w(i \rightarrow j)$ positive.

... in detail... for the spin $\frac{1}{2}$ XXZ model



$h_b = \frac{h}{zJ}$: magnetic field

Δ : spin space anisotropy parameter

region I: bounce free solution

$$a = \frac{1+\Delta}{4} + \frac{h_b}{2} + \frac{-b_1 - b_2 + b_3}{2},$$

$$b = \frac{1-\Delta}{4} - \frac{h_b}{2} + \frac{-b_1 + b_2 - b_3}{2},$$

$$c = \frac{\Delta-1}{4} + \frac{h_b}{2} + \epsilon + \frac{b_1 - b_2 - b_3}{2},$$

$$a' = \frac{1+\Delta}{4} - \frac{h_b}{2} + \frac{-b'_1 - b'_2 + b'_3}{2},$$

$$b' = \frac{1-\Delta}{4} + \frac{h_b}{2} + \frac{-b'_1 + b'_2 - b'_3}{2},$$

$$c' = \frac{\Delta-1}{4} + \frac{3h_b}{2} + \epsilon + \frac{b'_1 - b'_2 - b'_3}{2},$$

| | Bounce weights | | ϵ_{\min} |
|-----|-------------------------|--------------------------|----------------------|
| I | | | $(\Delta^- - h_b)/2$ |
| II | $b_2 = h_b - \Delta^-$ | $b'_2 = -h_b - \Delta^-$ | 0 |
| III | $b_2 = h_b - \Delta^-$ | | 0 |
| IV | $b_2 = h_b - \Delta^-$ | $b'_3 = h_b - \Delta^+$ | 0 |
| V | | $b'_3 = h_b - \Delta^+$ | $(\Delta^- - h_b)/2$ |
| VI | $b_3 = -h_b - \Delta^+$ | $b'_3 = h_b - \Delta^+$ | $-h_b - \Delta/2$ |

SSE estimators

- ▶ Energy

$$E = -\frac{\langle n \rangle}{\beta}, \quad \langle n \rangle : \text{average expansion order}$$

- ▶ Specific heat

$$C = [\langle n^2 \rangle - \langle n \rangle - \langle n \rangle^2]$$

This shows that the fluctuations of a quantity are not the same as the fluctuations of its estimator.

- ▶ Magnetization

$$m^z = \frac{1}{NL} \sum_{p=0}^{L-1} \langle S_r^z[p] \rangle$$

Due to the cyclic structure of the partition function one can average over propagation steps p to obtain more statistics.

- ▶ Helicity modulus, superfluid density

$$\Gamma_\alpha = \frac{k_B T}{2J^{xy}} L^{d-2} \langle w_\alpha^2 \rangle$$

where $w_\alpha = \sum_{b||\alpha} \left(\frac{N_b^+ - N_b^-}{L} \right)$ is the winding number for $\alpha = x, y, z$.

SSE estimators and extended ensemble techniques

- ▶ (Propagation-time) off-diagonal correlator: Ratio of a modified partition function and of the original partition function.

$$\langle \hat{S}_i^+[m] \hat{S}_{i+r}[0] \rangle = \frac{Z'}{Z}$$

Z' : modified partition function with two discontinuities of the worm ends at (i, m) and $(i + r, 0)$.

Transverse field Ising model with arbitrary interactions

- ▶ Hamiltonian

$$H = \sum_{ij} J_{ij} \sigma_i^z \sigma_j^z - h_x \sum_i \sigma_i^x$$

- ▶ J_{ij} arbitrary (long-range, frustrated, random)
- ▶ Define the bond operators in such a way that the state evolution is deterministic and all weights are positive:

$$H_{0,0} = 1$$

$$H_{i,0} = h(\sigma_i^+ + \sigma_i^-), \quad i > 0$$

$$H_{i,i} = h, \quad i > 0$$

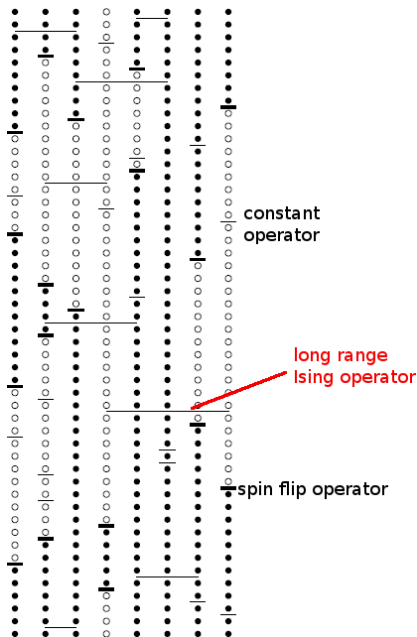
$$H_{i,j} = |J_{ij}| - J_{ij} \sigma_i^z \sigma_j^z, \quad i, j > 0, \quad i \neq j$$

- ▶ Observations: No loop update possible as there are no **off-diagonal pair interactions**.
Constants added in a clever way.

TFI model

Observations:

- ▶ The arbitrary-range interactions in space have been transformed into completely local constraints in imaginary time.
- ▶ Summing over all interactions requires $\approx N^2$ operations. Here the diagonal update at all positions in S_L requires $\approx L \ln(N) \approx \beta N \ln(N) I_N(J)$ operations.



The sign problem

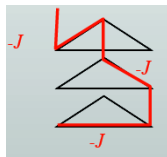
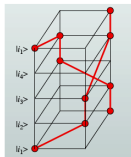
QMC cannot simulate

- ▶ fermions
- ▶ frustrated spin systems (i.e. AFM on non-bipartite lattices)

as the weights cannot be chosen to be positive definite, e.g.

$$H_{\text{XXZ}} = -\frac{1}{2} \underbrace{\sum_{\langle ij \rangle} (J^z S_i^z S_j^z + 2hS_i^z + C)}_{\geq 0} + \frac{1}{2} \sum_{\langle ij \rangle} \underbrace{J^{xy} (S_i^+ S_j^- + S_i^- S_j^+)}_{\text{sign problem}}$$

The sign problem affects only the **off-diagonal part** since here we cannot add a constant to make the matrix elements positive definite.



In a **bipartite lattice**, we need an off-diagonal bond operator to act an **even number** of times to restore the original configuration on $|\alpha\rangle$.

If the lattice is non-bipartite (e.g. triangular or $J_1 - J_2$ chain), there can be a product of an odd number of off-diagonal operators. Similarly for fermions.

The sign problem is NP-hard. [Troyer and Wiese, PRL 2005]

The sign problem

How does a negative sign in some configuration weights affect the QMC simulation ?

$$w(\alpha, S_L) \rightarrow |w(\alpha, S_L)|$$

Any average of an observable takes the form

$$\langle \hat{O} \rangle = \frac{\sum_{\alpha, S_L} O(\alpha, S_L) \text{sign}(\alpha, S_L) |w(\alpha, S_L)|}{\sum_{\alpha, S_L} \text{sign}(\alpha, S_L) |w(\alpha, S_L)|}$$

shift the sign from the weight onto the observable

$$\langle \hat{O} \rangle = \frac{\langle \text{sign}(\alpha, S_L) O(\alpha, S_L) \rangle_{|w|}}{\langle \text{sign}(\alpha, S_L) \rangle_{|w|}}$$

where $\langle \dots \rangle_{|w|}$ denotes the average $\frac{\sum_{\alpha, S_L} (\dots) |w(\alpha, S_L)|}{\sum_{\alpha, S_L} |w(\alpha, S_L)|}$.

The sign problem

In particular:

$$\begin{aligned}\langle \text{sign}(\alpha, S_L) \rangle_{|w|} &= \frac{\sum_{\alpha, S_L} w(\alpha, S_L)}{\sum_{\alpha, S_L} |w(\alpha, S_L)|} = \frac{Z_w}{Z_{|w|}} \\ &= \frac{e^{-\beta V f_w}}{e^{-\beta V f_{|w|}}} = \exp[-\beta V (f_w - f_{|w|})],\end{aligned}$$

where f_w and $f_{|w|}$ are the free energy densities of the systems with weight $w(\alpha, S_L)$ and $|w(\alpha, S_L)|$, respectively.

Now we have that

$$Z_w \leq Z_{|w|}$$

because $\sum_{\alpha, S_L} w(\alpha, S_L) \leq \sum_{\alpha, S_L} |w(\alpha, S_L)|$, and since $f = -\frac{1}{\beta V} \ln Z$,

$$f_w \geq f_{|w|}, \quad \Delta f = f_w - f_{|w|} \geq 0$$

so that $\langle \text{sign}(\alpha, S_L) \rangle_{|w|} = \exp(-\beta V \Delta f)$ is an **exponentially decreasing** quantity when $\beta, V \rightarrow \infty$.

The miserable signal-to-noise ratio of $\langle \text{sign} \rangle$ propagates on all the other estimates.

References

- ▶ [1] Phys. Rev. B 59, R14157 (1999)
- ▶ [2] Phys.Rev.E 68, 056701 (2003)