Introduction to quantum Monte Carlo: The Stochastic Series Expansion method

Anders Sandvik, Boston University

• Illustration of concept; classical Monte Carlo example
• Detailed account of SSE for the S=1/2 Heisenberg model

This presentation is based on material available at http://physics.bu.edu/~sandvik/programs/

A simple SSE program (Fortran90) for the 2D Heisenberg model can be downloaded from this site

References:
Quantum Monte Carlo for $S=1/2$ Heisenberg models

Classical Monte Carlo - we know the energies for all configurations

$$\langle A \rangle = \frac{1}{Z} \sum_n A_n e^{-E_n/T}, \quad Z = \sum_n e^{-E_n/T}$$

Quantum systems - we normally don't know the eigenstates/energies

- often the goal is to find the ground state $|0\rangle$
- for small $N (<25)$, we can diagonalize the Hamiltonian
- number of basis states $= 2^N$ (size of the matrix)
- block diagonalize using symmetries; up to $N=40$ possible
- not enough to study $N \to \infty$ limit

QMC based on quantum mechanical thermal expectation value

$$\langle A \rangle = \frac{1}{Z} \text{Tr}\{A e^{-H/T}\} = \frac{1}{Z} \sum_s \langle s | A e^{-H/T} | s \rangle$$

$$Z = \text{Tr}\{e^{-H/T}\} = \sum_s \langle s | e^{-H/T} | s \rangle$$

The trace can be written in any basis

- But the exponential operator is in general difficult to evaluate

Quantum Monte Carlo; rewrite the exponential operator

- Can be done in different ways; many different algorithms
Warm-up: SSE for a classical problem

Classical thermal expectation value

\[ \langle f \rangle = \frac{1}{Z} \sum_{\{\sigma\}} f(\sigma) e^{-\beta E(\sigma)}, \quad Z = \sum_{\{\sigma\}} e^{-\beta E(\sigma)} \]

Classical (e.g., Ising) spins: \( \sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_N\} \)

**Classical Monte Carlo:** Importance sampling of spin configurations

Probability of generating a configuration

\[ P(\sigma) = \frac{1}{Z} W(\sigma), \quad W(\sigma) = e^{-\beta E(\sigma)} \]

Estimate of expectation value based on sampled configurations

\[ \langle f \rangle = \langle f \rangle_W = \frac{1}{N_{\text{samples}}} \sum_i f(\sigma[i]) \]

Imagine that we are not able to evaluate the exponential function

How could we proceed then?
Use Taylor expansion of the exponential function

\[ \langle f \rangle = \frac{1}{Z} \sum_{\{\sigma\}} \sum_{n=0}^{\infty} f(\sigma) \frac{(-\beta E)^n}{n!}, \quad Z = \sum_{\{\sigma\}} \sum_{n=0}^{\infty} \frac{(-\beta E)^n}{n!} \]

Expansion power \( n \) is a new “dimension” of the configuration space
To ensure positive-definiteness we may have to shift \( E \) (must be < 0)

\[ E(\sigma) \rightarrow E(\sigma) - \epsilon \]

The sampling weight for the configurations \((\sigma,n)\) is

\[ W(\sigma, n) = \frac{\beta^n [\epsilon - E(\sigma)]^n}{n!} \]

The function to be averaged (estimator) \( f(\sigma) \) is the same as before; it does not depend on \( n \)

\[ \langle f \rangle = \langle f \rangle_W = \frac{1}{N_{\text{samples}}} \sum_{i} f(\sigma[i]) \]

However, if \( f(\sigma) \) is a function of the energy it can be rewritten as a function of \( n \) only!
Define: \( H(\sigma) = \epsilon - E(\sigma) \)

\[
\langle H \rangle = \frac{1}{Z} \sum_{\sigma, n} H(\sigma)W(\sigma, n), \quad Z = \sum_{\sigma, n} W(\sigma, n), \quad W(\sigma, n) = \frac{\beta^n H(\sigma)^n}{n!}
\]

Shift summation index: \( m = n + 1 \)

\[
\sum_{\sigma, n} H(\sigma)W(\sigma, n) = \sum_{\sigma, m} \frac{m}{\beta} W(\sigma, m)
\]

Therefore the energy expectation value is

\[
\langle H \rangle = \frac{1}{\beta} \langle n \rangle_W \quad \Rightarrow \quad E = \epsilon - \frac{1}{\beta} \langle n \rangle_W
\]

We can also easily obtain

\[
\langle H^2 \rangle = \frac{1}{\beta^2} \langle n(n - 1) \rangle_W
\]

And thus the specific heat \( C = \beta^{-1}(\langle E^2 \rangle - \langle E \rangle^2) \) is

\[
C = \frac{1}{\beta} \left( \langle n^2 \rangle - \langle n \rangle^2 - \langle n \rangle \right)
\]
What range of expansion orders $n$ is sampled?

From the preceding results we obtain

$$\langle n \rangle = \beta (\epsilon - E')$$

$$\langle n^2 \rangle - \langle n \rangle^2 = \beta (C + \epsilon - E')$$

Consider low $T$; $C \to 0$

$$\langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle$$

Thus, for a system with $N$ spins:

**Average expansion order** $\propto \beta N$

**Width of distribution** $\propto \sqrt{\beta N}$

These results hold true for quantum systems as well

$H^n$ requires more complicated treatment
Quantum-mechanical SSE

Thermal expectation value

\[ \langle A \rangle = \frac{1}{Z} \text{Tr}\{A e^{-\beta H}\}, \quad Z = \text{Tr}\{e^{-\beta H}\} \]

Choose a basis and Taylor expand the exponential operator

\[ Z = \sum_{\alpha} \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \langle \alpha | (-H)^n | \alpha \rangle \]

Write the hamiltonian as a sum of local operators

\[ H = - \sum_{a,b} H_{a,b} \quad a = \text{operator type (e.g., 1=diagonal, 2=off-diagonal)} \]

\[ b = \text{lattice unit (e.g., bond connecting sites i,j)} \]

such that for every a, b: \( H_{a,b} | \alpha \rangle = h_{a,b}(\alpha) | \alpha' \rangle \) (no branching)

Write the powers of H in terms of “strings” of these operators

\[ (-H)^n = \sum_{\{H_{a,b}\}} \prod_{p=1}^{n} H_{a(p),b(p)} \]
Operator strings of varying length $n$
• as in the classical case $\langle n \rangle = -\beta \langle H \rangle$

Fixed-length operator strings: introduce unit operator: $H_{0,0} = 1$

**Expansion cut-off $M$:** add $M-n$ unit operators to each string
• there are $M!/n!(M-n)!$ ways of doing this $\Rightarrow$

$$(-H)^n = \sum_{\{H_{ab}\}} \frac{(M-n)!n!}{M!} \prod_{p=1}^{M} H_{a(p),b(p)}$$

$n = \text{number of non-[0,0] operators}$

The truncation should not be considered an approximation
• $M$ can be chosen such that the truncation error is negligible

$$Z = \sum_{\alpha} \sum_{\{H_{ab}\}} \frac{\beta^n (M-n)!}{M!} \left\langle \alpha \left| \prod_{i=1}^{M} H_{a(i),b(i)} \right| \alpha \right\rangle$$

The terms $(\alpha, \{H_{ab}\})$ are sampled according to weight in this sum
• requires positive-definiteness
• to this end, a constant may have to be added to diagonal $H_{ab}$
• there can still be a “sign problem” arising from off-diagonal $H_{ab}$
SSE algorithm for the $S=1/2$ Heisenberg model

- The algorithm for this model is particularly simple and efficient

$$H = J \sum_{\langle i, j \rangle} S_i \cdot S_j$$

Consider **bipartite lattice** (sign problem for frustrated systems)

![Bipartite lattice](image)

- sublattice A
- sublattice B

Standard $z$-component basis:

$$|\alpha\rangle = |S_1^z, S_2^z, \ldots, S_N^z\rangle, \quad S_i^z = \pm \frac{1}{2}$$

Bond operators: bond $b$ connects sites $i(b), j(b)$

$$H = \sum_{b=1}^{B} \left[ S_{i(b)}^z S_{j(b)}^z + \frac{1}{2} \left( S_{i(b)}^+ S_{j(b)}^- + S_{i(b)}^- S_{j(b)}^+ \right) \right]$$
\[ H = - \sum_{b=1}^{B} \sum_{a=1}^{2} H_{a,b} \]

**Diagonal and off-diagonal** bond operators

\[ H_{1,b} = \frac{1}{4} - S_{i(b)}^{z} S_{j(b)}^{z}, \quad H_{2,b} = \frac{1}{2} \left( S_{i(b)}^{+} S_{j(b)}^{-} + S_{i(b)}^{-} S_{j(b)}^{+} \right) \]

A minus sign in front of the off-diagonal \( H_{2b} \) is neglected

- this corresponds to a **sublattice rotation**; 180 degree rotation in the xy-plane of the spin operators on sublattice B
- The **sign is irrelevant for a bipartite lattice** (will be shown later)

SSE operator string \( \prod_{p=1}^{M} H_{a(p),b(p)} \)

Represented in the computer program by

\[ \text{opstring}[p] = 2b(p) + a(p) - 1 \]

Spin state \( |\alpha> \) represented by

\[ \text{spin}[i] = 2S_{i}^{z} \]
SSE partition function

\[ Z = \sum_\alpha \sum_{\{H_{ab}\}} \frac{\beta^n (M - n)!}{M!} \left\langle \alpha \left| \prod_{i=1}^M H_{a(i),b(i)} \right| \alpha \right\rangle \]

Both \( H_{1b} \) and \( H_{2b} \) give 0 when acting on parallel spins
• non-zero matrix element = \( \frac{1}{2} \) in both cases

Define propagated states

\[ |\alpha(p)\rangle = \prod_{j=1}^p H_{a(j),b(j)} |\alpha\rangle \quad |\alpha\rangle = |\alpha(0)\rangle \]

For a contributing configuration: \( |\alpha(M)\rangle = |\alpha(0)\rangle \) (periodic)

The configuration weight is then

\[ W(\alpha, \{H_{ab}\}) = \left( \frac{\beta}{2} \right)^n \frac{(M - n)!}{M!} \]

Periodicity requires an even number of spin flips
• This is why the sign of \( H_{2b} \) is irrelevant for a bipartite lattice
• For a frustrated lattice an odd number of flips is possible
Graphical representation

- 1D example; 8 spins, M=12

1D: bond b connects sites b and b+1

spin[i] = +1 +1 -1 -1 +1 -1 +1 -1

<table>
<thead>
<tr>
<th>p</th>
<th>a(p)</th>
<th>b(p)</th>
<th>opstring[p]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>7</td>
<td>14</td>
</tr>
</tbody>
</table>
Linked-list representation

- vertex: operator and spins before and after the operator has acted

```
  i = 1 2 3 4 5 6 7 8
spin[i] = +1 +1 -1 -1 +1 -1 +1 -1
```

- replace spins between vertices by **links**

```
p  a(p)  b(p)  opstring[p]
1   1     2     4
2   0     0     0
3   2     4     9
4   2     6    13
5   1     3     6
6   0     0     0
7   0     0     0
8   1     2     4
9   2     6    13
10  0     0     0
11  2     4     9
12  1     7    14
```

- **linked vertex list** used in some parts of the program
A vertex has 4 “legs”, numbered \( l=0,1,2,3 \):

\[
\begin{array}{cccc}
0 & 1 & 0 & 1 \\
\circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ \\
2 & 3 & 2 & 3 \\
\end{array}
\]

position \( p \) of operator in operator string \( \text{opstring}[p] \), vertex leg \( l \)
\( \Rightarrow \) position \( v \) in linked vertex list: 
\[
v = 1 + l + 4 \times (p - 1)
\]

\( \text{vertexlist}[v] \) contains the element \( # \) to which \( v \) is linked

\[
\begin{array}{cccccc}
l = & 0 & 1 & 2 & 3 & p \\
 & [37] & 0 & [38] & 0 & [39] & 0 & [40] & 0 & 10 \\
\end{array}
\]
Sampling the SSE configurations; updates

1) **Diagonal update**
   - replace unit operator by diagonal operator, and vice versa
   \[ H_{0,0} \leftrightarrow H_{1,b} \]

2) **Off-diagonal update** (local or loop)
   - change the operator type, diagonal off-diagonal, for two (local) or several (loop) operators
   \[ \{ H_{a_1,b_1}, H_{a_2,b_2}, \ldots, H_{a_m,b_m} \} \leftrightarrow \{ H_{3-a_1,b_1}, H_{3-a_2,b_2}, \ldots, H_{3-a_m,b_m} \} \]

3) **Flip spins** in the state \(|\alpha\rangle\)
   - unconstrained “free” spins; weight unchanged after flip
   - only possible at high temperatures; strictly not necessary
   \[ S^z_i \rightarrow -S^z_i \]

Updates satisfy detailed balance:

\[
P_{\text{accept}}(A \rightarrow B) = \min \left( \frac{W(B)P_{\text{select}}(B \rightarrow A)}{W(A)P_{\text{select}}(A \rightarrow B)}, 1 \right)
\]
Diagonal update

- Carried out in `opstring[p]` for `p=1,...,M`
- State `|\alpha(p-1)>` stored in `spin[]`

**Insertion of diagonal operator**

<table>
<thead>
<tr>
<th>i = 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>spin[i] = -1 +1 -1 +1 +1 -1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Removal of diagonal operator**

<table>
<thead>
<tr>
<th>i = 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>spin[i] = -1 +1 -1 +1 +1 -1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Off-diagonal no change, propagate state**

<table>
<thead>
<tr>
<th>i = 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>spin[i] = -1 +1 +1 -1 +1 -1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Insertion of a diagonal operator if $\text{opstring}[p]=0$

Generate bond index $b$ at random, attempt $\text{opstring}[p]=2*b$

• can only be done if $\text{spin}[i(b)]\neq\text{spin}[j(b)]$
• $n$ increases by 1; weight ratio

$$\frac{W(n+1)}{W(n)} = \frac{\beta/2}{M-n}$$

Removal of a diagonal operator if $\text{opstring}[p]\neq0$

• $n$ decreases by 1; weight ratio

$$\frac{W(n-1)}{W(n)} = \frac{M-n+1}{\beta/2}$$

B ways of selecting $b$ but only one way of removing an operator;

$$\frac{P_{\text{select}}(b \rightarrow 0)}{P_{\text{select}}(0 \rightarrow b)} = B$$

Accept probabilities:

$$P_{\text{accept}}(n \rightarrow n+1) = \min\left(\frac{B\beta/2}{M-n}, 1\right)$$

$$P_{\text{accept}}(n \rightarrow n-1) = \min\left(\frac{M-n+1}{B\beta/2}, 1\right)$$
Local off-diagonal update (obsolete)

Change type of 2 operators on the same bond

- cannot always be done; check for constraining operators
- no weight change; accept with fixed probability (e.g., $P=1$)
**Note:** periodic boundary conditions in the “propagation” direction

- update spanning across the boundary affects the stored state $|\alpha>$

Local updates typically are **not very efficient**

- critical slowing-down
- no winding-number or particle-number fluctuations
**Loop update**

- carried out in the linked-vertex-list representation
- move “vertically” along links and “horizontally” on the same operator
- spins flipped at all vertex-legs visited; operator type changes; weight unchanged
- construct all loops, flip with probability 1/2 (as in Swendsen-Wang)

![Diagram showing loop update process](image)
**Monte Carlo step**
- a cycle of diagonal updates \((p=1, \ldots, M \text{ in } \text{opstring}[p])\)
- construction of the linked vertex list
- construct all loops, flip each with probability \(1/2\)
- map updated vertex list back to \(\text{opstring}[], \text{spin[]}\)

**Starting the simulation**
- “empty” operator string, \(\text{opstring}[p]=\emptyset, p=1, \ldots, M\)
- \(M\) is arbitrary, e.g., \(M=20\)
- random spin state; \(\text{spin}[p]=+1, -1\)

**Determining the cut-off \(M\)**
- after each, MC step, compare expansion order \(n\) with \(M\)
- if \(M-n<n/a\), with, e.g., \(a=3\), then \(M=n+n/a\)
Generalization of loop update; directed loops

In the case of the isotropic S=1/2 model
• There are only 4 non-0 vertices
• The operators uniquely define all loops
• Loops are non-self-intersecting

Directed loops
• In general, there are more than 4 allowed vertices
• A vertex is entered at some entrance leg
• The path can proceed (exit) through any of the 4 legs
• Exit probabilities are obtained from directed-loop equations
• Loops can back-track (“bounce”) and self-intersect
• Bounces can be avoided for some models (more efficient)

Can be used for spins, bsons, 1D fermions,…