Introduction to Bosonic QMC: World line and Stochastic Green Function algorithms

G. G. Batrouni (Nice, Singapore, Beijing)

Centre for Quantum **Technologies**

Outline

- Brief review of classical Monte Carlo
- World line algorithm
- Introduction to work-like algorithms

Preliminary Comments

- There is no "best" algorithm.
- Algorithms have advantages and disadvantages.
- Algorithms can be very well adapted to and optimized for a class of problems, but not work very well for other classes.
- The point is to choose an algorithm which is well adapted to your problem.
- Be pragmatic.

Here I will try to explain the basics building blocks of algorithms, but will not discuss important issues like auto-correlation, error bars etc. These are very interesting and important issues, but beyond our scope here.

Our main goal is to calculate averages:

$$
\langle \mathcal{A}[\{\phi\}]\rangle = \int D\phi \; \mathcal{A}[\{\phi\}] P[\{\phi\}]; \qquad P[\{\phi\}] = \frac{e^{-\beta E[\{\phi\}]}}{Z}
$$

P is the normalized (equilibrium) Boltzmann weight and Z the partition function. Complete numerical integration is impossible for large systems, even for a "simple" model like Ising where the number of configurations grows with the number of sites as 2^N .

Monte Carlo: Construct a Markov chain to generate the configurations such that in the stationary limit, configurations are given by the correct probability distribution: P. In a Markov chain, an event depends only on the event immediately before it.

The evolution equation for P can be written as:

- P[φ ,t] is the probability density for the configuration $\{\varphi\}$ at step t.
- \bullet $W_{\phi \to \phi^\prime} \;$ is the transition rate from ${\sf \small \textsf{p}}$ to ${\sf \small \textsf{p}}'$
- t is not necessarily real time.

In the stationary limit (**i.e. at equilibrium**):

$$
\sum_{\{\phi'\}} P[\phi'] W_{\phi' \to \phi} = \sum_{\{\phi'\}} P[\phi] W_{\phi \to \phi'}
$$

Not easy to implement!

Easier to implement the configuration-by-configuration condition:

 $P[\{\phi'\}|W_{\phi'\to\phi}=P[\{\phi\}|W_{\phi\to\phi'}]$

This condition is called **detailed balance**, it is sufficient to reach the correct equilibrium, but **not** necessary!

Many ways to do this. I give two examples:

Glauber dynamics:

 $W_{\phi \to \phi'} = \frac{P[\{\phi'\}]}{P[\{\phi'\}] + P[\{\phi\}]}$

Check detailed balance:

 $P[\{\phi'\}|W_{\phi'\to\phi}=P[\{\phi\}|W_{\phi\to\phi'}]$ $P[\{\phi'\}]\,W_{\phi'\to\phi} = \frac{P[\{\phi'\}]\,P[\{\phi\}]}{P[\{\phi\}]\,+\,P[\{\phi'\}]}$ $P[\{\phi\}]W_{\phi\to\phi'}=\frac{P[\{\phi\}]P[\{\phi'\}]}{P[\{\phi'\}]+P[\{\phi\}]}$

We see that detailed balance is indeed satisfied. Of course we have: $P[\{\phi\}] = \frac{1}{7}e^{-\beta H[\{\phi\}]}$

Z cancels out in the detailed balance condition: We do not need to know the partition function, we only need the exponential of the energy of the configuration. We have this!

Another very popular choice is Metropolis dynamics:

$$
W_{\phi_i \to \phi_f} = 1, \quad \text{if } P[\{\phi_f\}] \ge P[\{\phi_i\}]
$$

$$
W_{\phi_i \to \phi_f} = \frac{P[\{\phi_f\}]}{P[\{\phi_i\}]}, \quad \text{if } P[\{\phi_f\}] < P[\{\phi_i\}]
$$

Check detailed balance for transitions between two configurations φ and φ'. Two possibilities:

 $P[\{\phi'\}] > P[\{\phi\}]$ or $P[\{\phi'\}] < P[\{\phi\}]$ Take the first one. $P[\{\phi\}|W_{\phi\rightarrow\phi'}=P[\{\phi\}]$

$$
P[\{\phi'\}]W_{\phi'\to\phi} = P[\{\phi'\}\frac{P[\{\phi\}]}{P[\{\phi'\}]} = P[\{\phi\}]
$$

As before, we do not need Z. Choose the dynamics you want, iterate using the appropriate W.

Example: Ising with Metropolis.

$$
H[\{S\}] = -J \sum_{\langle i,j \rangle} S_i S_j, \quad S_\ell = \pm 1, \quad Z = \sum_{\{s\}} e^{-\beta H[\{S\}]}
$$

- Generate an arbitrary initial configuration of spins: $\{S\}_{init}$
- Visit the spins in any order (sequentially or randomly).
- At each spin, try to change it: $S_i \rightarrow -S_i$
- Calculate the ratio: $\mathcal{R} = P[\{S'\}]/P[\{S\}]$

•
$$
\mathcal{R} = \frac{e^{-\beta JS_i \sum_j' S_j}}{e^{\beta JS_i \sum_j' S_j}} = e^{-2\beta JS_i \sum_j' S_j}
$$

• Accept the new value of the spin and move to the next one if: $\mathcal{R} \geq 1$.

- Otherwise, generate a random number, r , uniformly distributed between 0 and 1. Accept the new value of the spin if: $r < \mathcal{R}$
- Otherwise reject the attempted change: Put back the value of the spin to its original value.
- Visit the next spin and repeat.
- When you visit all the spins, you have completed a SWEEP.
- Start another sweep. And so on.
- After many sweeps, the system will be at equilibrium and the spin configurations given by the correct Boltzmann weight.
- Continue generating new configurations now to calculate averages.
- To calculate physical quantities, visit the configuration and calculate.
- Average over all measurements. The average is arithmetic because the measurements have the correct weight built in.

Important considerations:

- We need to make sure the system has thermalized, in other words P is stationary. Otherwise, the configurations are not drawn from equilibrium.
- The update scheme I described is local: In the Ising example I showed, we saw that only the close neighbors of a given spin participate in its evolution
- Away from a critical point, this is fine and evolution is efficient because correlation the length is short.
- But near a critical point, the correlation length is very long and one needs many more iterations for the effect of a flipped spin to be felt by all the spins correlated with it. This is called **critical slowing down**.
- In some cases it is possible to fight it by constructing "cluster algorithms", e.g. the Wolff algorithm.

Quantum Monte Carlo

Apply the Metropolis method to quantum systems:

 $Z = \text{Tr} e^{-\beta \hat{H}}$

But the partition function here is a trace over an operator: How can we apply Metropolis to something like this? Recall that for Metropolis, we changed c-numbers and probabilities; how can one "change operators"?

We change matrix elements of operators!

The trace of an operator is invariant, we can calculate it in any representation, |ψ>:

$$
Z = \sum \langle \Psi_1 | e^{-\beta H} | \Psi_1 \rangle
$$

=
$$
\sum \langle \Psi_1 | e^{-\Delta \tau H} e^{-\Delta \tau H} e^{-\Delta \tau H} \dots e^{-\Delta \tau H} | \Psi_1 \rangle
$$

Lexponentials: $\beta = L \Delta \tau$

L is chosen such that $\Delta\tau$ times the largest energy scale in H is very small.

Quantum Monte Carlo

So:
\n
$$
Z = \sum \langle \Psi_1 | e^{-\Delta \tau H} e^{-\Delta \tau H} e^{-\Delta \tau H} \dots e^{-\Delta \tau H} | \Psi_1 \rangle
$$
\n
$$
= \sum \langle \Psi_1 | e^{-\Delta \tau H} | \Psi_L \rangle \langle \Psi_L | e^{-\Delta \tau H} | \Psi_{L-1} \rangle \langle \Psi_{L-1} | e^{-\Delta \tau H} | \Psi_{L-2} \rangle \langle \Psi_{L-2} | \dots
$$
\n
$$
\dots \dots | \Psi_2 \rangle \langle \Psi_2 | e^{-\Delta \tau H} | \Psi_1 \rangle
$$
\nUsed:

\n
$$
1 = \sum |\Psi\rangle\langle \Psi|
$$

- The partition function is now a sum over a product of matrix elements, i.e. c-numbers!
- Time evolution operator: e^{-iHt}
- Matrix elements formally look like (imaginary) time evolution between the two successive states: $|\Psi_n\rangle \rightarrow |\Psi_{n+1}\rangle$
- If the product of the matrix elements is positive, it can be used as a Boltzmann weight and the Metropolis algorithm can be used.
- If the product is negative or, worse, complex, we are in trouble!
- Sometimes, we can avoid this with a good choice of states. But **NOT ALWAYS!**

Refreneces:

- ``Monte Carlo simulations of one-dimensional fermion systems'', J. E. Hirsch, R. L. Sugar, D. J. Scalapino, and R. Blankenbecler Phys. Rev. B26 (1982) 5033.
- ``World-line quantum Monte Carlo algorithm for a onedimensional Bose model'', G. G. Batrouni et R. T. Scalettar, Phys. Rev. B46 (1992) 9051.

This is a relatively easy algorithm to understand intuitively and to implement in code. You can easily calculate all diagonal quantities (we will see what this means) but not non-diagonal ones, such as Green functions.

A large class of Hamiltonians can be written in the form:

 $H = K + U$

Where K is nondiagonal, for example:

$$
\mathsf{K} = -t \sum_{\langle ij \rangle} (\hat{a}_{i}^{\dagger} \hat{a}_{j} + \hat{a}_{j}^{\dagger} \hat{a}_{i})
$$

And U is diagonal, for example:

$$
U = V_0 \sum_i \hat{n}_i (\hat{n}_i - 1) + V_1 \sum_{\langle ij \rangle} \hat{n}_i \hat{n}_j
$$

$$
\hat{n}_i = \hat{a}_i^{\dagger} \hat{a}_i, \qquad [\hat{a}_i, \hat{a}_j^{\dagger}] = \delta_{i,j}
$$

This type of model will be our focus.

To fix ideas, we will take the one-dimensional Bose-Hubbard model:

$$
H = -t \sum_{i} (\hat{a}_{i}^{\dagger} \hat{a}_{i+1} + \hat{a}_{i+1}^{\dagger} \hat{a}_{i}) + V_{0} \sum_{i} \hat{n}_{i} (\hat{n}_{i} - 1) + V_{1} \sum_{i} \hat{n}_{i} \hat{n}_{i+1} + \dots
$$

\nWrite:
\n
$$
H_{1} = -t \sum_{i \text{ odd}} (\mathbf{a}_{i}^{\dagger} \mathbf{a}_{i+1} + \mathbf{a}_{i+1}^{\dagger} \mathbf{a}_{i}) + \frac{1}{2} \mathbf{U}
$$
\n
$$
H = H_{1} + H_{2}
$$
\n
$$
H_{2} = -t \sum_{i \text{ even}} (\mathbf{a}_{i}^{\dagger} \mathbf{a}_{i+1} + \mathbf{a}_{i+1}^{\dagger} \mathbf{a}_{i}) + \frac{1}{2} \mathbf{U}
$$
\n
$$
H = H_{1} + H_{2}
$$

 $H^{}_1$ connects odd site to even site after it, and $H^{}_2$ connects even site to odd site after it. This allows us to write the first approximation:

 $e^{-\Delta \tau H} \approx e^{-\Delta \tau H_2} e^{-\Delta \tau H_1} + \mathcal{O}\{(\Delta \tau)^2 [H_1, H_2]\}.$ (Checkerboard decomposition)

where we used the Baker-Hausdorff formula

$$
e^A e^B = e^{A+B+\frac{1}{2}[A,B]+\dots}
$$

Then, the partition function becomes:

$$
Z = \text{Tr} e^{-\beta H} = \text{Tr} e^{-\Delta \tau H} e^{-\Delta \tau H} e^{-\Delta \tau H} \dots e^{-\Delta \tau H} e^{-\Delta \tau H} \cdot L_{\tau} \text{ exponentials}
$$

\n
$$
\approx \text{Tr} e^{-\Delta \tau H_2} e^{-\Delta \tau H_1} e^{-\Delta \tau H_2} e^{-\Delta \tau H_1} \dots
$$

\n
$$
\dots e^{-\Delta \tau H_2} e^{-\Delta \tau H_1} \dots e^{-\Delta \tau H_2} e^{-\Delta \tau H_1}
$$
 $2L_{\tau} \text{ exponentials}$

Now consider the occupation number states:

$$
a_i^{\dagger} | n_1, n_2, \dots, n_i, \dots, n_S \rangle = \sqrt{n_i + 1} | n_1, n_2, \dots, n_i + 1, \dots, n_S \rangle
$$

\n
$$
a_i | n_1, n_2, \dots, n_i, \dots, n_S \rangle = \sqrt{n_i} | n_1, n_2, \dots, n_i - 1, \dots, n_S \rangle
$$

\n
$$
n_i | n_1, n_2, \dots, n_i, \dots, n_S \rangle = n_i | n_1, n_2, \dots, n_i, \dots, n_S \rangle
$$

The subscripts label the sites on the 1-dimensional lattice. To evaluate the trace, introduce between each pair of exponentials the identity:

$$
I = \sum_{\{n\}} |\mathbf{n}\rangle \langle \mathbf{n}| = \sum_{n_1, n_2, \dots, n_i \dots n_S} |n_1, n_2, \dots, n_i, \dots, n_S\rangle \langle n_1, n_2, \dots, n_i, \dots, n_S|
$$

Sites on shaded squares are connected.

Two consecutive matrix elements advance the entire system by one imaginary time step: All sites will be connected.

Consider a typical move:

In a typical move, only four matrix elements change. **They can be evaluated independently**. That is why we broke up the Hamiltonian as we did.

- Visit a shaded square.
- Randomly choose: Try a left jump or a right jump.
- Check if the choice is possible.
- Not possible: move to next square.
- Possible: Evaluate the product of the 4 elements before the move.
- Then: Evaluate the product of the 4 elements after the move.
- Accept/reject using the Metropolis criterion.
- Move to next shaded square.
- And so on.
- The problem is now reduced to a product of independent 4-site matrix elements which can be evaluated exactly.

$$
n'_{j}
$$
\n
$$
e^{-\Delta \tau H_{\ell}}
$$
\n
$$
n_{j+1}
$$
\n
$$
n_{j+1}
$$

$$
\langle n'_j, n'_{j+1}|e^{-\Delta\tau H_\ell}|n_j, n_{j+1}\rangle = \langle n'_j, n'_{j+1}|e^{-\Delta\tau(K_\ell + \frac{1}{2}U_\ell)}|n_j, n_{j+1}\rangle
$$

\n
$$
\approx \langle n'_j, n'_{j+1}|e^{-\Delta\tau \frac{1}{4}U_\ell}e^{-\Delta\tau K_\ell}e^{-\Delta\tau \frac{1}{4}U_\ell}|n_j, n_{j+1}\rangle
$$

\n
$$
\approx e^{-\Delta\tau U'/4}e^{-\Delta\tau U/4}\langle n'_j, n'_{j+1}|e^{-\Delta\tau K_\ell}|n_j, n_{j+1}\rangle
$$

$$
U = \frac{V_0}{2} \left[n_j(n_j - 1) + n_{j+1}(n_{j+1} - 1) \right]
$$

The entire update sweep is comprised of evaluations of such matrix elements.

$$
\Pi_i = \langle 0, 0 | e^{-\Delta \tau H_1} | 0, 0 \rangle \langle 1, 0 | e^{-\Delta \tau H_2} | 1, 0 \rangle
$$

$$
\langle 0, 1 | e^{-\Delta \tau H_1} | 0, 1 \rangle \langle 1, 0 | e^{-\Delta \tau H_2} | 1, 0 \rangle
$$

$$
= \langle 0, 0 | e^{-\Delta \tau \mathsf{K}_1} | 0, 0 \rangle \langle 1, 0 | e^{-\Delta \tau \mathsf{K}_2} | 1, 0 \rangle
$$

$$
\langle 0, 1 | e^{-\Delta \tau \mathsf{K}_1} | 0, 1 \rangle \langle 1, 0 | e^{-\Delta \tau \mathsf{K}_2} | 1, 0 \rangle
$$

The matrix elements can be evaluated exactly numerically. But we can also expand the exponential:

$$
\langle 0, 1 | e^{-\Delta \tau \mathsf{K}_i} | 0, 1 \rangle \approx \langle 0, 1 | (1 - \Delta \tau \mathsf{K}_i) | 0, 1 \rangle
$$

\n
$$
= 1 - \Delta \tau \langle 0, 1 | (a_i^\dagger a_{i+1} + a_{i+1}^\dagger a_i) | 0, 1 \rangle = 1
$$

\n
$$
\langle 0, 1 | e^{-\Delta \tau \mathsf{K}_i} | 1, 0 \rangle \approx \langle 0, 1 | (1 - \Delta \tau \mathsf{K}_i) | 1, 0 \rangle
$$

\n
$$
= \langle 0, 1 | 1, 0 \rangle + t \Delta \tau \langle 0, 1 | (a_i^\dagger a_{i+1} + a_{i+1}^\dagger a_i) | 1, 0 \rangle
$$

\n
$$
= t \Delta \tau
$$

So, we have for our typical move:

$$
\Pi_i = 1 \qquad \qquad \Pi_f = (t\Delta\tau)^2
$$
\n
$$
\frac{\Pi_f}{\Pi_i} = (t\Delta\tau)^2 < 1
$$

Generate random number from uniform distribution [0,1)

Accept the new configuration if: $r \leq \frac{\prod_f}{\prod_i}$

$$
\Pi_{i} = \langle 1, 0 | e^{-\Delta \tau H_{1}} | 1, 0 \rangle \langle 1, 1 | e^{-\Delta \tau H_{2}} | 1, 1 \rangle
$$

$$
\langle 0, 1 | e^{-\Delta \tau H_{1}} | 0, 1 \rangle \langle 1, 1 | e^{-\Delta \tau H_{2}} | 1, 1 \rangle
$$

$$
= \langle 1, 0 | e^{-\Delta \tau \mathsf{K}_1} | 1, 0 \rangle \langle 1, 1 | e^{-\Delta \tau \mathsf{K}_2} | 1, 1 \rangle
$$

$$
\langle 0, 1 | e^{-\Delta \tau \mathsf{K}_1} | 0, 1 \rangle \langle 1, 1 | e^{-\Delta \tau \mathsf{K}_2} | 1, 1 \rangle
$$

 $\overline{1}$ (Calculated exactly as before) \equiv

 $\Pi_f = \langle 2, 0 | e^{-\Delta \tau H_1} | 2, 0 \rangle \langle 1, 1 | e^{-\Delta \tau H_2} | 0, 2 \rangle$ $\langle 0,0|e^{-\Delta \tau H_1}|0,0\rangle\langle 0,2|e^{-\Delta \tau H_2}|1,1\rangle$

$$
= e^{-\frac{\Delta \tau V_0}{4}} e^{-\frac{\Delta \tau V_0}{4}} \langle 2, 0 | e^{-\Delta \tau K_1} | 2, 0 \rangle \langle 1, 1 | e^{-\Delta \tau K_2} | 0, 2 \rangle e^{-\frac{\Delta \tau V_0}{4}}
$$

$$
\langle 0, 0 | e^{-\Delta \tau K_1} | 0, 0 \rangle \langle 0, 2 | e^{-\Delta \tau K_2} | 1, 1 \rangle e^{-\frac{\Delta \tau V_0}{4}}
$$

$$
= e^{-\Delta \tau V_0} \langle 2, 0 | e^{-\Delta \tau K_1} | 2, 0 \rangle \langle 1, 1 | e^{-\Delta \tau K_2} | 0, 2 \rangle
$$

$$
\langle 0, 0 | e^{-\Delta \tau K_1} | 0, 0 \rangle \langle 0, 2 | e^{-\Delta \tau K_2} | 1, 1 \rangle
$$

The matrix elements can be evaluated as before:

$$
\langle 2, 0 | e^{-\Delta \tau K_i} | 1, 1 \rangle \approx \langle 2, 0 | (1 - \Delta \tau K_i) | 1, 1 \rangle
$$

=
$$
\langle 2, 0 | 1, 1 \rangle + t \Delta \tau \langle 2, 0 | (a_i^{\dagger} a_{i+1} + a_{i+1}^{\dagger} a_i) | 1, 1 \rangle
$$

=
$$
t \Delta \tau \langle 2, 0 | a_i^{\dagger} a_{i+1} | 1, 1 \rangle
$$

=
$$
t \sqrt{2} \Delta \tau
$$

So, we have for our typical move:

$$
\Pi_i = 1 \qquad \qquad \Pi_f = 2e^{-\Delta \tau U} (t\Delta \tau)^2
$$
\n
$$
\frac{\Pi_f}{\Pi_i} = 2e^{-\Delta \tau U} (t\Delta \tau)^2 < 1
$$

Generate random number from uniform distribution [0,1)

Accept the new configuration if: $r \leq \frac{\prod_f}{\prod_i}$

On the other hand:

 $\mathbf{1}$

Then:

$$
\Pi_i = 2e^{-\Delta \tau U} (t\Delta \tau)^2
$$
\n
$$
\frac{\Pi_f}{\Pi_i} = \left[2e^{-\Delta \tau U} (t\Delta \tau)^2\right]^{-1} > 1
$$
\n
$$
\Pi_f =
$$

Accept!

So, we now know how to make the Monte Carlo moves. We do many sweeps until the system thermalizes. Then we need to measure physical quantities and average them.

Why split H into H₁ and H₂? Recall, in my examples, I used expressions like: $e^{-\Delta \tau K_i} \approx 1 - \Delta \tau (a_i^{\dagger} a_{i+1} + a_{i+1}^{\dagger} a_i)$

to calculate the matrix elements of the 4-site problem. But in fact, for this 4-site problem, the matrix element of the **exponential** is calculated exactly:

- Calculate the matrix elements of K, assuming a maximum multiple occupancy on a site (say 6),
- Fourier transform (diagonalize),
- Exponentiate.
- Invert Fourier transform.

But to see how quickly this becomes complicated, go to second order:

$$
e^{-\Delta \tau K_{i}} \approx 1 - \Delta \tau (a_{i}^{\dagger} a_{i+1} + a_{i+1}^{\dagger} a_{i})
$$
\n
$$
+ \frac{(\Delta \tau)^{2}}{2} (a_{i}^{\dagger} a_{i+1} a_{i}^{\dagger} a_{i+1} + a_{i}^{\dagger} a_{i+1} a_{i+1}^{\dagger} a_{i})
$$
\n
$$
+ a_{i+1}^{\dagger} a_{i} a_{i}^{\dagger} a_{i+1} + a_{i+1}^{\dagger} a_{i} a_{i+1}^{\dagger} a_{i})
$$
\n
$$
+ a_{i+1}^{\dagger} a_{i} a_{i}^{\dagger} a_{i+1} + a_{i+1}^{\dagger} a_{i} a_{i+1}^{\dagger} a_{i})
$$
\n
$$
n_{j+1}
$$

The term: $a_i^{\dagger}a_{i+1}a_i^{\dagger}a_{i+1}$ destroys 2 particles on $i+1$ and creates 2 on i The term: $a_{i+1}^{\dagger}a_i a_i^{\dagger} a_{i+1}$ hops a particle from $i+1$ to i and back to $i+1$.

The higher the power, the more such terms are included. In fact we need to keep all of them because the sum up to finite contributions.

That is why we need to calculate the transition matrix exactly, and that is not hard for a 4-site problem.

What if we do not split H into H $_{\rm_1}$ and H $_{\rm_2}$?

How will we evolve the particles from the first slice to the next?

- Suggest sequential moves for one particle at a time and apply Metropolis? We can but the matrix is much larger than before. Why not just do the 4-site problem?
- Suggest a move for all particles? The acceptance rate will be very low. Think of Ising: We only change one site at a time because if we change many dependent sites, the acceptance will not be high.
- This suggests that it is more efficient to do the checkerboard splitting like I showed before.

Measurements:

$$
\langle O \rangle = \frac{1}{Z} \text{Tr} \left(O e^{-\beta H} \right)
$$

\n
$$
= \frac{1}{Z} \sum_{\{\mathbf{n}\}} \langle \mathbf{n}^1 | e^{-\Delta \tau H_2} | \mathbf{n}^{2L} \rangle \langle \mathbf{n}^{2L} | e^{-\Delta \tau H_1} | \mathbf{n}^{2L-1} \rangle ...
$$

\n
$$
\dots \langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} O | \mathbf{n}^{2i} \rangle \langle \mathbf{n}^{2i} | e^{-\Delta \tau H_1} | \mathbf{n}^{2i-1} \rangle ...
$$

\n
$$
\dots \langle \mathbf{n}^3 | e^{-\Delta \tau H_2} | \mathbf{n}^2 \rangle \langle \mathbf{n}^2 | e^{-\Delta \tau H_1} | \mathbf{n}^1 \rangle
$$

\n
$$
= \frac{1}{Z} \sum_{\{\mathbf{n}\}} \langle \mathbf{n}^1 | e^{-\Delta \tau H_2} | \mathbf{n}^{2L} \rangle \langle \mathbf{n}^{2L} | e^{-\Delta \tau H_1} | \mathbf{n}^{2L-1} \rangle ...
$$

\n
$$
\dots \langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} O | \mathbf{n}^{2i} \rangle \frac{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} | \mathbf{n}^{2i} \rangle}{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} | \mathbf{n}^{2i} \rangle} \langle \mathbf{n}^{2i} | e^{-\Delta \tau H_1} | \mathbf{n}^{2i-1} \rangle ...
$$

\n
$$
\dots \langle \mathbf{n}^3 | e^{-\Delta \tau H_2} | \mathbf{n}^2 \rangle \langle \mathbf{n}^2 | e^{-\Delta \tau H_1} | \mathbf{n}^1 \rangle
$$

Measurements:

$$
\langle O \rangle = \frac{1}{Z} \sum_{\{\mathbf{n}\}} \langle \mathbf{n}^{1} | e^{-\Delta \tau H_{2}} | \mathbf{n}^{2L} \rangle \langle \mathbf{n}^{2L} | e^{-\Delta \tau H_{1}} | \mathbf{n}^{2L-1} \rangle ...
$$

\n
$$
\dots \langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_{2}} | \mathbf{n}^{2i} \rangle \frac{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_{2}} O | \mathbf{n}^{2i} \rangle}{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_{2}} | \mathbf{n}^{2i} \rangle} \langle \mathbf{n}^{2i} | e^{-\Delta \tau H_{1}} | \mathbf{n}^{2i-1} \rangle ...
$$

\n
$$
\dots \langle \mathbf{n}^{3} | e^{-\Delta \tau H_{2}} | \mathbf{n}^{2} \rangle \langle \mathbf{n}^{2} | e^{-\Delta \tau H_{1}} | \mathbf{n}^{1} \rangle
$$

\n
$$
= \frac{1}{Z} \sum_{\{\mathbf{n}\}} \langle \mathbf{n}^{1} | e^{-\Delta \tau H_{2}} | \mathbf{n}^{2L} \rangle \langle \mathbf{n}^{2L} | e^{-\Delta \tau H_{1}} | \mathbf{n}^{2L-1} \rangle ...
$$

\n
$$
\dots \langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_{2}} | \mathbf{n}^{2i} \rangle \langle \mathbf{n}^{2i} | e^{-\Delta \tau H_{1}} | \mathbf{n}^{2i-1} \rangle ...
$$

\n
$$
\dots \langle \mathbf{n}^{3} | e^{-\Delta \tau H_{2}} | \mathbf{n}^{2} \rangle \langle \mathbf{n}^{2} | e^{-\Delta \tau H_{1}} | \mathbf{n}^{1} \rangle
$$

\n
$$
\frac{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_{2}} O | \mathbf{n}^{2i} \rangle}{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_{2}} | \mathbf{n}^{2i} \rangle}
$$

\n
$$
\rangle
$$

\nMeasurable

The previous expression takes the form:

$$
\langle \mathbf{O} \rangle = \sum_{\{\mathbf{n}\}} P_B[\{n_i\}] \frac{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} \mathbf{O} | \mathbf{n}^{2i} \rangle}{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} | \mathbf{n}^{2i} \rangle}
$$

- We need to calculate the ratio of the matrix elements
- For quantities which are diagonal in occupation, this is very simple.

$$
\langle \mathcal{F}(\hat{n}) \rangle = \sum_{\{\mathbf{n}\}} P_B[\{n_i\}] \frac{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} \mathcal{F}(\hat{n}) | \mathbf{n}^{2i} \rangle}{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} | \mathbf{n}^{2i} \rangle}
$$

\n
$$
= \sum_{\{\mathbf{n}\}} P_B[\{n_i\}] \frac{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} \mathcal{F}(n) | \mathbf{n}^{2i} \rangle}{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} | \mathbf{n}^{2i} \rangle}
$$

\n
$$
= \sum_{\{\mathbf{n}\}} \mathcal{F}(n) P_B[\{n_i\}] \frac{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} | \mathbf{n}^{2i} \rangle}{\langle \mathbf{n}^{2i+1} | e^{-\Delta \tau H_2} | \mathbf{n}^{2i} \rangle}
$$

- We can calculate static (equal imaginary "time") correlation functions.
- We can also calculate dynamic (separated in imaginary "time") correlation functions.
- Use maximum entropy, for example, to get dynamic properties, like excited states etc.
- Correlations in the "time" direction also give us the superfluid density.
- We cannot calculate Green functions with this algorithm because the world lines are continuous, they do not break.

Avery interesting quantity for Bose systems is the superfluid density:

where W is the winding number of the world line configuration.

In two dimensions, we do the same. However, one full time step is now accomplished with 4 exponentials due to the breakup of the H.

References:

- ``Worm Algorithm for Problems of Quantum and Classical Statistics'', Nikolay Prokof'ev and Boris Svistunov, Understanding Quantum Phase Transitions, CRC Press, Taylor and Francis Group, Lincoln D. Carr (editor), (2011).
- ``Quantum Monte Carlo simulation in the canonical ensemble at finite temperature'', K. Van Houcke, S. M. A. Rombouts, and L. Pollet, Phys. Rev. E73, 056703 (2006).
- ``Stochastic Green function algorithm", V. G. Rousseau, Phys. Rev. E77, 056705 (2008).
- ``Directed update for the stochastic Green function algorithm", V. G. Rousseau, Phys. Rev. E78, 056707 (2008).

General features:

- Extension of the canonical worm algorithm.
- Works well in canonical and grand canonical ensembles.
- Continuous imaginary time: No Δτ errors.
- World lines are broken and ends move around: Can calculate Green functions.
- Good for complicated Hamiltonians.

Examples of Hamiltonians where SGF is useful:

Spin-1/2 Bose-Hubbard model:

$$
H = -t \sum_{\sigma, \langle \mathbf{r}, \mathbf{r}' \rangle} \left(a_{\sigma \mathbf{r}}^{\dagger} a_{\sigma \mathbf{r}'} + a_{\sigma \mathbf{r}'}^{\dagger} a_{\sigma \mathbf{r}} \right) - \mu \sum_{\sigma, \mathbf{r}} \hat{n}_{\sigma \mathbf{r}} + \frac{U_0}{2} \sum_{\sigma, \mathbf{r}} \hat{n}_{\sigma \mathbf{r}} (\hat{n}_{\sigma \mathbf{r}} - 1) + (U_0 + U_2) \sum_{\mathbf{r}} \hat{n}_{\uparrow}(\mathbf{r}) \hat{n}_{\downarrow}(\mathbf{r}) - \frac{U_2}{2} \sum_{\mathbf{r}} \left(a_{\uparrow \mathbf{r}}^{\dagger} a_{\uparrow \mathbf{r}}^{\dagger} a_{\downarrow \mathbf{r}} + a_{\downarrow \mathbf{r}}^{\dagger} a_{\downarrow \mathbf{r}}^{\dagger} a_{\uparrow \mathbf{r}} \right)
$$

The one dimensional Rabi-Hubbard model:

$$
H = -J\sum_{i} (a_i^{\dagger} a_{i+1} + a_{i+1}^{\dagger} a_i) + \sum_{i} \left(\omega_s \sigma_i^+ \sigma_i^- + \omega_l a_i^{\dagger} a_i\right)
$$

$$
-g\sum_{i} (\sigma_i^- + \sigma_i^+) (a_i + a_i^{\dagger})
$$

As usual, we start with the partition function:

 $Z(\beta) = \text{Tr} \, \mathrm{e}^{-\beta H}$

Now we will write the H in the form:

 $H=U-T$

Where U is the diagonal part and depends only on number operators. T is the nondiagonal part which can be more complicated than just the hopping term: It can have ring exchange operators, conversion from one particle type to another, etc ...

The minus sign in front of T is needed next page. T itself now is assumed to have positive signs.

Define the imaginary time-dependent operator:

$$
T(\tau) \equiv e^{\tau U} T e^{-\tau U}
$$

In the interaction picture we can write:

$$
e^{-\beta H} = e^{-\beta U} \exp\left(\int_0^{\beta} d\tau \ T(\tau)\right)
$$

= $e^{-\beta U} \left\{ 1 + \int_0^{\beta} d\tau \ T(\tau) + \int_0^{\beta} d\tau_1 \int_{\tau_1}^{\beta} d\tau_2 \ T(\tau_2) T(\tau_1) + \dots \right\}$
= $e^{-\beta U} \sum_{n>0} \int_{0<\tau_1<\tau_2<\dots<\tau_n<\beta} d\tau_1 \dots d\tau_n T(\tau_n) \dots T(\tau_2) T(\tau_1)$

The exponential of the nondiagonal part is **time ordered**. It takes into account all possible (imaginary) times when the T can act. So, the partition function can be written as:

$$
Z(\beta) = \text{Tr}\left\{e^{-\beta U} \sum_{n\geq 0} \int_{0<\tau_1<\tau_2<\cdots<\tau_n<\beta} d\tau_1 \dots d\tau_n \mathsf{T}(\tau_n) \dots \mathsf{T}(\tau_2) \mathsf{T}(\tau_1)\right\}
$$

$$
Z(\beta) = \sum_{\{\Psi_0\}, n \ge 0} \int_{0 < \tau_1 < \tau_2 < \cdots < \tau_n < \beta} d\tau_1 \dots d\tau_n \langle \Psi_0 | e^{-\beta U} \mathsf{T}(\tau_n) \dots \mathsf{T}(\tau_2) \mathsf{T}(\tau_1) | \Psi_0 \rangle
$$
\n
$$
= \sum_{\{\Psi_n\}, n \ge 0} \int_{0 < \tau_1 < \tau_2 < \cdots < \tau_n < \beta} d\tau_1 \dots d\tau_n \langle \Psi_0 | e^{-\beta U} \mathsf{T}(\tau_n) | \Psi_{n-1} \rangle
$$
\n
$$
\langle \Psi_{n-1} | \mathsf{T}(\tau_{n-1}) | \Psi_{n-2} \rangle \langle \Psi_{n-2} | \mathsf{T}(\tau_{n-2}) | \Psi_{n-3} \rangle \dots
$$
\n
$$
\dots \langle \Psi_2 | \mathsf{T}(\tau_2) | \Psi_1 \rangle \langle \Psi_1 | \mathsf{T}(\tau_1) | \Psi_0 \rangle
$$

So far this is just the usual partition function written is a slightly different form.

Note: The imaginary time insertions are not at discretized imaginary time steps, they can be anywhere between 0 and β. That is what makes this a continuous time algorithm.

Now let us write an "extended" partition function:

 $Z(\beta,\tau) \equiv \text{Tr} \, \mathrm{e}^{-(\beta-\tau)H} \mathsf{G} \, \mathrm{e}^{-\tau H}$ G is an arbitrary operator for now.

Defining,
$$
G(\tau) \equiv e^{\tau U} G e^{-\tau U}
$$
 we can show as before:
\n
$$
Z(\beta, \tau) = \sum_{\{\Psi_n\}, n \ge 0} \int_{0 < \tau_1 < \tau_2 < \dots < \tau_n < \beta} d\tau_1 \dots d\tau_n \langle \Psi_0 | e^{-\beta U} \mathsf{T}(\tau_n) | \Psi_{n-1} \rangle
$$
\n
$$
\langle \Psi_{n-1} | \mathsf{T}(\tau_{n-1}) | \Psi_{n-2} \rangle \langle \Psi_{n-2} | \mathsf{T}(\tau_{n-2}) | \Psi_{n-3} \rangle \dots
$$
\n
$$
\dots \langle \Psi_{L+1} | \mathsf{T}(\tau_L) | \Psi_L \rangle \langle \Psi_L | G(\tau) | \Psi_R \rangle \langle \Psi_R | \mathsf{T}(\tau_r) | \Psi_{R-1} \rangle
$$
\n
$$
\dots \langle \Psi_2 | \mathsf{T}(\tau_2) | \Psi_1 \rangle \langle \Psi_1 | \mathsf{T}(\tau_1) | \Psi_0 \rangle
$$

This is the same as the previous partition function, except we have this G acting at τ which we introduced but not yet explained.

Define:

$$
\mathsf{A}^\dagger \equiv \mathsf{a}^\dagger \frac{1}{\sqrt{\mathsf{n}+1}}, \quad \mathsf{A} \equiv \frac{1}{\sqrt{\mathsf{n}+1}} \mathsf{a} \implies \mathsf{A}^\dagger |n\rangle = |n+1\rangle, \quad \mathsf{A} |n\rangle = |n-1\rangle
$$

Then, G is defined by:

$$
\mathsf{G} \equiv \sum_{p,q=0}^{\infty} g_{pq} \sum_{\{i_p | j_q\}} \prod_{k=1}^{p} \mathsf{A}^{\dagger}_{i_k} \prod_{\ell=1}^{q} \mathsf{A}_{j_\ell}
$$

- $g_{_{\rho q}}$ is a matrix of our choice which depends on the application.
- It determines how many creation and annihilation operators can be be inserted.
- Typically, it decreases rapidly with number of insertions.
- $(p+q\leq 2)$ • Example: g_{pq} = 1 $= e^{-4(2-p-q)^2}$ $(p+q>2)$

$$
\mathsf{G} \equiv \sum_{p,q=0}^{\infty} g_{pq} \sum_{\{i_p | j_q\}} \prod_{k=1}^{p} \mathsf{A}^{\dagger}_{i_k} \prod_{\ell=1}^{q} \mathsf{A}_{j_{\ell}}
$$

- i_{κ} and $j_{_{\ell}}$ are site indices where we insert the operators
- $\{i_\rho|j_\rho\}$ denotes that all site indices in subset i are different from all indices in subset *j*. But one or more indices in a subset can be equal.
- The operator G inserts inserts breaks in the world lines.
- These are the "worms" in the algorithm.
- When the world lines close, we have a configuration of continuous lines.
- We can then measure diagonal quantities.

- For illustration, I have taken T to be a hop between near neighbors.
- In general, the nondiagonal operator can be more complicated.
- How far to shift? It is chosen according to a probability distribution, typically exponential.

Now, shift left and create: The shift must stop before $τ_$ because of time ordering

If we shift left and destroy, then:

- The head and tail of all worms are at the same imaginary time.
- Insures canonical ensemble.
- The canonical worm algorithm: Only one head and tail (one worm) and at the same imaginary time.
- Worm algorithm: Only the head or tail moves and is not at the same time.
- Worm algorithm can change the number of world lines, it is grand canonical.
- Measure Green functions: Just count how often a worm of a certain length appears.
- Measure diagonal quantities: Only when worms close. Then measurement is trivial.

Detailed Balance:

- Let P_{i} (P_{f}) be the probability of the initial (final) configuration.
- S_{i} the transition probability from *i* to *f.*
- A_{i} the acceptance rate for the transition from *i* to *f.*

Detailed balance condition is: $P_iS_{i\rightarrow f}A_{i\rightarrow f}=P_fS_{f\rightarrow i}A_{f\rightarrow i}$

Typically we choose the Metropolis solution:

$$
A_{i \to f} = \min(1, q)
$$

with
$$
q = \frac{P_f S_{f \to i}}{P_i S_{i \to f}}
$$

For details of the probabilities, P, S and the probabilities to choose left or right, create or destroy, see the references. They are discussed in detail.

Quantum Monte Carlo: Some Last comments

- The world line algorithm is intuitive and easy to code, especially in one dimension.
- It is still very useful for simpler bosonic models.
- Cannot measure Green functions.
- The SGF and other worm algorithms are powerful and, depending on the application, should be chosen.
- Always choose an algorithm that is adapted to the problem.
- Choose if it is more practical to simulate in the canonical or grand canonical ensembles.
- Algorithm development is an active, and fun, part of what we do!