Nuclear interactions and quantum Monte Carlo methods

Maria Piarulli—Washington University, St. Louis

July 11-13, 2022

2022 National Nuclear Physics Summer School (NNPSS)—MIT

Lecture 2: The many-body problem and Quantum Monte Carlo methods

- The nuclear many-body problem
- Nuclear quantum Monte Carlo
- Variational Monte Carlo
- Green's function Monte Carlo
- Auxiliary Field diffusion Monte Carlo

The nuclear many-body problem

Many-body Schrödinger equation:

$$H\,\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_A; s_1, s_2, \ldots, s_A; t_1, t_2, \ldots, t_A)$$

$$= E\,\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_A; s_1, s_2, \ldots, s_A; t_1, t_2, \ldots, t_A)$$

Erwin Schrödinger

where $\mathbf{r}_i$, s_i, and t_i are the nucleon coordinates, spins, and isospins , respectively

This corresponds to solve

$2^A \times \binom{A}{Z}$ coupled second-order differential equations in 3A dimensions.

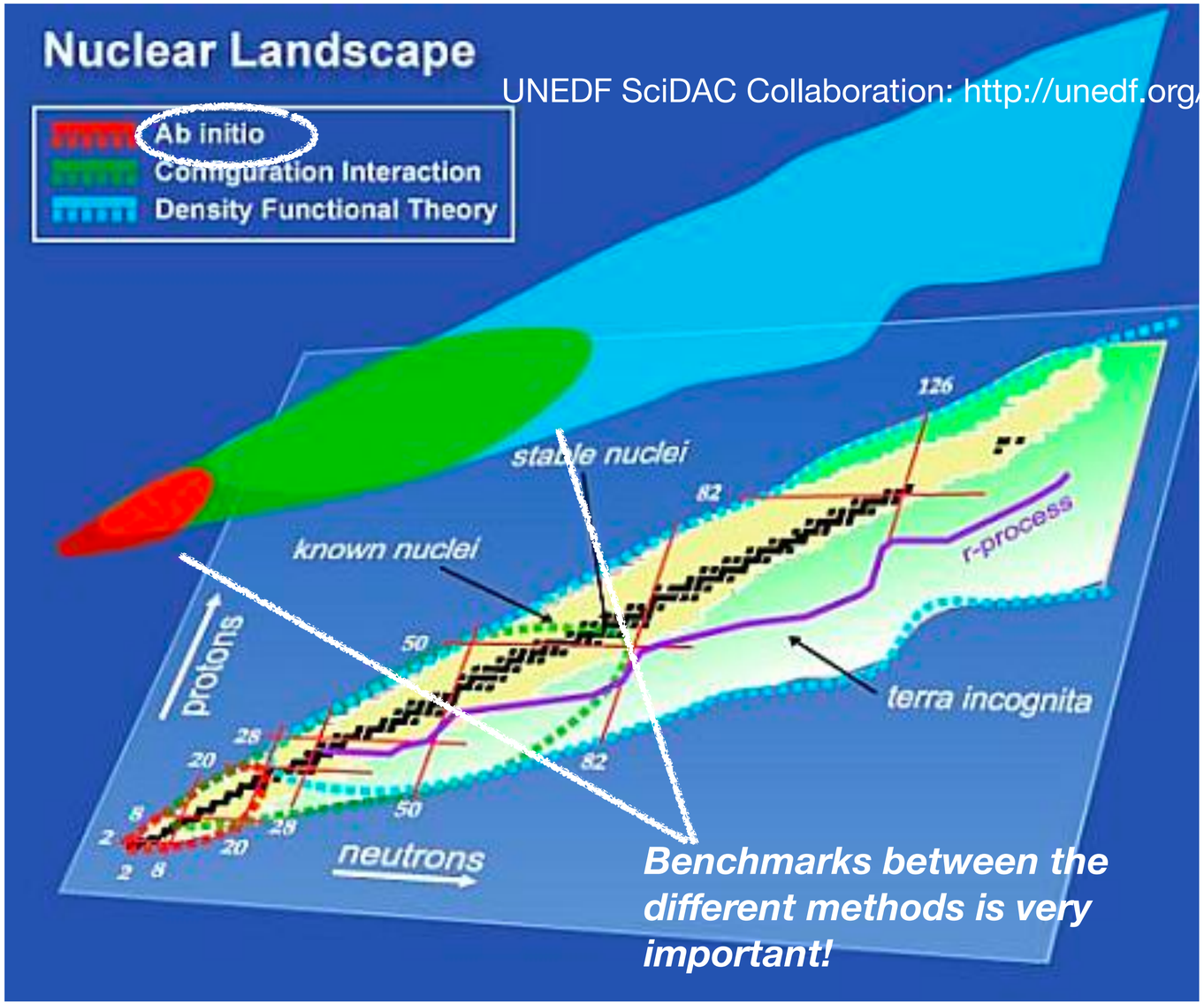
96 for ^{4}He

17,920 for ^{8}Be

3,784,704 for ^{12}C

This is a challenging many-body problem!

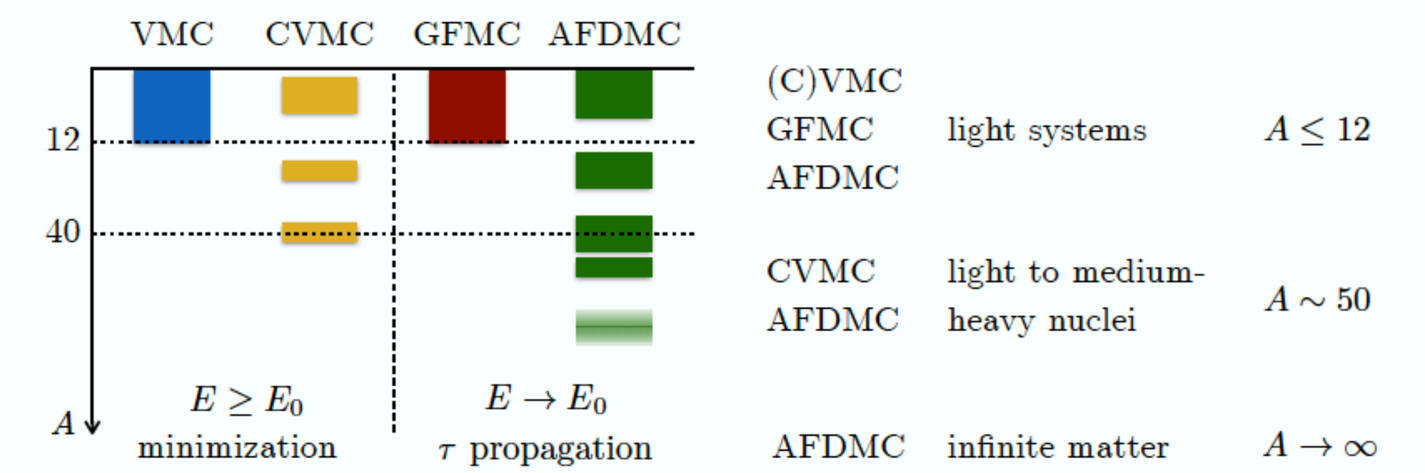
The nuclear landscape



Definition: the *ab-initio* methods seek to describe atomic nucleus from the ground up by solving the non-relativistic Schrödinger equation for all constituent nucleons and the forces between them

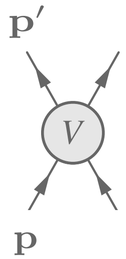
Nuclear quantum Monte Carlo methods

- Quantum Monte Carlo (QMC) methods: a large family of computational methods whose common aim is the study of complex quantum systems—J. Carlson et al., RMP. 87, 1067 (2015); J.E. Lynn et al., Ann. Rev. Nucl. Part. Sci 279, 69 (2019); S. Gandolfi, MP et. al., Front.in Phys. 8 (2020) 117



Computational resources awarded by the DOE ALCC and INCITE programs

- Work with bare interactions but local r-space representation of the Hamiltonian



$$\mathbf{k} = \mathbf{p}' - \mathbf{p}$$

Local

$$\mathbf{K} = (\mathbf{p}' + \mathbf{p})/2$$

Non-Local

- Stochastic method: based on recursive sampling of a probability density, statistical errors quantifiable and systematically improvable

Variational Monte Carlo

R.B. Wiringa, PRC **43**, 1585 (1991)

- In variational Monte Carlo, one minimize the expectation value of *H*: $\frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} = E_T \geq E_0$

- Trial wave function involves variational parameters $\{\alpha\}$ and minimization algorithms are used to search the parameter space: 12-50 parameters

- One assumes a suitable form for the trial wave function: $|\Psi_T\rangle = \left(1 - \sum_{i<j<k} F_{ijk}\right)\left(\mathcal{S}\prod_{i<j} F_{ij}\right)|\Phi_J\rangle$

$|\Phi_J\rangle$ is the antisymmetric Jastrow wave function depends on the nuclear state under investigation

For s-shell nuclei: $|\Phi_J\rangle = \left[\prod_{i<j<k} f^c_{ijk}(\mathbf{r}_{ik}, \mathbf{r}_{jk}) \prod_{i<j} f_c(r_{ij})\right] |\Phi_A(JMTT_3)\rangle$

$f_c(r_{ij})$ and $f^c_{ijk}(\mathbf{r}_{ik}, \mathbf{r}_{jk})$ are central two- and three-body correlations induced by the NN potential

$f^c_{ijk} = 1 - q^c_1(\mathbf{r}_{ij}\cdot\mathbf{r}_{ik})(\mathbf{r}_{ij}\cdot\mathbf{r}_{jk})(\mathbf{r}_{ik}\cdot\mathbf{r}_{jk})e^{-q^c_2(r_{ij}+r_{ik}+r_{jk})}$, where q^c_1 and q^c_2 are variational parameters

(reduce the two body correlations when a third particle comes between the correlated pairs)

Single particle wf (no spatial dependence for s-shell nuclei) ex. $|\Phi_\alpha(0000)\rangle = \mathcal{A}|\uparrow p \downarrow p \uparrow n \downarrow n\rangle$

Variational Monte Carlo

R.B. Wiringa, PRC **43**, 1585 (1991)

- In variational Monte Carlo, one minimize the expectation value of *H*: $\frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} = E_T \geq E_0$

- One assumes a suitable form for the trial wave function: $|\Psi_T\rangle = \left(1 - \sum_{i<j<k} F_{ijk}\right)\left(\mathcal{S}\prod_{i<j} F_{ij}\right)|\Phi_J\rangle$

F_{ij} are spin-isospin dependent pair correlations induced by the NN: $F_{ij} = \sum_{i=2,6}\left[\prod_{i<j<k} f^p_{ijk}(\mathbf{r}_{ik}, \mathbf{r}_{jk})\right] u_p(r_{ij}) O^p_{ij}$

$f^p_{ijk}(\mathbf{r}_{ij}, \mathbf{r}_{ik}) = 1 - q^p_1(1 - \hat{\mathbf{r}}_{ik} \cdot \hat{\mathbf{r}}_{jk}) e^{-q^p_2(r_{ij}+r_{ik}+r_{jk})}$ with q^p_1 and q^p_2 being variational parameters

$S\Pi_{i<j}$ denotes a symmetrized product over nucleon pairs since, in general, the F_{ij} do not commute

F_{ijk} are three-body correlations induced by the NNN potential $F_{ijk} = \sum_q \epsilon_q V^q_{ijk}(y_q r_{ij}, y_q r_{ik}, y_q r_{jk})$

Functions $f_c(r_{ij})$ and $u_p(r_{ij})$ obtained from coupled differential equations with v_{ij}.

p-shell nuclei: Pudliner at al., Phys. Rev. Lett. 74 (1995) 4396-4399

- The *LS* coupling scheme is used to obtain the desired *JM* value of a given state, as suggested by the shell-model studies of light p-shell nuclei. Different possible *LS* combinations lead to multiple components in the Jastrow wave function.

$$|\Phi_J\rangle = \mathscr{A}\left\{\prod_{i<j<k} f^c_{ijk} \prod_{i<j\leq 4} f_{ss}(r_{ij}) \sum_{LS[n]} \beta_{LS[n]} \prod_{k\leq 4<l\leq A} f_{sp}(r_{kl}) \prod_{4<l<m\leq A} f^{[n]}_{pp}(r_{lm}) \; |\Phi_A(LS[n]JJ_zT_z)_{1234:5\ldots A}\rangle\right\}$$

$$\left|\Phi_\alpha(0000)_{1234} \prod_{4<l\leq A} \phi_p^{LS[n]}(R_{\alpha l}) \left\{[Y_1^{m_l}(\Omega_{\alpha l})]_{LM_L} \otimes [\chi_l(\tfrac{1}{2}m_s)]_{SM_S}\right\}_{JM} [\nu_l(\tfrac{1}{2}t_3)]_{TT_3}\right\rangle$$

- The operator $\mathscr{A}$ indicates an antisymmetric sum over all possible partitions of the A particles into 4 s-shell and (A – 4) p-shell ones.
- The central correlation $f_{ss}(r_{ij})$ comes from the structure of an α particle (it is the $f_c(r_{ij})$ from the ^{4}He wave function).
- The $f_{sp}(r_{kl})$ is similar to the $f_c(r_{ij})$ at short range, but with a long-range tail going to unity; this helps the wave function factorize to a cluster structure like α + d in ^{6}Li or α + t in ^{7}Li at large cluster separations.

p-shell nuclei: Pudliner at al., Phys. Rev. Lett. 74 (1995) 4396-4399

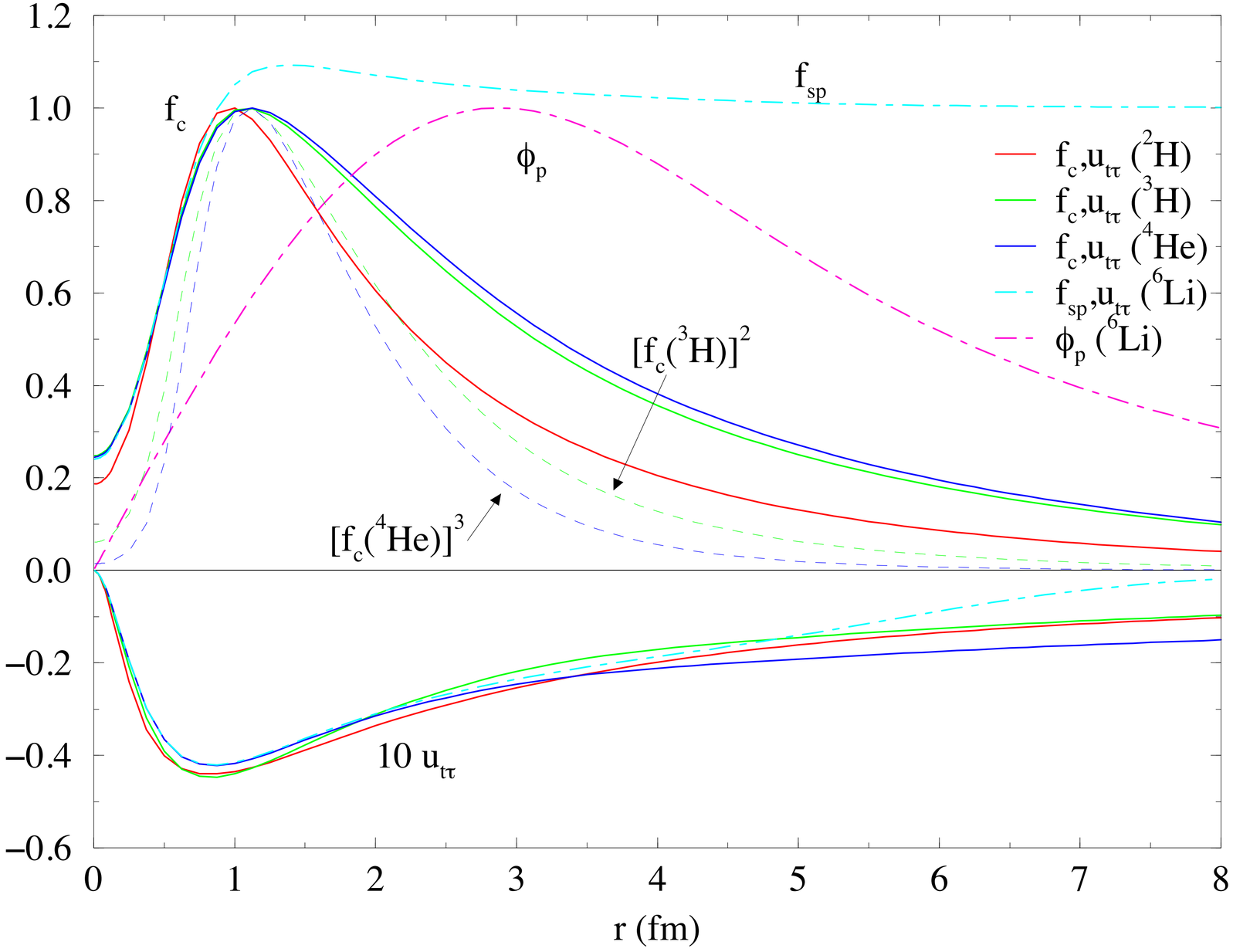
- The *LS* coupling scheme is used to obtain the desired *JM* value of a given state, as suggested by the shell-model studies of light p-shell nuclei. Different possible *LS* combinations lead to multiple components in the Jastrow wave function.

$$|\Psi_J\rangle = \mathscr{A}\left\{\prod_{i<j<k} f^c_{ijk} \prod_{i<j\leq 4} f_{ss}(r_{ij}) \sum_{LS[n]} \beta_{LS[n]} \prod_{k\leq 4<l\leq A} f_{sp}(r_{kl}) \prod_{4<l<m\leq A} f^{[n]}_{pp}(r_{lm}) \; |\Phi_A(LS[n]JJ_zT_z)_{1234:5\ldots A}\rangle\right\}$$

$$\left|\Phi_\alpha(0000)_{1234} \prod_{4<l\leq A} \phi_p^{LS[n]}(R_{\alpha l}) \left\{[Y_1^{m_l}(\Omega_{\alpha l})]_{LM_L} \otimes [\chi_l(\tfrac{1}{2}m_s)]_{SM_S}\right\}_{JM} [\nu_l(\tfrac{1}{2}t_3)]_{TT_3}\right\rangle$$

- The $f_{pp}(r_{lm})$ is set to give the appropriate cluster structure outside the α particle core, for example is similar to the deuteron (triton) $f_c(r_{ij})$ in the case of ^{6}Li (^{7}Li).
- $\phi_p^{LS[n]}(R_{\alpha l})$ are p-wave solutions of a particle in an effective α + N Woods-Saxon potential and are functions of the distance between the center of mass of the α core and nucleon l; they may be different for different LS[n] components.
- Except for closed-shell nuclei, the complete trial wave function is constructed by taking a linear set of states $\beta_{LS[n]}$ with the same total angular momentum and parity. Typically these correspond to the lowest shell-model states of the system.

Correlation functions in light nuclei:



- The $f_c(r)$ is small at short distances, to reduce the contribution of the repulsive core of v_{ij} , and peaks at an intermediate distance corresponding to the maximum attraction of v_{ij} .
- The noncentral $u_p(r)$ are all relatively small; the most important is the long-range tensor-isospin part $u_{t\tau}(r)$, induced by the OPEP
- The $f_{sp}(r)$, shown in Fig. for ^{6}Li nuclei, is similar to the $f_c(r)$ at short range, but with a long-range tail going to unity; this helps the wave function factorize to a cluster structure like α + d in ^{6}Li at large cluster separations.

Computational implementation:

- Since the nuclear interaction is spin-isospin dependent, the trial state is a sum of complex amplitudes for each spin-isospin state of the system: $|\Psi_T\rangle = \sum_{i_s \le n_s, i_t \le n_t} a(i_s, i_t; \mathbf{R}) \, |\chi_{i_s} \nu_{i_t}\rangle \, .$

- The $n_s = 2^A$ many-body spin states can be written as $|\chi_1\rangle = |\downarrow_1, \downarrow_2, \ldots, \downarrow_A\rangle$, $|\chi_2\rangle = |\uparrow_1, \downarrow_2, \ldots, \downarrow_A\rangle$, $|\chi_3\rangle = |\downarrow_1, \uparrow_2, \ldots, \downarrow_A\rangle, \ldots, |\chi_{n_s}\rangle = |\uparrow_1, \uparrow_2, \ldots, \uparrow_A\rangle$ and the isospin ones can be recovered by replacing $\downarrow$ with n and $\uparrow$ with p.

- To construct the trial state, one starts from the mean-field component $|\Phi_A(LS[n]JJ_zT_z)_{1234:5\ldots A}\rangle$. For fixed spatial coordinates $\mathbf{R}$, the spin-isospin independent correlations needed to retrieve $|\Phi_J\rangle$ are simple multiplicative factors, common to all spin amplitudes. The symmetrized product of pair correlation operators is evaluated by successive operations for each pair, sampling their ordering.

NOTE: sampling the order of the pairs help reducing the computational cost since the number of possible orders is $P!$, where $P = 1/2\, A(A-1)$ is the number of pairs. This introduces relatively little statistical variance, because the different orders contain the same linear terms and differ only at $O(u_p^2)$ and above.

- Consider the application of the operator $\sigma_1 \cdot \sigma_2$ on a three-body spin state (for simplicity we neglect the isospin components). Noting that $\sigma_i \cdot \sigma_j = 2\,\mathscr{P}^{\sigma}_{ij} - 1$ exchanges the spin of particles i and j, we obtain:

$$\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 \begin{pmatrix} a_{\uparrow\uparrow\uparrow} \\ a_{\uparrow\uparrow\downarrow} \\ a_{\uparrow\downarrow\uparrow} \\ a_{\uparrow\downarrow\downarrow} \\ a_{\downarrow\uparrow\uparrow} \\ a_{\downarrow\uparrow\downarrow} \\ a_{\downarrow\downarrow\uparrow} \\ a_{\downarrow\downarrow\downarrow} \end{pmatrix} = \begin{pmatrix} a_{\uparrow\uparrow\uparrow} \\ a_{\uparrow\uparrow\downarrow} \\ 2a_{\downarrow\uparrow\uparrow} - a_{\uparrow\downarrow\uparrow} \\ 2a_{\downarrow\uparrow\downarrow} - a_{\uparrow\downarrow\downarrow} \\ 2a_{\uparrow\downarrow\uparrow} - a_{\downarrow\uparrow\uparrow} \\ 2a_{\uparrow\downarrow\downarrow} - a_{\downarrow\uparrow\downarrow} \\ a_{\downarrow\downarrow\uparrow} \\ a_{\downarrow\downarrow\downarrow} \end{pmatrix}$$

- Most of the computing time is spent on spin-isospin operations like the one just described.
- They amount to an iterative sequence of large sparse complex matrix multiplications which mainly rely on three useful matrices:
  - $m(i, i_s)$: z-component of the spin of particle i associated to the many-body spin-state i_s.
  - $n_{\text{exch}}(k_{ij}, i_s)$: number of the many-body spin state obtained by exchanging the spins of particles i and j, belonging to the pair labeled k_{ij} in the state i_s
  - $n_{\text{flip}}(i, i_s)$: the number of the spin state obtained by flipping the spin of particle i in the spin state

- The action of $\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2$ can be expressed as:

$$\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 \sum_{i_s,i_t} a(i_s, i_t; \mathbf{R}) \, |\chi_{i_s} \chi_{i_t}\rangle = \sum_{i_s,i_t} \left[2a(i_s, i_t; \mathbf{R}) - a(n_{\text{exch}}(k_{ij}, i_s), i_t; \mathbf{R})\right] |\chi_{i_s} \chi_{i_t}\rangle$$

subroutine for $\sigma_i \cdot \sigma_j$

```
subroutine sigdotsig( cwvout,cwvin,i,j)
complex(kind=kind(0.d0)),dimension(0:nspin0m,ntau0)
do is=0,nspin0m
iex=ispex(is,i,j)  ! exchange spins i and j in is,
cwvout(is,:)= 2.d0*cwvin(iex,:)-cwvin(is,:)
enddo
end subroutine
```

- By utilizing this representation, we only need to evaluate 2^A operations for each pair, instead of the $2^A \times 2^A$ operations that are required using a simple matrix representation in spin space.
- The tensor operator is slightly more complicated to evaluate and requires both matrices $m(i, i_s)$ and $n_{\text{flip}}(i, i_s)$.

Expectation values:

- The expectation values contain multi-dimensional integrals over all particle positions, $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_A)$,

$$\langle \mathcal{O} \rangle = \frac{\int d\mathbf{R} \Psi_T^\dagger(\mathbf{R}) \mathcal{O} \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T^\dagger(\mathbf{R}) \Psi_T(\mathbf{R})}$$

  where for simplicity the sum over the spin and isospin states is implied.

- Stochastic integration methods are useful to calculate such integrals over many variables. They are based on the central limit theorem of Riemann integrals, which asserts that:

$$\int \mathcal{I}(\mathbf{R}) d\mathbf{R} = \int \frac{\mathcal{I}(\mathbf{R})}{\mathcal{W}(\mathbf{R})} \mathcal{W}(\mathbf{R}) d\mathbf{R} = \left[\frac{1}{N_c} \sum_{I=1}^{N_c} \frac{\mathcal{I}(\mathbf{R}_I)}{\mathcal{W}(\mathbf{R}_I)} \right] \int \mathcal{W}(\mathbf{R}) d\mathbf{R} \quad \text{for} \quad N_c \to \infty$$

  - $\mathcal{I}(\mathbf{R})$ is the integrand, $\mathcal{W}(\mathbf{R})$ is a suitably chosen, normalizable, positive, real function of $\mathbf{R}$ called weight function (representing a probability distribution)
  - The N_c configurations $\mathbf{R}_{I=1,N_c}$ are distributed accordingly with the probability $\mathcal{W}(\mathbf{R})$
  - We sample points $\mathbf{R}_I$ from $\mathcal{W}(\mathbf{R})$ and evaluate $\frac{\mathcal{I}(\mathbf{R}_I)}{\mathcal{W}(\mathbf{R}_I)}$ for each point
  - The error typically goes like $1/\sqrt{N_c}$, and depends critically upon the choice of $\mathcal{W}(\mathbf{R})$

The Metropolis algorithm:

- To obtain the configurations $\mathbf{R}_{I=1,N_c}$ distributed accordingly with the probability $\mathscr{W}(\mathbf{R})$. These configurations are called "samples of $\mathscr{W}(\mathbf{R})$" (see: Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E., "Equations of state calculations by fast computing machines", J. Chem. Phys. 2121(6) 1087 (1953)).

- Based on a Markov Chain Monte Carlo algorithm where a Markov process is a random walk with a selected probability for making a move. The new move is independent of the previous history of the system.
- The reason for choosing a Markov process is that when it is run for a long enough time (steps) starting with a random state, we will eventually reach the most likely state of the system.

- The algorithm satisfies the detailed balance equation:

$$\mathscr{W}(\mathrm{R}_i)T(\mathrm{R}_i \rightarrow \mathrm{R}_j) = \mathscr{W}(\mathrm{R}_j)T(\mathrm{R}_j \rightarrow \mathrm{R}_i)$$

- A simple solution of the equation above is given by $T(\mathbf{R}_i \rightarrow \mathbf{R}_j) = min\left[1, \frac{\mathscr{W}(\mathbf{R}_j)}{\mathscr{W}(\mathbf{R}_i)}\right]$
  - For instance if $\mathscr{W}(\mathbf{R}_i) > \mathscr{W}(\mathbf{R}_j)$ then $T(\mathbf{R}_i \rightarrow \mathbf{R}_j) = \frac{\mathscr{W}(\mathbf{R}_j)}{\mathscr{W}(\mathbf{R}_i)}$ and $T(\mathbf{R}_j \rightarrow \mathbf{R}_i) = 1$ and the detailed balanced equation is satisfied

The Metropolis algorithm: Acceptance-Rejection method

$$\int \mathscr{I}(\mathbf{R})d\mathbf{R} = \int \frac{\mathscr{I}(\mathbf{R})}{\mathscr{W}(\mathbf{R})}\mathscr{W}(\mathbf{R})d\mathbf{R} = \left[\frac{1}{N_c}\sum_{I=1}^{N_c}\frac{\mathscr{I}(\mathbf{R}_I)}{\mathscr{W}(\mathbf{R}_I)}\right]\int \mathscr{W}(\mathbf{R})d\mathbf{R} \quad \text{for} \quad N_c \to \infty$$

- Consider $\mathbf{R}_I$ being the Ith configuration. To obtain the next move in the walk we make a random step $\mathbf{R}' = \underbrace{\mathbf{R} + \Delta\mathrm{R}}$ and accept it with probability $T(\mathrm{R}_I \to \mathrm{R}') = min\left[1, \frac{\mathscr{W}(\mathrm{R}')}{\mathscr{W}(\mathrm{R}_I)}\right]$.

  $x_i' = x_{i,I} + (\epsilon_i - 0.5)l$ for $i = x, y, z$ coordinates of the A particles, ϵ_i random number uniformly distributed in the interval 0 and 1.

  - If the step is accepted then $\mathbf{R}_{I+1} = \mathbf{R}'$, if rejected $\mathbf{R}_{I+1} = \mathbf{R}_I$

- Practically generate a random number s uniformly distributed in the interval 0 and 1. If $w = \frac{\mathscr{W}(\mathrm{R}')}{\mathscr{W}(\mathrm{R}_I)} \geq s$ then the step is accepted otherwise is rejected

- In order to maximize the efficiency of the algorithm, you are shooting for an acceptance rate of 25-50%. The size step l is chosen to satisfy this condition.

- The configurations generated by a Metropolis walk are strongly correlated: a number of substeps are often made to obtain the configuration $\mathbf{R}_{I+1}$ from $\mathbf{R}_I$. Note the initial configurations of the walk are dependent upon the initial value of $\mathbf{R}$ where the walk started; hence the first few hundred steps of the walk are generally discarded.

Expectation values:

- In our case we need to stochastically evaluate:

$$\langle \mathcal{O} \rangle = \frac{\sum_{p,q} \int d\mathbf{R} \frac{\Psi^{\dagger}_{T,p}(\mathbf{R}) \mathcal{O} \Psi_{T,q}(\mathbf{R})}{W_{pq}(\mathbf{R})} W_{pq}(\mathbf{R})}{\sum_{p,q} \int d\mathbf{R} \frac{\Psi^{\dagger}_{T,p}(\mathbf{R}) \Psi_{T,q}(\mathbf{R})}{W_{pq}(\mathbf{R})} W_{pq}(\mathbf{R})} = \frac{\sum_{p,q} \sum_{I=1}^{N_c} \frac{\Psi^{\dagger}_{T,p}(\mathbf{R}_I) \mathcal{O} \Psi_{T,q}(\mathbf{R}_I)}{W_{pq}(\mathbf{R}_I)}}{\sum_{p,q} \sum_{I=1}^{N_c} \frac{\Psi^{\dagger}_{T,p}(\mathbf{R}_I) \Psi_{T,q}(\mathbf{R}_I)}{W_{pq}(\mathbf{R}_I)}} \quad \text{for} \quad N_c \to \infty$$

- The subscripts p and q specify the order of operators in the left- and right-hand-side wave functions. Again a complete sum over all spin-isospin variables is implied.
- In VMC calculations, one usually takes $W_{pq}(\mathbf{R}) = |\Re(\Psi^{\dagger}_{T,p}(\mathbf{R}) \Psi_{T,q}(\mathbf{R}))|$. The Metropolis algorithm produces a set of configurations $[\mathbf{R}, p, q]$ whose density is proportional to $W_{pq}(\mathbf{R})$.

- Expectation values have a statistical error which can be estimated by the standard deviation σ:

$\sigma = \left[\frac{\langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2}{N_c - 1} \right]^{1/2}$ where N_c is the number of statistically independent samples. Block averaging schemes can be used to estimate the autocorrelation times and determine the statistical error.

Expectation values:

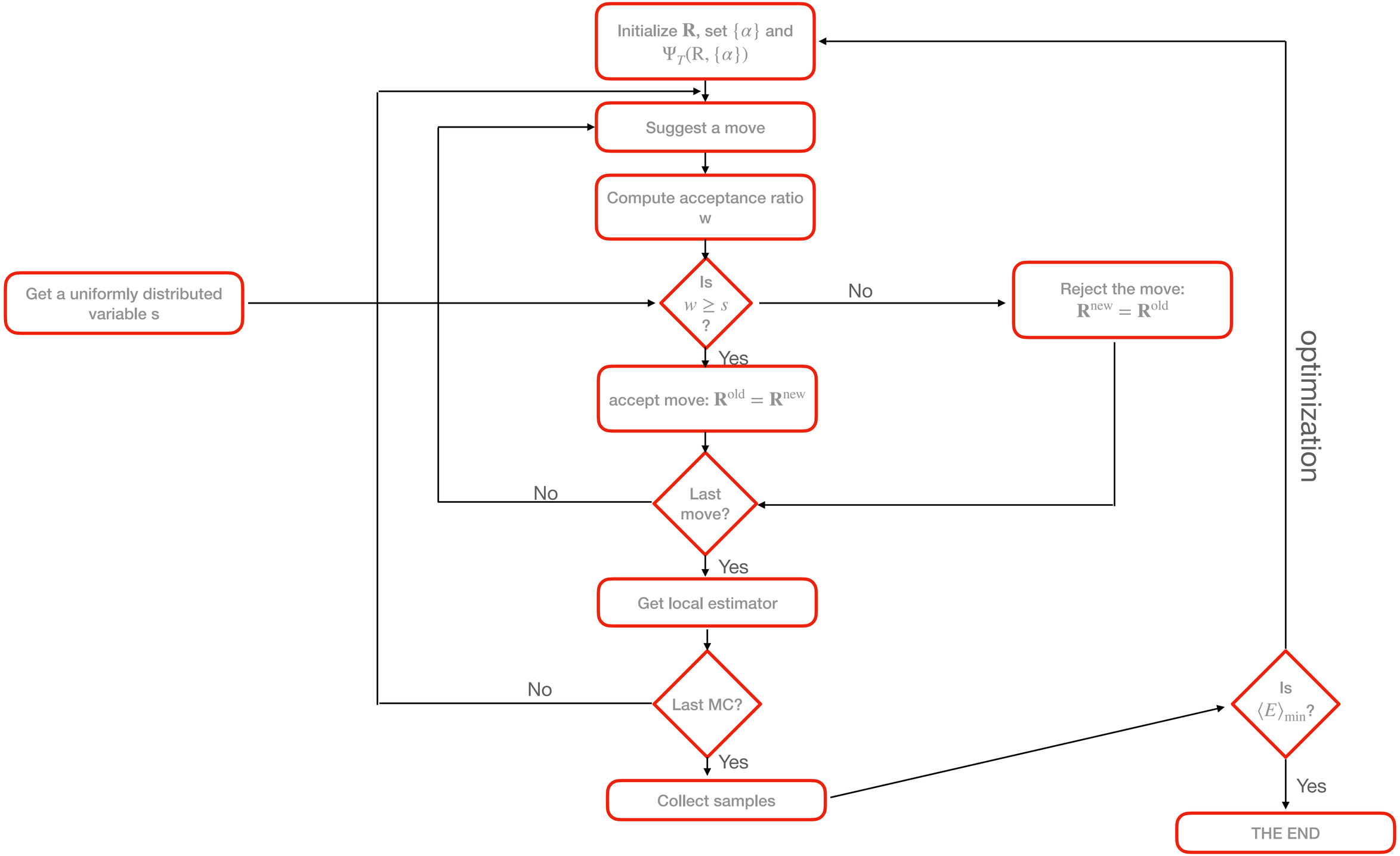
- The expectation values are evaluated by having the operators act entirely on the right hand side of $\Psi_T(\mathbf{R})$. The matrix machinery used to apply the spin-dependent correlation operators is also used to evaluate $\mathcal{O}\,|\Psi_{T,p}\rangle$. A simple scalar product of this quantity with $\langle\Psi_{T,q}\,|$, provides the numerator of the local estimate $\Psi^{\dagger}_{T,q}(\mathbf{R})\mathcal{O}\Psi_{T,p}(\mathbf{R})/W_{pq}(\mathbf{R})$.

- The first and second derivatives of the wave function are numerically computed by means of the two- and three-point stencil, respectively. To determine the kinetic energy, $\sim$ 6A evaluations of $\Psi_T(\mathbf{R})$ are needed.

- Using some tricks, the action of the square of angular momentum terms in the potential can be evaluated on $\Psi_T(\mathbf{R})$ an additional 3A(A-1)/2 times.

A simplified VMC calculation:

1. Initialization: Fix the number of Monte Carlo steps and thermalization. Choose an initial $\mathbf{R}$ and variational parameters $\{\alpha\}$ and calculate the weight function $W(\mathbf{R})$. Define also the value of the stepsize to be used when moving from one value of $\mathbf{R}$ to a new one.
2. Initialize the energy and the variance.
3. Start the Monte Carlo calculation with a loop over a given number of Monte Carlo cycles
   (a) Calculate a trial position $\mathbf{R}' = \mathbf{R} + \Delta\mathbf{R}$ where $\Delta\mathbf{R}$ depends on a random variable $\epsilon \in [0, 1]$ and stepsize
   (b) Use then the Metropolis algorithm to accept or reject this move by calculating the ratio $w = W(\mathbf{R}')/W(\mathbf{R})$. If $w \geq s$, where s is a random number s $\in$ [0, 1], the new position is accepted, else we stay at the same place.
   (c) If the step is accepted, then we set $\mathbf{R} = \mathbf{R}'$.
   (d) Update the local estimator and the variance.
4. When the Monte Carlo sampling is finished, we calculate the mean and the standard deviation.
5. Finally, we may print our results to a specified file.


Initialize $\mathbf{R}$, set $\{\alpha\}$ and $\Psi_T(\mathrm{R}, \{\alpha\})$
Suggest a move
Compute acceptance ratio w
Get a uniformly distributed variable s
Is $w \geq s$?
No
Reject the move: $\mathbf{R}^{\mathrm{new}} = \mathbf{R}^{\mathrm{old}}$
Yes
accept move: $\mathbf{R}^{\mathrm{old}} = \mathbf{R}^{\mathrm{new}}$
No
Last move?
Yes
Get local estimator
No
Last MC?
Yes
Collect samples
Is $\langle E \rangle_{\mathrm{min}}$?
optimization
Yes
THE END

Scaling of VMC calculation time with the nucleus

Scales with # particles (6A w.f. calculations for kinetic energy) × # pairs (operations to construct w.f.) × spin×isospin (size of w.f. vector):

	A	Pairs	Spin×Isospin	$\prod(/^8\mathrm{Be})$
^{4}He	4	6	16×2	0.001
^{5}He	5	10	32×5	0.010
^{6}Li	6	15	64×5	0.036
^{7}Li	7	21	128×14	0.33
^{8}Be	8	28	256×14	1.
^{9}Be	9	36	512×42	8.7
^{10}Be	10	45	1024×90	52.
^{11}B	11	55	2048×132	200.
^{12}C	12	66	4096×132	530.
^{14}C	14	91	16384×1001	26,000.
^{16}O	16	120	65536×1430	220,000.
^{40}Ca	40	780	$1.1\times10^{12}\times6.6\times10^{9}$	2.8×10^{20}
^{8}n	8	28	256×1	0.071
^{14}n	14	91	16384×1	26.

Green's Function Monte Carlo

- The Green's function Monte Carlo (GFMC) overcomes the limitations of VMC by using a projection technique to determine the true ground-state: $|\Psi_0\rangle \equiv lim_{\tau\to\infty}|\Psi(\tau)\rangle = lim_{\tau\to\infty}e^{-(H-E_T)\tau}|\Psi_T\rangle$
  - τ is the imaginary time, and E_T is a parameter used to control the normalization of the evolved state
  - excited states can be computed within GFMC since imaginary-time diffusion yields to the lowest-energy eigenstate with the same quantum numbers as $|\Psi_T\rangle$

- The method relies on the observation that the trial wave function can be expanded in the complete set of eigenstates of H, $|\Psi_n\rangle$ with eigenvalues E_n: $|\Psi_T\rangle = \sum_n c_n|\Psi_n\rangle$ with $H|\Psi_n\rangle = E_n|\Psi_n\rangle$.
  - $\lim_{\tau\to\infty} e^{-(H-E_T)\tau}|\Psi_T\rangle = \lim_{\tau\to\infty}\left(\sum_n c_n e^{(E_T-E_n)\tau}|\Psi_n\rangle\right) = c_0 e^{(E_T-E_0)\tau}|\Psi_0\rangle$ since $E_T - E_0 > E_T - E_{i>0}$.
  - When $E_T = E_0$ than the norm of $\Psi(\tau\to\infty)$ is c_0^2 independent of τ

- Except for some specific cases, the direct computation of the propagator $e^{-H\tau}$ for arbitrary values of τ is typically not possible.

Imaginary short-time propagation

- For small imaginary times $\delta\tau = \tau/N$ with N large, the calculation is tractable, and the full propagation to large imaginary times τ can be recovered through the following path integral (neglecting spin-isospin indices)

$$\Psi(\tau, \mathbf{R}_N) = \int \prod_{i=0}^{N-1} \mathrm{d}\mathbf{R}_i \, \langle \mathbf{R}_N | e^{-(H-E_T)\delta\tau} | \mathbf{R}_{N-1} \rangle \cdots \langle \mathbf{R}_1 | e^{-(H-E_T)\delta\tau} | \mathbf{R}_0 \rangle \langle \mathbf{R}_0 | \Psi_T \rangle$$

- We defined the short-time propagator, or Green's function, $G_{\delta\tau}(\mathbf{R}_{i+1}, \mathbf{R}_i) = \langle \mathbf{R}_{i+1} | e^{-H\delta\tau} | \mathbf{R}_i \rangle$

$$\Psi(\tau, \mathbf{R}_N) = \int \mathrm{d}\mathbf{R}_{N-1} \ldots \mathrm{d}\mathbf{R}_2 \mathrm{d}\mathbf{R}_1 \mathrm{d}\mathbf{R}_0 G_{\delta\tau}(\mathbf{R}_N, \mathbf{R}_{N-1}) \ldots G_{\delta\tau}(\mathbf{R}_2, \mathbf{R}_1) G_{\delta\tau}(\mathbf{R}_1, \mathbf{R}_0) \Psi_V(\mathbf{R}_0)$$

- The wave function at imaginary time $\tau + \delta\tau$ can be written in an integral form

$$\Psi(\tau + \delta\tau, \mathbf{R}_{i+1}) = \int d\mathbf{R}_i G_{\delta\tau}(\mathbf{R}_{i+1}, \mathbf{R}_i) \Psi(\tau, \mathbf{R}_i)$$

- The short-time propagator should allow as large a time step $\delta\tau$ as possible, because the total computational time for propagation is proportional to $1/\delta\tau$

Imaginary short-time propagation: Trotter-Suzuki expansion

- A common approximation for the short-time propagator is based upon the Trotter-Suzuki expansion:

$$G_{\delta\tau}(\mathbf{R}_{i+1}, \mathbf{R_i}) = e^{-V(\mathbf{R}_{i+1})\delta\tau/2}\langle \mathbf{R}_{i+1} \,|\, e^{-T\delta\tau} \,|\, \mathbf{R}_i\rangle e^{-V(\mathbf{R}_i)\delta\tau/2} + o(\delta\tau^3)$$

- T is the kinetic energy giving rise to the free-particle propagator; for non-relativistic systems, it can be expressed as a simple Gaussian in configuration space (this is easy to see it in 1D)

$$\langle \mathbf{R}_{i+1} \,|\, e^{-T\delta\tau} \,|\, \mathbf{R}_i\rangle = G^0_{\delta\tau}(\mathbf{R_{i+1}}, \mathbf{R}_i) = \left[\frac{1}{\lambda^3\pi^{3/2}}\right]^A e^{-(\mathbf{R}_{i+1}-\mathbf{R}_i)^2/\lambda^2} \text{ with } \lambda^2 = 4\frac{\hbar^2}{2m}\delta\tau$$

- The exponentials of the two-body potentials (difficult to calculate directly in nuclei since the interactions do not commute) can be approximated to first order by turning the sums over pairs (remember that $V(\mathbf{R}) \equiv \sum_{i<j} v_{ij} = \sum_{i<j}\sum_p v_p(r_{ij})O^p_{ij}$ and we are neglecting three nucleon potential for now) in the exponent into a symmetrized product of exponentials of the individual pair potentials

$$\langle \mathbf{R} \,|\, e^{-V\delta\tau/2} \,|\, \mathbf{R}\rangle \sim \mathcal{S}\prod_{i<j}\exp\left(-\frac{\delta\tau}{2}\sum_{p=1,6} v_p(r_{ij})O^p_{ij}\right) \text{ where } \mathcal{S} \text{ indicates a symmetrization over orders of pairs.}$$

Comments about the Trotter-Suzuki expansion:

NOTE:
- The first six terms of the potential can be easily exponentiated since they are diagonal in coordinate space.
- Momentum dependent terms cannot be treated this way: a way is to expand the exponential of the momentum terms to first order in $\delta\tau$ and use integration by parts to let the derivatives act on the free-particle Green's function. This is applied to the terms in the potential that are linear in momentum, such as $\mathbf{L}\cdot\mathbf{S}$ and $(\mathbf{L}\cdot\mathbf{S})\,\tau_{\mathbf{i}}\cdot\tau_{\mathbf{j}}$
- Terms that are quadratic in the momentum cannot be evaluated to first order in this manner.

- We use approximations to the full NN potentials, such as the AV8' interaction, projection on the first eight operators (reproduce phase shifts in $^1S_0$, $^1P_1$, ^3P_J, and $^3S_1 - ^3D_1$) . (Note: The error due to this approximation is estimated by direct comparison with Faddeev calculations with AV8' and AV18. It seems to be less than 0.3%).
- The isoscalar part of the Coulomb interaction can also be easily included in the propagator; however the small isospin breaking terms of the electromagnetic and strong interactions are treated perturbatively.

- The main error in this Trotter-Suzuki expansion approximation comes from terms in $e^{-H\tau}$ having multiple $v_{i,j}$ —like $v_{i,j}Tv_{i,j}(\delta\tau)^3$— where T is the kinetic energy, which can become large when particles i and j are very close due to the large repulsive core in v_{ij} . This requires a rather small $\delta\tau$.

Imaginary short-time propagation: exact two-body propagator

- It has been shown that that including the exact two-body propagator allows much larger time steps (a factor of 5-10 larger time steps $\delta\tau$) than the simple approximation, so the computational time will be reduced.

- The two-body interactions are replaced with the exact two-body Green's function g_{ij}.

- The most common one (the Schmidt-Lee many-body Green's function) consists in building the Green's function operator as a product of exact two-body propagators

$$G_{\delta\tau}(\mathbf{R}_{i+1}, \mathbf{R}_i) = \left(\mathcal{S} \prod_{j<k} \frac{g_{jk}(\mathbf{r}_{jk,i}, \mathbf{r}_{jk,i+1})}{g^0_{jk}(\mathbf{r}_{jk,i}, \mathbf{r}_{jk,i+1})} \right) G^0_{\delta\tau}(\mathbf{R}_{i+1}, \mathbf{R}_i)$$

- $g_{jk}(\mathbf{r}_{jk,i}, \mathbf{r}_{jk,i+1}) = \langle \mathbf{r}_{jk,i} \,|\, e^{-H_{jk}\delta\tau} \,|\, \mathbf{r}_{jk,i+1} \rangle$ is the exact two-body propagator where $H_{jk} = -\frac{1}{m}\nabla^2_{jk} + v_{jk}$
- $g^0_{jk}(\mathbf{r}_{jk,i}, \mathbf{r}_{jk,i+1})$ is the two-body free-particle propagator

Imaginary short-time propagation: exact two-body propagator

- Terms quadratic in the angular momentum can in principle be accounted for into the exact pair propagator. However, the Monte Carlo sampling can lead to large variance. Thus, simplified AV8' potentials are also used in the pair propagator, even though in this case no approximations in treating $\mathbf{L}\cdot\mathbf{S}$ and $(\mathbf{L}\cdot\mathbf{S})\,\boldsymbol{\tau}_i\cdot\boldsymbol{\tau}_j$ terms are necessary.

- The 3N interaction V_{ijk} is included symmetrically, and the full propagation for each step

$$G_{\delta\tau}(\mathbf{R}_{i+1},\mathbf{R}_i)=\left(1-\sum_{i<j<k}V_{ijk}(\mathbf{R}_{i+1})\right)\left(\mathcal{S}\prod_{j<k}\frac{g_{jk}(\mathbf{r}_{jk,\,i},\mathbf{r}_{jk,\,i+1})}{g^0_{jk}(\mathbf{r}_{jk,\,i},\mathbf{r}_{jk,\,i+1})}\right)G^0_{\delta\tau}(\mathbf{R_{i+1}},\mathbf{R_i})\left(1-\sum_{i<j<k}V_{ijk}(\mathbf{R}_i)\right)$$

Mixed estimates:

- One we have defined the propagator, we will be interested to calculate expectation values of operators:

$$\langle \mathcal{O} \rangle = \frac{\langle \Psi(\tau) \,|\, \mathcal{O} \,|\, \Psi(\tau) \rangle}{\langle \Psi(\tau) \,|\, \Psi(\tau) \rangle} \qquad \tau \to \infty$$

- These are difficult to calculate, so we will use mixed estimates

$$\langle \mathcal{O} \rangle_M = \frac{\langle \Psi_T \,|\, \mathcal{O} \,|\, \Psi(\tau) \rangle}{\langle \Psi_T \,|\, \Psi(\tau) \rangle} = \frac{\int d\mathbf{R}_N \Psi_T(\mathbf{R}_N) \mathcal{O} \Psi(\tau, \mathbf{R}_N)}{\int d\mathbf{R}_N \Psi_T(\mathbf{R}_N) \Psi(\tau, \mathbf{R}_N)} = \frac{\int d\mathbf{P}_N \Psi_T(\mathbf{R_N}) \mathcal{O} \mathbf{G}_{\delta\tau}(\mathbf{R_N}, \mathbf{R_{N-1}}) \ldots \mathbf{G}_{\delta\tau}(\mathbf{R_1}, \mathbf{R_0}) \mathbf{\Psi_T}(\mathbf{R_0})}{\int d\mathbf{P}_N \Psi_T(\mathbf{R}_N) G_{\delta\tau}(\mathbf{R}_N, \mathbf{R}_{N-1}) \ldots G_{\delta\tau}(\mathbf{R}_1, \mathbf{R}_0) \Psi_T(\mathbf{R}_0)}$$

where we are always considering $\tau \to \infty$

- The set of configurations $\mathbf{P}_N = \mathbf{R}_N, \ldots, \mathbf{R}_1, \mathbf{R}_0$ make a path in the 3A dimensional configuration space and the integrals over $\mathbf{P}_N$ are carried out using stochastic methods.

Mixed estimates:

- In practice, a set of configurations, typically called walkers, are simultaneously evolved in imaginary time, and then used to calculate observables once convergence is reached.

- In the GFMC method, each walker contains the nucleon positions and a complex amplitude for each spin/isospin state of the nucleus, implying an unfavorable exponential scaling with the number of nucleons.

- Since the Hamiltonian commutes with the imaginary time propagator, we can

$\langle H \rangle_M = \dfrac{\langle \Psi_T | H | \Psi(\tau) \rangle}{\langle \Psi_T | \Psi(\tau) \rangle} = \dfrac{\langle \Psi(\tau/2) | H | \Psi(\tau/2) \rangle}{\langle \Psi(\tau/2) | \Psi(\tau/2) \rangle} = E(\tau/2)$ where $E(\tau/2)$ is the expectation of H in the state $\Psi(\tau/2)$. It is therefore $\geq E_0$ and approaches E_0 from above as $\tau \to \infty$.

Mixed estimates: operators

- Let consider $\delta\Psi = \Psi(\tau) - \Psi_T$; we can re-write the expectation value of $\langle \mathcal{O} \rangle$ as

$$\langle \mathcal{O} \rangle = 2\frac{\langle \Psi_T | \mathcal{O} | \Psi(\tau) \rangle}{\langle \Psi_T | \Psi(\tau) \rangle} - \frac{\langle \Psi_T | \mathcal{O} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} + \delta\Psi^2$$

  where the the second term is the variational matrix element.

- When $\delta\Psi$ is small, therefore the difference between the mixed and the variational estimates is small, we can use this equation to calculate the ground-state expectation value neglecting $\delta\Psi^2$ term.

- If the $\langle \mathcal{O} \rangle_M$ and $\langle \mathcal{O} \rangle_V$ are significantly different so that the extrapolation may not be valid, Ψ_T is clearly poor.

Sampling of the paths:

- The path integrals in $\langle \mathcal{O} \rangle_M$ are evaluated stochastically. Consider an ensemble of paths denoted as $\{\mathbf{P}\}$, that contains N_p paths, be sampled by a normalized weight function $P(\mathbf{P}_N)$.
- Each path consists of N steps, where each step contains a sample of 3A particle coordinates (and full spin-isospin state), as well as a sets of operators orders used to sample the symmetrized product for the pair operators in the Ψ_T

- For a given path $\mathbf{P}_N$, consider:

$$\mathcal{N}_{\mathbf{P}_N} = \frac{\Psi_T(\mathbf{R}_N)\mathcal{O}G_{\delta\tau}(\mathbf{R}_N,\mathbf{R}_{N-1})\ldots G_{\delta\tau}(\mathbf{R}_1,\mathbf{R}_0)\Psi_T(\mathbf{R}_0)}{P(\mathbf{P_N})}\,,\ \mathscr{D}_{\mathbf{P}_N} = \frac{\Psi_T(\mathbf{R}_N)G_{\delta\tau}(\mathbf{R}_N,\mathbf{R}_{N-1})\ldots G_{\delta\tau}(\mathbf{R}_1,\mathbf{R}_0)\Psi_T(\mathbf{R}_0)}{P(\mathbf{P_N})}$$

- The average value of $\langle \mathcal{O} \rangle_M$ is given by: $\langle \mathcal{O} \rangle_M = \frac{\mathcal{N}}{\mathscr{D}} = \frac{\sum_{\{\mathbf{P}\}} \mathcal{N}_{\mathbf{P}_N}}{\sum_{\{\mathbf{P}\}} \mathscr{D}_{\mathbf{P}_N}}$ with a statistical error being proportional to $1/\sqrt{N_p}$. Also in this case to improve the statistical error block averaging on $\{\mathbf{P}\}$ is performed.

Sampling of the paths:

- Most calculations use the weight function:

$$P(\mathbf{P}) = \prod_{i=1,N} \left[I(\mathbf{R}_i) G(\mathbf{R}_i, \mathbf{R}_{i-1}) \frac{1}{I(\mathbf{R}_{i-1})} \right] I(\mathbf{R}_0) \, |\Psi_T(\mathbf{R}_0)| = I(\mathbf{R}_N) \prod_{i=1,N} \left[G(\mathbf{R}_i, \mathbf{R}_{i-1}) \right] |\Psi_T(\mathbf{R}_0)|$$

- The importance function $I(\mathbf{R})$ is used in sampling and hence should be positive definite

- The probability of the path $P(\mathbf{P})$ depends implicitly upon all of the steps in the path, but is decomposed into an initial weight $I(\mathbf{R}_0) \, |\Psi_T(\mathbf{R}_0)|$, times a product of weights for each step.

- The idea is that the initial configurations are sampled from $I(\mathbf{R}_0) \, |\Psi_T(\mathbf{R}_0)|$ and the quantity in bracket is referred as the importance sampled Green's function $G_I(\mathbf{R}_i, \mathbf{R}_{i-1}) = \left[I(\mathbf{R}_i) G(\mathbf{R}_i, \mathbf{R}_{i-1}) \frac{1}{I(\mathbf{R}_{i-1})} \right]$

- For a given path $\mathbf{P}_N$, note that: $\mathcal{N}_{\mathbf{P}_N} = \frac{\Psi_T(\mathbf{R}_N) \mathcal{O} \Psi_T(\mathbf{R}_0)}{I(\mathbf{R}_N) \, |\Psi_T(\mathbf{R}_0)|}$ and $\mathcal{D}_{\mathbf{P}_N} = \frac{\Psi_T(\mathbf{R}_N) \Psi_T(\mathbf{R}_0)}{I(\mathbf{R}_N) \, |\Psi_T(\mathbf{R}_0)|}$

Sampling of the paths:

- Implementing the algorithm to sample the paths is straightforward. For simplicity choosing $I(\mathbf{R}) = |\Psi_T(\mathbf{R})|$, the initial ($\tau = 0$) configuration $\mathbf{R}_0$ for each path is obtained as in the VMC by sample $|\Psi_T(\mathbf{R})|^2$ using Metropolis method.

- The subsequent configurations $\mathbf{R}_i$, at $\tau = i\delta\tau$, are obtained sequentially from $\mathbf{R}_{i-1}$, by iterating with the importance-sampled Green's function G_I,

$$I(\mathbf{R}_i)\Psi(\mathbf{R}_i) = \int G_I(\mathbf{R}_i, \mathbf{R}_{i-1}) I(\mathbf{R}_{i-1})\Psi(\mathbf{R}_{i-1}) d\mathbf{R}_{i-1}$$

- Describes the evolution of the density $I(\mathbf{R}_i)|\Psi_T(\mathbf{R}_i)|$ with $\tau = i\delta\tau$, hence the configurations are $\mathbf{R}_i$ distributed with this density.

- This implies that we can sample the points along the path directly from $G_I(\mathbf{R}_i, \mathbf{R}_{i-1})$ but this is typically not possible. One must sample from an approximate $\tilde{G}_I(\mathbf{R}_i, \mathbf{R}_{i-1})$ and then use weighting and branching techniques.

Sampling of the paths:

- If points are sampled from an approximate $\tilde{G}_I(\mathbf{R}_i, \mathbf{R}_{i-1})$, it is convenient to define a weight

$$\tilde{w}(\mathbf{R}_i, \mathbf{R}_{i-1}) = \frac{G_I(\mathbf{R}_i, \mathbf{R}_{i-1})}{\tilde{G}_I(\mathbf{R}_i, \mathbf{R}_{i-1})}$$

- Choosing $P(\mathbf{P}_N) = \prod_{i=1,N} \tilde{G}_I(\mathbf{R}_i, \mathbf{R}_{i-1}) I(\mathbf{R}_0) \,|\, \Psi_T(\mathbf{R}_0) \,|$ will modify expressions for the the numerator and denominator, by multiplying the contribution of each path by the product of the product of $\tilde{w}(\mathbf{R}_i, \mathbf{R}_{i-1})$

$$\mathcal{N}_{\mathbf{P}_N} = W(\mathbf{P}_N) \frac{\Psi_T(\mathbf{R}_N) \mathcal{O} \Psi_T(\mathbf{R}_0)}{I(\mathbf{R}_N) \,|\, \Psi_T(\mathbf{R}_0) \,|}, \qquad \mathcal{D}_{\mathbf{P}_N} = W(\mathbf{P}_N) \frac{\Psi_T(\mathbf{R}_N) \Psi_T(\mathbf{R}_0)}{I(\mathbf{R}_N) \,|\, \Psi_T(\mathbf{R}_0) \,|} \quad \text{with } W(\mathbf{P}_N) = \prod_{i=1,N} \tilde{w}(\mathbf{R}_i, \mathbf{R}_{i-1})$$

Sampling of the paths:

- What we really do is the following:
  - We sample a number of points $\mathbf{R}'_j$ for $j = 1, n_{\text{samp}}$ from $G_0(\mathbf{R}'_j, \mathbf{R}_{i-1})$
  - We define a scalar spin-independent importance sample Green's function $G_I^S(\mathbf{R}_i, \mathbf{R}_{i-1})$, positive and fast to calculate and approximate $G_I(\mathbf{R}_i, \mathbf{R}_{i-1})$
  - For each of the n_{samp} points we calculate $G_I^S(\mathbf{R}'_j, \mathbf{R}_{i-1})$ and the $\mathbf{R}_i$ is picked from the set $\mathbf{R}'_j$ with probability proportional to $\dfrac{G_I^S(\mathbf{R}'_j, \mathbf{R}_{i-1})}{G_0(\mathbf{R}'_j, \mathbf{R}_{i-1})}$
  - This procedure implicitly defines a $\tilde{G}_I$, and requires a weight: $\tilde{w}(\mathbf{R}_i, \mathbf{R}_{i-1}) = \left[\dfrac{1}{n_{\text{samp}}} \sum_{j=1,\text{samp}} \dfrac{G_I^S(\mathbf{R}'_j, \mathbf{R}_{i-1})}{G_0(\mathbf{R}'_j, \mathbf{R}_{i-1})}\right] \dfrac{G_I(\mathbf{R}_i, \mathbf{R}_{i-1})}{G_I^S(\mathbf{R}_i, \mathbf{R}_{i-1})}$
- G_I^S contains approximations to the dominant physics in the propagator and the trial wave function

$$G_I^S(\mathbf{R}, \mathbf{R}') = |\Psi_J(\mathbf{R})| G^S(\mathbf{R}, \mathbf{R}') \frac{1}{|\Psi_J(\mathbf{R}')|}$$

  where G^S is spin-isospin independent interactions average central potential in the S-wave

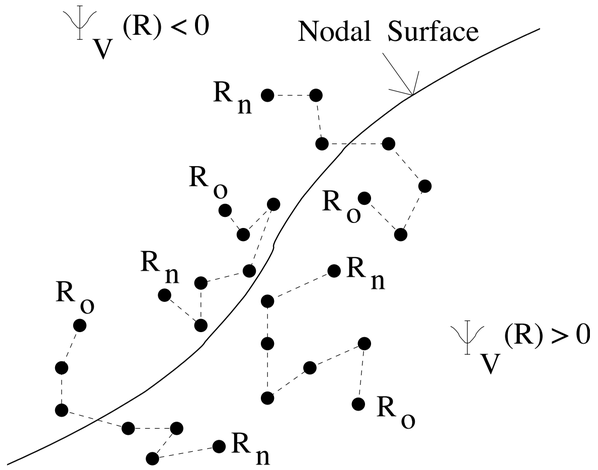
- We use $I[\Psi_T(\mathbf{R}_i, \Psi_i(\mathbf{P}_i)] = \left|\sum_\alpha \Psi^\dagger_{T,\alpha}(\mathbf{R}_i)\Psi_{i,\alpha}(\mathbf{P}_i)\right| + \epsilon \sum_\alpha |\Psi^\dagger_{T,\alpha}(\mathbf{R}_i)\Psi_{i,\alpha}(\mathbf{P}_i)|$

Sign problem:

- As in standard Fermion diffusion Monte Carlo algorithms, the GFMC method suffers from the Fermion sign problem that arises from stochastically evaluating the mixed estimates.

- The imaginary-time propagator is a local operator, but antisymmetry is global property of the system. $|\Psi(\tau)\rangle$ can have bosonic components with lower energy than Fermionic ones, which are exponentially amplified during propagation.

- When the dot product with the antisymmetric Ψ_T is taken, the desired Fermionic component is projected out in the expectation values, but the variance (the statistical error) grows exponentially with τ.



At small τ , few paths are long enough to cross nodal surfaces and the variance is small.
As τ increases, many paths cross nodal surfaces, the variance increases and the average value of $\mathscr{D}$ decreases.

- Because the number of pairs that can be exchanged grows with A, the sign problem also grows exponentially with the number of nucleons.

"contrained-path" method: to discard those configurations that, in future generations, will contribute only noise to expectation values.

- If we knew the exact ground state, we could discard any walker for which $\Psi_0^\dagger(\mathbf{R}_i)\Psi(\tau,\mathbf{R}_i)=0$ where a sum over spin-isospin states is implied.
- The sum of these discarded configurations can be written as a state $|\Psi_d\rangle$, which has zero overlap with the ground state.
- Disregarding $|\Psi_d\rangle$ is justified because it only contains excited-states components and should decay away as $\tau\to\infty$.
- However, in general, the exact ground state is not known, and the constraint is approximately imposed using Ψ_T in place of Ψ_0: $\langle\Psi_T|\Psi_d\rangle=0$

- GFMC trial wave function is a vector in spin-isospin space, and there are no coordinates for which all the spin-isospin amplitudes will vanish: the overlap $\Psi_{T,p}^\dagger(\mathbf{R}_i)\Psi(\tau,\mathbf{R}_i)$ is complex and depends on the particular sampled order p.

- To circumvent these difficulties, we define the overlap $O_{T,p} = \Re[\Psi^{\dagger}_{T,p}(\mathbf{R}_i)\Psi(\tau, \mathbf{R}_i)]$ and introduce a probability for discarding a configuration in terms of the ratio $O_{T,p}/I_p$ where

$$P[\Psi^{\dagger}_{T,p}(\mathbf{R}_i), \Psi(\tau, \mathbf{R}_i)] = \begin{cases} 0 & O/I > \alpha_c \\ \frac{\alpha_C - O/I}{\alpha_c - \beta_c} & \alpha_c > O/I > \beta_c \\ 1 & O/I < \beta_c \end{cases}$$

- According to this algorithm configurations with O/I less than β_c are always discarded, configurations with O/I greater than α_c are never discarded, and there is a linear interpolation in between.
- The constants α_c and β_c are adjusted such that the average of the overlap $O_{T,p}/I_p$ is zero within statistical errors.

- In a few cases the constrained propagation converges to the wrong energy (either above or below the correct energy). Therefore, a small number, $n_u = 10$ to 80, of unconstrained steps are made before evaluating expectation values. These few unconstrained steps appear to be sufficient to remove the bias introduced by the constraint but do not greatly increase the statistical error.

Auxiliary field diffusion Monte Carlo

- The GFMC method works very well for calculating the low lying states of nuclei up to ^{12}C. Its major limitation is that the computational costs scale exponentially with the number of particles, because of the full summations of the spin-isospin states.

- Over the last two decades, the auxiliary field diffusion Monte Carlo method has become a mainstay for studying medium mass atomic nuclei and infinite neutron matter.

- The AFDMC overcomes the exponential scaling with the number of nucleons of the GFMC by:

  1. using a spin-isospin basis given by the outer product of single-nucleon spinors

$$|\chi_{i_s}\,\chi_{i_t}\rangle \rightarrow |S\rangle \equiv |s_1\rangle \otimes |s_2\rangle \otimes \ldots \otimes |s_A\rangle\,,\ |s_i\rangle = a_{i,\uparrow p}|\uparrow p\rangle + a_{i,\downarrow p}|\downarrow p\rangle + a_{i,\uparrow n}|\uparrow n\rangle + a_{i,\downarrow n}|\downarrow n\rangle$$

  where the state vector is fully specified by a set of $4A$ complex coefficients.

NOTE: Single-nucleon spinor:

Example: A=3 $$|\Phi_{A=3}\rangle = \begin{pmatrix} a_{1\uparrow} \\ a_{1\downarrow} \end{pmatrix}_1 \otimes \begin{pmatrix} a_{2\uparrow} \\ a_{2\downarrow} \end{pmatrix}_2 \otimes \begin{pmatrix} a_{3\uparrow} \\ a_{3\downarrow} \end{pmatrix}_3 \quad \text{with} \quad a_{k\uparrow} = {}_k\langle\uparrow|\Phi_{A=3}\rangle \, .$$

Advantage: including the isospin total number of entries for $|\Psi_{A=3}\rangle$ is thus 12 (instead of 24 for the GFMC)

Issue: it is not closed with respect to the application of quadratic spin (isospin) operators

$$\begin{aligned} \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 \Big[& \left(a_{1,\uparrow}|\uparrow\rangle + a_{1,\downarrow}|\downarrow\rangle\right) \otimes \left(a_{2,\uparrow}|\uparrow\rangle + a_{2,\downarrow}|\downarrow\rangle\right) \otimes \left(a_{3,\uparrow}|\uparrow\rangle + a_{3,\downarrow}|\downarrow\rangle\right) \Big] = \\ & = 2 \Big[\left(a_{2,\uparrow}|\uparrow\rangle + a_{2,\downarrow}|\downarrow\rangle\right) \otimes \left(a_{1,\uparrow}|\uparrow\rangle + a_{1,\downarrow}|\downarrow\rangle\right) \otimes \left(a_{3,\uparrow}|\uparrow\rangle + a_{3,\downarrow}|\downarrow\rangle\right) \Big] \\ & - \Big[\left(a_{1,\uparrow}|\uparrow\rangle + a_{1,\downarrow}|\downarrow\rangle\right) \otimes \left(a_{2,\uparrow}|\uparrow\rangle + a_{2,\downarrow}|\downarrow\rangle\right) \otimes \left(a_{3,\uparrow}|\uparrow\rangle + a_{3,\downarrow}|\downarrow\rangle\right) \Big] . \end{aligned}$$

If using the standard DMC algorithm, the imaginary- time propagator generates a sum of single particles wave functions at each time step. The number of these functions will grows very quickly during the imaginary time evolution, destroying the gain in computational time obtained using a smaller multicomponent trial wave function.

- The AFDMC overcomes the exponential scaling with the number of nucleons of the GFMC by:

  2. a simpler trial wave function contains a linearized version of spin/isospin-dependent two-body correlations $|\Psi_T\rangle = \left(1 - \sum_{i<j} F_{ij} - \sum_{i<j<k} F_{ijk}\right)|\Phi_J\rangle$

     - The Jastrow component of $|\Psi_T\rangle$ is also simpler: $|\Phi_J\rangle = \prod_{i<j} f^c_{ij} \prod_{i<j<k} f^c_{ijk} |\Phi_A(J^\pi, J_z, T_z)\rangle$,

     - The mean-field component is modeled by a sum of Slater determinants,

$$\langle X|\Phi(J^\pi, J_z, T_z)\rangle = \sum_n c_n \left[\sum_{JJ_z} C_{JJ_z} \mathscr{A}\left[\phi_{\alpha_1}(x_1)\ldots\phi_{\alpha_A}(x_A)\right]\right]_{JJ_z}$$

       - $X = \{x_1, \ldots, x_A\}$ with $x_i \equiv \{\mathbf{r}_i, s_i\}$ representing both the position $\mathbf{R} = \mathbf{r}_1, \ldots, \mathbf{r}_A$ and the spin-isospin coordinates $S = s_1, \ldots, s_A$ of the A nucleons,
       - the determinants are coupled with Clebsch-Gordan coefficients C_{JJ_z} in order to reproduce the total angular momentum, total isospin, and parity,
       - the single-particle orbitals are given by $\phi_\alpha(x_i) = R_{nl}(r_i)\, Y_{ll_z}(\hat{r}_i)\, \chi_{ss_z}(\sigma)\, \chi_{tt_z}(\tau)$

Imaginary short-time propagation

- Small steps as in the GFMC but now the generalized coordinate X is used instead of $\mathbf{R}$ and the spin-isospin degrees of freedom are also sampled.

- The AFDMC wave function at imaginary time $\tau + \delta\tau$ can be written in an integral form analogous to the GFMC one

$$\Psi(\tau + \delta\tau, X_{i+1}) = \sum_{S_i} \int d\mathbf{R}_i G_{\delta\tau}(X_{i+1}, X_i) \Psi(\tau, X_i)$$

- Using the Trotter decomposition of

$$G_{\delta\tau}(X_{i+1}, X_i) = G^0_{\delta\tau}(\mathbf{R_{i+1}}, \mathbf{R_i}) \langle \mathbf{S_{i+1}} | \, \mathbf{e}^{-(\mathbf{V(R_{i+1})/2 + V(R_i)/2 - E_T})\delta\tau} \, | \mathbf{S_i} \rangle + \mathbf{o}(\delta\tau^3)$$

- In order to preserve the single-particle representation, the short-time propagator is linearized utilizing the Hubbard-Stratonovich transformation

$$e^{-\lambda \mathcal{O}^2 \delta\tau/2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \, e^{-x^2/2} \, e^{x\sqrt{-\lambda\delta\tau}\,\mathcal{O}},$$

where x are the auxiliary fields and and $\mathcal{O}$ can be any type of operator included in the propagator.

NOTE: Hubbard-Stratonovich transformation

- The first six terms defining the NN potential can be conveniently separated in a spin/isospin-dependent V_{SD} and spin/isospin-independent V_{SI} contributions. If we consider $\boldsymbol{\tau}_i \cdot \boldsymbol{\tau}_j = 1$ (neutron systems):

$$V_{SD} = \frac{1}{2} \sum_{i\alpha j\beta} A_{i\alpha,j\beta}\, \sigma_i^\alpha\, \sigma_j^\beta = \frac{1}{2} \sum_{n=1}^{3A} \mathcal{O}_n^2\, \lambda_n\,,$$

  - O_n are defined as $\mathcal{O}_n = \sum_{i,\alpha} \sigma_i^\alpha\, \psi_{i\alpha}^n$ with λ_n and $\psi_{i\alpha}^n$ are the eigenvalues and eigenvectors of the matrix A

- Applying the exponential of the spin-dependent terms of the NN interaction amounts to rotating the spin-isospin states of nucleons

$$e^{-V(\mathbf{R_i})\delta\tau/2}\,|\,S_i\rangle \sim \prod_n \frac{1}{\sqrt{2\pi}} \int dx_n e^{-x_n^2/2} e^{x_n\sqrt{-\lambda_n\delta\tau}\,O_n}\,|\,S_i\rangle$$

- The imaginary-time propagation is performed by sampling the auxiliary fields $\bar{x}_n$ from the Gaussian probability distribution

$$|\,S_{i+1}\rangle \sim \prod_n e^{\bar{x}_n\sqrt{-\lambda\delta\tau}\,O_n}\,|\,S_i\rangle$$