II. VARIATIONAL AND PROJECTOR MONTE CARLO METHODS

Markus Holzmann
LPMMC, UJF, Grenoble, and LPTMC, UPMC, Paris
markus@lptl.jussieu.fr
http://www.lptl.jussieu.fr/users/markus/
(Dated: January 24, 2012)

A. Variational Monte Carlo

Zero Variance Principle. Within VMC, we use Monte Carlo integration methods to sample the probability distribution $\sim |\Psi_T(\mathbf{R})|^2$ of particles at position $\mathbf{R} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$ for an *explicitly known, unnormalized* many-body wave function $\Psi_T(\mathbf{R})$. Expectation values of any operator in position space, $O(\mathbf{R})$, can then be calculated

$$\langle O \rangle_T = \langle O(\mathbf{R}) \rangle_\pi \equiv \frac{1}{Z} \int d\mathbf{R} O(\mathbf{R}) \pi(\mathbf{R}), \quad \pi(\mathbf{R}) = |\Psi_T(\mathbf{R})|^2, \quad Z = \int d\mathbf{R} \pi(\mathbf{R}) \tag{1}$$

The most important observable is the total energy of the system,

$$E_T = \langle H \rangle_T = \langle E_L(\mathbf{R}) \rangle_\pi , \quad E_L(\mathbf{R}) = \frac{H\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \tag{2}$$

where $E_L(\mathbf{R})$ is called *local energy*. According to the variational principle we have to minimize the energy to approach the true ground state energy. Closer to the ground state, we also expect to lower the variance of the local energy

$$\sigma^2_{E,T} = \left\langle (H - E_T)^2 \right\rangle_T = \left\langle [E_L(\mathbf{R}) - E_T]^2 \right\rangle_\pi \tag{3}$$

as $E_T(\mathbf{R}) = E_n$ (constant independent of $\mathbf{R}$) for any exact eigenstate $\Psi_T = \Psi_n$. Therefore, optimization of the trial wave function is in general based on minimizing a combination of energy and variance. As a side-effect of improving the trial wave function by optimization of parameter set is that we also improve the statistical accuracy of the energy, as the error, ϵ_E, of a finite simulation scales as

$$\epsilon_E \approx \sqrt{\frac{\sigma^2_E}{P^*}} \tag{4}$$

where P^* is the number of *uncorrelated* configurations created within the time of the simulation. If we reduce the variance by a factor two using a better wave function, we do not only have improved on getting a lower energy, but we also need only half of the simulation time for the same *statistical* accuracy. This may also compensate the slowing down of the calculations with increasing complexity of more accurate wave functions.

Off-diagonal operators. The sampling of general operators which are local in space and can be written in the form $O(\mathbf{R})$, e.g. the pair correlation function or static structure factor, is straightforward. Sometimes one is also interested in off-diagonal quantities, for example the single particle momentum distribution, $n_\mathbf{k}$,

$$n_\mathbf{k} = \frac{N}{V} \frac{1}{Z} \int d\mathbf{r}'_1 \int d\mathbf{r}_1 \cdots \int d\mathbf{r}_N e^{-i\mathbf{k}(\mathbf{r}'_1 - \mathbf{r}_1)} \Psi^*_T(\mathbf{r}'_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \Psi_T(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \tag{5}$$

which can be obtained by reweighting

$$n_\mathbf{k} = \langle n_\mathbf{k}(\mathbf{R}) \rangle_\pi , \quad n_\mathbf{k}(\mathbf{R}) = \frac{1}{V} \sum_i \int d\mathbf{r}'_i e^{i\mathbf{k}(\mathbf{r}'_i - \mathbf{r}_i)} \frac{\Psi^*_T(\mathbf{r}_1, \ldots, \mathbf{r}'_i, \ldots, \mathbf{r}_N)}{\Psi_T(\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_N)} \tag{6}$$

To calculate $n_\mathbf{k}$ within VMC, one therefore adds an additional particle at $\mathbf{r}'_i$, e.g. uniformely distributed, so that the rhs can be evaluated and add to the expectation value when observables are calculated. The underlying random walk to create $\pi(\mathbf{R})$ is not affected by this reweighting method.

Generalized partition functions. However, if we are interested in off-diagonal quantities of the general structure $\Psi^*_T(\mathbf{R})\Psi_T(\mathbf{R}')$, instead of doing a random walk to create the probability $\pi(\mathbf{R}) \sim |\Psi_T(\mathbf{R})|^2$ and reweight this probability when evaluating the off-diagonal operator, we can also directly perform a random walk in an extended space

$\{\mathbf{R},\mathbf{R}'\}$ to create the probability $\pi_{\rm off}(\mathbf{R}',\mathbf{R}) \sim |\Psi_T(\mathbf{R}')\Psi_T(\mathbf{R})|$ correponding to a partition function $Z_{\rm off}$. Whereas this random walk in extended space reduces the variance of off-diagonal observables, the overall-normalization Z may be more difficult to obtain as it involves only diagonal configurations $\mathbf{R}'=\mathbf{R}$ which are never sampled in practice. This normalization problem can be avoided by extending the random walk over both, diagonal, and off-diagonal space with a generalized partition function Z' given by

$$Z' = \int d\mathbf{R}\pi(\mathbf{R}) + \frac{z}{V}\int d\mathbf{R}\int d\mathbf{R}'\pi_{\rm off}(\mathbf{R}',\mathbf{R}), \quad \pi(\mathbf{R}) = |\Psi_T(\mathbf{R})|^2, \quad \pi_{\rm off}(\mathbf{R}',\mathbf{R}) = |\Psi_T(\mathbf{R}')\Psi_T(\mathbf{R})| \tag{7}$$

The Monte Carlo moves must now include also the possibility to "open" the diagonal configurations $\{\mathbf{R},\mathbf{R}\} \to \{\mathbf{R},\mathbf{R}'\}$ and close the off-diagonal ones $\{\mathbf{R},\mathbf{R}'\} \to \{\mathbf{R},\mathbf{R}\}$. The parameter z can be adapted to modify the frequency of switching between diagonal and off-diagonal configurations.

Standard VMC algorithm. Our general MC algorithm last time, was based on a uniform displacement of the particles. Here we will discuss the standard implementation within VMC based on force-bias MC. Therefore, let us try to improve the Metropolis acceptance rate using a gaussian a-priori probability

$$\mathcal{A}_D(\mathbf{R}\to\mathbf{R}') \sim e^{-(\mathbf{R}'-\mathbf{R}-\tau\mathbf{F}(\mathbf{R}))^2/2\tau} \tag{8}$$

where $\mathbf{F}(\mathbf{R})$ is a generalized force vector which only depends on $\mathbf{R}$. Let us look at the factor

$$r = \frac{|\psi_T(\mathbf{R}')|^2}{\mathcal{A}_{\mathcal{D}}(\mathbf{R}\to\mathbf{R}')}\frac{\mathcal{A}_D(\mathbf{R}'\to\mathbf{R})}{|\psi_T(\mathbf{R})|^2} \tag{9}$$

which enters into the Metropolis acceptance rate, $p = min[1,r]$, and expand $\log|\Psi_T(\mathbf{R}')| \simeq \log|\psi_T(\mathbf{R})]| + (\mathbf{R}-\mathbf{R}')\cdot\nabla\log|\psi_T(\mathbf{R})|$, we have

$$\frac{|\psi_T(\mathbf{R}')|^2}{\mathcal{A}_D(\mathbf{R}\to\mathbf{R}')} \simeq |\Psi_T(\mathbf{R}')\Psi_T(\mathbf{R})|\frac{e^{(\mathbf{R}'-\mathbf{R})\cdot\nabla\log|\psi_T(\mathbf{R})|}}{e^{-(\mathbf{R}'-\mathbf{R})^2/2\tau+(\mathbf{R}'-\mathbf{R})\cdot\mathbf{F}(\mathbf{R})/\tau-\mathbf{F}(\mathbf{R})^2/2\tau}} \tag{10}$$

and similar (exchanging $\mathbf{R}'$ with $\mathbf{R}$ and inverse)

$$\frac{\mathcal{A}_D(\mathbf{R}'\to\mathbf{R})}{|\psi_T(\mathbf{R})|^2} \simeq |\Psi_T(\mathbf{R}')\Psi_T(\mathbf{R})|^{-1}\frac{e^{-(\mathbf{R}'-\mathbf{R})^2/2\tau-(\mathbf{R}'-\mathbf{R})\cdot\mathbf{F}(\mathbf{R}')-\tau\mathbf{F}(\mathbf{R}')^2/2}}{e^{-(\mathbf{R}'-\mathbf{R})\cdot\nabla\log|\psi_T(\mathbf{R}')|}} \tag{11}$$

We get r close to one by choosing

$$\mathbf{F}(\mathbf{R}) = \nabla\log|\Psi_T(\mathbf{R})|, \quad r \simeq e^{\tau[\mathbf{F}(\mathbf{R})-\mathbf{F}(\mathbf{R}')][\mathbf{F}(\mathbf{R})+\mathbf{F}(\mathbf{R}')]/2} \simeq 1+\mathcal{O}\left(\tau^{3/2}\right) \tag{12}$$

Therefore, in the limit $\tau\to 0$, the drifted random walk based on

$$\mathcal{A}_D(\mathbf{R}\to\mathbf{R}';\tau) \sim e^{-(\mathbf{R}'-\mathbf{R}-\tau\nabla\log|\Psi_T(\mathbf{R})|)^2/2\tau} \tag{13}$$

directly samples $|\Psi_T(\mathbf{R})|^2$ as a stationary distribution

$$|\Psi_T(\mathbf{R}')|^2 = \int d\mathbf{R}\mathcal{A}_D(\mathbf{R}\to\mathbf{R}';\tau)|\Psi_T(\mathbf{R})|^2 \tag{14}$$

used later as an important ingredient in Diffusion Monte Carlo. However, the Metropolis Monte Carlo step is in general necessary to enforce detailed balance. Without detailed balance, small corrections, although formally of higher order in τ, may destroy the stationary solution. The VMC algorithm including the drift-forces writes

- subroutine **MCstep including drift**
- # fist move the old configuration $\mathbf{R}$ to sample a new one $\mathbf{R}_{new}$
  - Calculate forces $F(\mathbf{R}) = \nabla\log|\Psi_T(\mathbf{R})|$
  - create gaussian random numbers, δ, of variance τ
  - $\mathbf{R}_{new} = \mathbf{R} + \tau F(\mathbf{R}) + \delta$
  - calculate new forces $F(\mathbf{R}_{new})$
  - calculate δ_{new} which realizes the inverse: $\delta_{new} = \mathbf{R} - \mathbf{R}_{new} - \tau F(\mathbf{R}_{new})$

- # MC acceptance/ rejection using Metropolis algorithm
  - $p = \exp[-\delta_{new}^2/2\tau + \delta^2/2\tau]|\psi_T(\mathbf{R}_{new})|^2/|\psi_T(\mathbf{R})|^2$
  - IF $p >$rnd():
    - # accept move
    - use $\mathbf{R}_{new}$ as new configuration: $\mathbf{R} = \mathbf{R}_{new}$
- return configuration $\mathbf{R}$

Dynamical Interpretation of VMC with drift. Notice that the propagator $G_D(\mathbf{R} \to \mathbf{R}'; \tau)$ satisfying

$$-\frac{d}{d\tau} G_D(\mathbf{R} \to \mathbf{R}'; \tau) = -\frac{\hbar^2}{2m} \sum_i \nabla_i \left[\nabla_i - 2\nabla \log \Psi(\mathbf{R})\right] G_D(\mathbf{R} \to \mathbf{R}'; \tau) \tag{15}$$

leads to the stationary distribution $|\Psi_T(\mathbf{R})|^2$ as can be seen by direct insertion. Further, from the comparison with $\mathcal{A}_D(\alpha\tau)$ for $\tau \to 0$, we see that $G_D(\tau) \sim \mathcal{A}_D(\hbar^2\tau/m)$. The Monte-Carlo time in VMC based on gaussian a-priori probilities containing the drift $\hbar^2\tau/m\nabla \log \Psi_T$ thus realized the time-evolution given by Eq. (15), at least in the $\tau \to 0$ limit.

B. Projector Monte Carlo

In principle, any trial wave function can be improved by applying the operatore $\exp[-\beta H]$, and it is straightforward to show that

$$|\Psi_\beta\rangle = \frac{e^{-\beta H}|\Psi_T\rangle}{\left(\langle\Psi_T|e^{-2\beta H}|\Psi_T\rangle\right)^{1/2}} \simeq \frac{e^{-\beta E_0} c_0 |\Psi_0\rangle + e^{-\beta E_1} c_1 |\Psi_1\rangle \ldots}{\left(e^{-2\beta E_0} c_0^2 + e^{-2\beta E_1} c_1^2 \ldots\right)^{1/2}} \tag{16}$$

converges exponentially to the true ground state of the system for large projection time β. To sample $\Psi_\beta(\mathbf{R}) \equiv \langle \mathbf{R}|\Psi_\beta\rangle$, we need an explicit expression for the propagator $G(\mathbf{R}, \mathbf{R}'; \beta) \equiv \langle \mathbf{R}| \exp[-\beta H]|\mathbf{R}'\rangle$, which can be obtained from path-integral representation

$$\Psi_\beta(\mathbf{R}) \;\propto\; \int d\mathbf{R}'\, G(\mathbf{R}, \mathbf{R}'; \beta) \Psi_T(\mathbf{R}') \tag{17}$$

$$G(\mathbf{R}, \mathbf{R}'; \beta) \;=\; \lim_{M\to\infty} \int d\mathbf{R}_1 \cdots \int d\mathbf{R}_{M-1} G(\mathbf{R}, \mathbf{R}_1; \beta/M) \cdots G(\mathbf{R}_{M-1}, \mathbf{R}'; \beta/M) \tag{18}$$

together with explicit short-time expressions of the propagator, $G(\tau)$, valid in the limit $\tau \equiv \beta/M \to 0$. Projector Monte Carlo methods use in general a fixed discretization, M, and sample the integrals in Eq. (18) by Monte Carlo methods based on an explicitly known short-time approximation of the propagator $G(\tau)$. The extrapolation with respect to $\tau \to 0$ is done separately.

Primitive approximation. Conceptually simple is the primitive approximation, which neglects effects due to the commutator between the kinetic energy $T_N = -\sum_i \nabla_i^2/2m$, the external potential $U_N = \sum_i u(\mathbf{r}_i)$, and the interaction energy $V_N = \sum_i \sum_{i<j} v(r_{ij})$ with $H = T_N + U_N + V_N$,

$$G(\mathbf{R}, \mathbf{R}'; \tau) \;\approx\; e^{-\tau[V_N(\mathbf{R}) + U_N(\mathbf{R})]/2} G_0(\mathbf{R}, \mathbf{R}'; \tau) e^{-\tau[V_N(\mathbf{R}') + U_N(\mathbf{R}')]/2} e^{\mathcal{O}(\tau^2)} \tag{19}$$

$$G_0(\mathbf{R}, \mathbf{R}'; \tau) \;=\; \langle \mathbf{R}| e^{-\tau T_N} |\mathbf{R}'\rangle \tag{20}$$

$$\;=\; \left(\frac{m}{2\pi\hbar^2\tau}\right)^{dN/2} \exp\left[-\frac{m}{2\hbar^2\tau} \sum_i (r_i - r_i')^2\right] \tag{21}$$

In practice, the primitive approximation converges slowly with $\tau \to 0$, and may also not converge at all for long-range or hard core potentials. However, let us first consider the practical implementation of projector Monte Carlo methods.

C. Diffusion Monte Carlo

Diffusion Monte Carlo directly addresses the $\beta \to \infty$ limit, noticing that the (unnormalized) ground state distribution

$$\Psi_0(\mathbf{R}) = \lim_{\beta\to\infty} e^{\beta E_T} \int d\mathbf{R}' G(\mathbf{R}, \mathbf{R}'; \beta)\Psi_T(\mathbf{R}') \tag{22}$$

is also stationary under application of the short-time evolutio for $E_T = E_0$,

$$\Psi_0(\mathbf{R}) = e^{\tau E_T} \int d\mathbf{R}' G(\mathbf{R}, \mathbf{R}'; \tau)\Psi_0(\mathbf{R}') \tag{23}$$

or, using the primitive approximation

$$\Psi_0(\mathbf{R}) = \int d\mathbf{R}' G_0(\mathbf{R}, \mathbf{R}'; \tau) e^{\tau E_T - \tau[V(\mathbf{R}')+V(\mathbf{R})]/2}\Psi_0(\mathbf{R}') \tag{24}$$

Notice, that $G_0(\mathbf{R}, \mathbf{R}'; \tau)$ involves a pure diffusion from $\mathbf{R}$ to $\mathbf{R}'$ which we can realize using a displacement based on gaussian random numbers of variance $\hbar\tau/2m$. The remaining positive weights $w(\mathbf{R}, \mathbf{R}') = e^{\tau(E_T - V(\mathbf{R})/2 - V(\mathbf{R}')/2]}$ can then be interpreted as a change of weigth when moving to the new configuration.

Imagine now that we have $M \gg 1$ replicas of our system realizing an ensemble of configurations, $\{\mathbf{R}_1, \mathbf{R}_2, \ldots \mathbf{R}_M\}$. Each system ($\mathbf{R}_i$), called *walker*, is independent of the others. We can now diffuse one walker from $\mathbf{R}_i$ to a new configuration $\mathbf{R}'_i$ according to G_0. If the weight factor $w(\mathbf{R}_i, \mathbf{R}'_i)$ is exactly given by the non negative integer $[w]$, we can simple add $[w]$ new walkers at $\mathbf{R}_{M+i} \equiv \mathbf{R}'_i$ to the ensemble. But what should we do in the general case where $1 > p = w - [w] \geq 0$ (here $[w]$ is the closest integer to w with $w - [w] \geq 0$)? In this case, p realizes the probability that the original walker survives (or, equivalently, $(1-p)$ the probability for dying). For large enough ensembles, the mean number of walkers will fluctuate, due to this *branching* process. Successif iterations will lead to an exponential grow or death of the population, whenever E_T differs from the true ground state energy. Therefore, E_T must be adapted in time in order to reach a stationary process, which determines E_0. Alternatively, we have

$$E_0 = \frac{H\Psi_0(\mathbf{R})}{\Psi_0(\mathbf{R})} \to \lim_{\beta\to\infty} \langle V(\mathbf{R})\rangle_{\psi_\beta} \tag{25}$$

A sketch of the basic DMC program is

- program **DMC**
- call **initialize**($\mathbf{R}$,E_T)
- LOOP over MC time
  - LOOP over all walkers
    - call **Diffuse**($\mathbf{R}$)
    - call **Branch**($\mathbf{R}$,E_T)
  - END LOOP walkers
  - average over observables
  - call **PopulationControl**(E_T)
- END LOOP MC time

In **Diffuse**($\mathbf{R}$) we just diffuse $\mathbf{R} \to \mathbf{R} + \eta$ where η is gaussian distributed with variance $\hbar^2\tau/m$, whereas **Branch** creates or destroys walkers

- subroutine **Branch**($\mathbf{R}$,E_T)
  - calculate weight $w = \exp[\tau(E_T - V(\mathbf{R}))]$
  - total number of copies is $n = [w + rand()]$
  - $n > 0$? $\to$ add $n - 1$ walkers
  - $n = 0$? $\to$ delete walker

The population control can be done by assuming simple exponential law for the change of the population. If we aim for a stable number of walkers, $\overline{N_w}$, and N_W is the actual number, we want

$$\overline{N_w} = N_W e^{t\delta E_T} \tag{26}$$

after a certain time t. We then adapt the new trial energy $E_T + \Delta E_T$ according to

$$\delta E_T = \frac{1}{t} \log \overline{N_w}/N_w \tag{27}$$

A finite number of walker in the ensemble will introduce a small *population bias* which should be estimated by changing the population size.

Importance Sampling. The DMC algorithm describes above, clearly suffers from a large variance of of the potential energy around its average value which results in rather uncontrolled branching. Since our trial wave function $\Psi_T(\mathbf{R})$ should correspond to the best approximation of $\Psi_0(\mathbf{R})$, we should try to include this physical information in our DMC algorithm. Since our latest VMC algorithm provided an almost rejection free sampling to obtain $|\Psi_T(\mathbf{R})|^2$ in the limit $\tau \to 0$, we should implement this purely diffusive dynamics to sample

$$f(\mathbf{R}) \equiv \lim_{\beta\to\infty} \Psi_T(\mathbf{R})\Psi_\beta(\mathbf{R}) \tag{28}$$

Similar to Eq. (23), $f(\mathbf{R})$ is the stationary solution of

$$f(\mathbf{R}) = e^{\tau E_T} \int d\mathbf{R}' G_T(\mathbf{R}' \to \mathbf{R}; \tau) f(\mathbf{R}') \tag{29}$$

where the importance sampling propator is

$$G_T(\mathbf{R}' \to \mathbf{R}; \tau) = \frac{1}{\Psi_T(\mathbf{R}')} G(\mathbf{R}', \mathbf{R}; \tau)\Psi_T(\mathbf{R}) \tag{30}$$

In the case where we have the exact eigen function, $\Psi_T(\mathbf{R}) = \Psi_0(\mathbf{R})$, we know that G_T is exactly given by G_D which described the drifted random walk used in VMC. In that case, branching is not any more necessary. This motivates further explicit simplification of the importance sampling propagator G_T in a form where G_D is separated. Indead, we have

$$-\frac{d}{d\tau} G_T(\mathbf{R} \to \mathbf{R}'; \tau) = \frac{1}{\Psi_T(\mathbf{R})} \left\langle \mathbf{R}|e^{-\tau H}H|\mathbf{R}'\right\rangle \Psi_T(\mathbf{R}') = \frac{\Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} \left[H_{\mathbf{R}'} \left\langle \mathbf{R}|e^{-\tau H}|\mathbf{R}'\right\rangle\right] \tag{31}$$

or inserting $\left\langle \mathbf{R}|e^{-\tau H}|\mathbf{R}'\right\rangle = \Psi_T(\mathbf{R})G_T(\mathbf{R} \to \mathbf{R}')/\Psi_T(\mathbf{R}')$ we get

$$-\frac{d}{d\tau} G_T(\mathbf{R} \to \mathbf{R}'; \tau) = \Psi_T(\mathbf{R}')H \frac{1}{\Psi_T(\mathbf{R}')} G_T(\mathbf{R} \to \mathbf{R}'; \tau) \tag{32}$$

where H is applied with respect to $\mathbf{R}'$ on the right (but a similar equation can be obtained by applying H to $\mathbf{R}$). Now the only non-comuting part is the kinetic energy involving

$$\begin{aligned}
\Psi_T\nabla^2 \frac{1}{\Psi_T} G_T &= \nabla \left[\Psi_T \nabla \frac{1}{\Psi_T} G_T\right] - [\nabla\Psi_T]\nabla[\frac{1}{\Psi_T} G_T] && (33)\\
&= \nabla \left[-(\nabla \log \Psi_T)G_T + \nabla G_T\right] + \frac{[\nabla\Psi_T]^2}{\Psi_T^2} G_T - (\nabla \log \Psi_T)\nabla G_T && (34)\\
&= \nabla \left[\nabla - 2(\nabla \log \Psi_T)\right] G_T + [\nabla^2 \log \Psi_T]G_T + (\nabla \log \Psi_T)^2 G_T && (35)
\end{aligned}$$

but $\nabla^2 \log \Psi_T = \Psi_T^{-1}\nabla^2\Psi_T - (\nabla \log \Psi_T)^2$ so that we get

$$-\frac{d}{d\tau} G_T(\mathbf{R} \to \mathbf{R}'; \tau) = -\frac{\hbar}{2m} \sum_i \nabla_i' \left[\nabla_i' - 2\nabla_i' \log \Psi_T(\mathbf{R}')\right] G_T(\mathbf{R} \to \mathbf{R}'; \tau) + E_L(\mathbf{R}')G_T(\mathbf{R}' \to \mathbf{R}; \tau) \tag{36}$$

Neglecting the term containing the local energy, $E_L(\mathbf{R})$, we recover the drifted random walk used in a standard VMC calculation, Eq. (15). Applying the Trotter breakup, we thus obtain

$$G_T(\mathbf{R} \to \mathbf{R}'; \tau) = G_D(\mathbf{R} \to \mathbf{R}'; \tau) e^{-\tau[E(\mathbf{R})+E(\mathbf{R}')]/2} \tag{37}$$

for $\tau \to 0$. Therefore, for updating a standard VMC algorithm to include DMC, just the branching term, creating $[e^{\tau[E_T - E(\mathbf{R})/2 + E(\mathbf{R}')/2]} + rand()]$ copies of one configuration, must be added.

Within DMC, the ground state energy can be obtained from the expectation value of the local energy

$$E_0 = \langle E_L(\mathbf{R}) \rangle = \frac{\int d\mathbf{R} E_L(\mathbf{R}) f(\mathbf{R})}{\int d\mathbf{R} f(\mathbf{R})} = \frac{\int d\mathbf{R} \Psi_T(\mathbf{R}) H \Psi_0(\mathbf{R})}{\int d\mathbf{R} \Psi_T(\mathbf{R}) \Psi_0(\mathbf{R})} \tag{38}$$

Other observables can be adressed by so-called *mixed estimators*,

$$\langle O \rangle = \frac{\int d\mathbf{R} \Psi_T(\mathbf{R}) O(\mathbf{R}) \Psi_0(\mathbf{R})}{\int d\mathbf{R} \Psi_T(\mathbf{R}) \Psi_0(\mathbf{R})} \tag{39}$$

which lead to a systematic small bias.

D. Variational Path Integral and Reptation Monte Carlo

Sampling the DMC wavefunction, three sources of systematic bias are important:

- short time discretization $\to$ extrapolation $\tau \to 0$
- population bias $\to$ extrapolation $\overline{N_w} \to \infty$
- mixed estimator bias for observables different from energy

The following Projector Monte Carlo methods avoid the last two sources of systematic errors. They are based on the path-integral representation of Eq. (16) for fixed β, and introduce a finite projection time which necessites the explicit extrapolation $\beta \to \infty$.

Variational Path Integral. Since we have an explicit expression for $\Psi_\beta(\mathbf{R})$ in terms of high dimensional integrals with a non-negative integrand, we can simply try to sample $|\Psi_{\beta/2}(\mathbf{R})|^2$, but now in a configuration space $\mathcal{R} \equiv \{\mathbf{R}_0, \mathbf{R}_1, \cdots, \mathbf{R}_M\}$ which is $M+1$ times bigger than just $\mathbf{R} \equiv \mathbf{R}_{M/2}$

$$\langle O \rangle_{\beta/2} = \left\langle O(\mathbf{R}_{M/2}) \right\rangle_\pi = \int d\mathbf{R}_0 \ldots \int d\mathbf{R}_M O(\mathbf{R}_{M/2}) \pi(\mathcal{R}), \tag{40}$$

$$\pi(\mathcal{R}) = \Psi_T(\mathbf{R}_0) \exp\left[-\sum_{t=0}^{M} S(\mathbf{R}_t; \mathbf{R}_{t+1}) \right] \Psi_T(\mathbf{R}_M) \tag{41}$$

where $S(\mathbf{R}, \mathbf{R}') = \log G(\mathbf{R}, \mathbf{R}'; \beta/M)$. However, in contrast to simple VMC moves, updates which change the whole configuration space $\mathcal{R}$ will be most likely rejected, and we have to create new moves in order to equilibrate the $(M+1)N$ particles in the simulation. I will not discuss typical moves as they are in general borrowed from finite temperature Path-Integral calculations beyond the scope of this lecture.

Reptation Monte Carlo. A particular update in Variational Path Integral is given by simple shifting the old configurations $\mathcal{R}$ in the time directions, e.g.

$$\mathcal{R}' = \{\mathbf{R}'_0, \ldots \mathbf{R}'_M\} \quad \text{with} \quad \mathbf{R}'_0 \equiv \mathbf{R}_1, \mathbf{R}'_1 \equiv \mathbf{R}_2, \ldots, \mathbf{R}'_{M-1} \equiv \mathbf{R}_M \tag{42}$$

together with a new proposition for $\mathbf{R}_M$ based on the drifted random walk used in DMC, $G_D(\mathbf{R}_M \to \mathbf{R}'_M; \tau)$. The inverse move is realized by the shift in the opposite direction which involves a drivted random walk $\mathbf{R}'_0 \equiv \mathbf{R}_1 \to \mathbf{R}_0$. Choosing randomly the direction before any move, the Metropolis acceptance probability then writes

$$p = \min\left\{ 1, \frac{\pi(\mathcal{R}')}{\pi(\mathcal{R})} \frac{G_D(\mathbf{R}_1 \to \mathbf{R}_0, \tau)}{G_D(\mathbf{R}_M \to \mathbf{R}'_M; \tau)} \right\} \tag{43}$$

Further use of the importance sampling propagator, Eq. (30), in the statistical weight of the path, we have

$$\pi(\mathcal{R}) = \Psi_T(\mathbf{R}_0) \prod_{i=1}^{M} G(\mathbf{R}_{i-1}, \mathbf{R}_i; \tau) \Psi_T(\mathbf{R}_M) \tag{44}$$

$$= \Psi_T^2(\mathbf{R}_0) \prod_{i=1}^{M} G_T(\mathbf{R}_{i-1} \to \mathbf{R}_i; \tau) \tag{45}$$

so that

$$\frac{\pi(\mathcal{R}')}{\pi(\mathcal{R})} = \frac{\Psi_T(\mathbf{R}_1)}{\Psi_T(\mathbf{R}_0)G(\mathbf{R}_0,\mathbf{R}_1;\tau)} \frac{G(\mathbf{R}_M,\mathbf{R}'_M;\tau)\Psi_T(\mathbf{R}'_M)}{\Psi_T(\mathbf{R}_M)} = \frac{G_T(\mathbf{R}_M \to \mathbf{R}'_M;\tau)}{G_T(\mathbf{R}_0 \to \mathbf{R}_1;\tau)} \frac{|\Psi_T(\mathbf{R}'_M)|^2}{\Psi_T(\mathbf{R}_M)|^2} \tag{46}$$

Inserting the explicit expression of the DMC propagator, Eq. (37), the acceptance ratio finally simplifies

$$p = \min\left\{1, \frac{\Psi_T^2(\mathbf{R}'_M)}{\Psi_T^2(\mathbf{R}_M)} \frac{\exp\left(-\tau[E_L(\mathbf{R}'_M)+E_L(\mathbf{R}_M)]/2\right)}{\exp\left(-\tau[E_L(\mathbf{R}_1)+E_L(\mathbf{R}_0)]/2\right)}\right\} \tag{47}$$

E. Many-body propagator

Zero-temperature DMC propagator. Whereas the primitive approximation may contain singularities, e.g. for systems with Coulomb or hard core interactions, importance sampling based on Ψ_T eliminates this problem. If we are mainly interested in ground state properties, we can also use the expression of the importance sampling propagator used in DMC, Eq. (37), and use Eq. (30) to obtain

$$G(\mathbf{R},\mathbf{R}';\tau) = \frac{1}{\Psi_T(\mathbf{R})} G_T(\mathbf{R}' \to \mathbf{R};\tau)\Psi_T(\mathbf{R}') \tag{48}$$

Note that the propagator, G, is not any more symmetric with respect to interchanging $\mathbf{R}$ and $\mathbf{R}'$, however, within the order of the approximation, we can conserve this important property by explicit symmetrization, either

$$G(\mathbf{R}',\mathbf{R},\tau) \approx \left[G_T(\mathbf{R}',\mathbf{R};\tau)G_T(\mathbf{R},\mathbf{R}';\tau)\right]^{1/2} \tag{49}$$

or

$$G(\mathbf{R}',\mathbf{R},\tau) \approx \min\left[\Psi_T(\mathbf{R}')G_T(\mathbf{R}',\mathbf{R};\tau)\Psi_T^{-1}(\mathbf{R}), \Psi_T(\mathbf{R})G_T(\mathbf{R},\mathbf{R}';\tau)\Psi_T^{-1}(\mathbf{R}')\right] \tag{50}$$

together with the explicit expression, Eq. (37).

Pair product approximation. Better approximations for the potential energy can be found, e. g. by considering the effective pair potential, u_p, given by the solution of the two-particle problem

$$e^{-u_p(\mathbf{r}_i,\mathbf{r}_j;\mathbf{r}'_i,\mathbf{r}'_j;\tau)} \equiv \frac{\langle \mathbf{r}_i\mathbf{r}_j|e^{-\tau(T_2+V_2)}|\mathbf{r}'_i,\mathbf{r}'_j\rangle}{\langle \mathbf{r}_i\mathbf{r}_j|e^{-\tau T_2}|\mathbf{r}'_i,\mathbf{r}'_j\rangle} \tag{51}$$

and the whole propagator is then approximated by

$$G(\mathbf{R},\mathbf{R}';\tau) \approx G_0(\mathbf{R},\mathbf{R}';\tau)e^{-\tau[U_N(\mathbf{R})+U_N(\mathbf{R}')]/2}e^{-\sum_{i<j} u_p(\mathbf{r}_i,\mathbf{r}_j;\mathbf{r}'_i,\mathbf{r}'_j;\tau)} \tag{52}$$

where T_2 is the kinetic energy and V_2 the interatomic potential. This propagator is frequently employed in finite temperature path-integral Monte Carlo calculations. As the time-step error is small, less discretization points M are needed, however, the parametrization of $u_p(\mathbf{r},\mathbf{r}';\tau)$ for a given interaction potential v as a function of the relative coordinates $\mathbf{r}$ and $\mathbf{r}'$ is a majeur complication.

F. Fixed-node approximation

Quite generally, the ground state of any (regular) Hamiltonian, is nodeless and symmetric with respect to particle exchange. Thus, all QMC methods described above can be directly applied to obtain the ground state of a system containing N Bosons. This is not the case for Fermions, since the ground state wave function of a Fermi system must be antisymmetric,

$$\Psi_F(\ldots,\mathbf{r}_i,\ldots,\mathbf{r}_j) = -\Psi_F(\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_i), \quad \text{for any } i,j \tag{53}$$

leading to nodes where $\Psi(\mathbf{R}) = 0$. Thus, in general, the ground state of fermions is never the lowest eigenstate of the Hamiltonian of the system. Only in particular situations which we do not adress here, e.g. for some particular Hamiltonian in one spatial dimension, it can be degenerate with the Bose ground state.

It is possible to extend the variational principle for the energy to some excited states, $|\Psi_m\rangle$ with $E_m > E_0$, imposing orthogonality of the trial wave function to all lower eigenfunctions, $\langle\Psi_T|\Psi_n\rangle = 0$ for all n with $E_n < E_m$. Since the Fermion ground state wave function is the lowest eigenfunction in the space of anti-symmetric wave functions, orthogonality to states is guaranteed by symmetry. Since VMC based methods are based on sampling $\Psi_T(\mathbf{R})|^2 \geq 0$, they can be directly applied to Fermions by using antisymmetric trial wave functions which obey Eq. (53).

Fermion sign problem. In contrast to VMC, Projection Monte Carlo methods stochastically sample $\Psi_0(\mathbf{R})$ which now contains negative regions where the wave function cannot be interpreted as probability. It poses no problem to adapt to sample $|\Psi(\mathbf{R})|$ and to treat the sign $s(\mathbf{R}) \equiv \Psi(\mathbf{R})/|\Psi(\mathbf{R})| = \pm 1$ as observable. However since $|\Psi(\mathbf{R})|$ has a non-vanishing overlap with the bosonic ground state of energy E_B, we have

$$\overline{s} \equiv \langle s\rangle_\beta \sim \exp[-N\beta(E_F - E_B)/N] \tag{54}$$

which enters in the normalization of any expectation value of an observable. Assuming a finite gap between the fermionic and bosonic ground state energy per particle, $\overline{s}$ vanishes exponentially in $N\beta \gg 1$. Since $\langle s^2\rangle = 1$, the variance is one, and the signal to noise ratio prevents any direct sampling involving $\overline{s}$.

Antisymmetry and nodes. Importance sampling DMC was based on the overlap

$$f(\mathbf{R}) = \Psi_T(\mathbf{R})\Psi_0(\mathbf{R}) \tag{55}$$

Let us imaging that we have found a trial wave function with exacly the same positive and negative regions as the Fermion ground state we are looking for, $f(\mathbf{R}) \geq 0$ for $\Psi_0 \in \Psi_F$. In this case, if we impose $f(\mathbf{R}) \geq 0$ during the time evolution in DMC, we expect that DMC converges to the exact Fermion ground state. What happens? Looking at the drifted random walk created by the importance sampling, imposing $f(\mathbf{R}) \geq 0$ for all $\mathbf{R}$, we reject any move $\mathbf{R} \to \mathbf{R}'$ with $\Psi_T(\mathbf{R})\Psi_T(\mathbf{R}') < 0$. Our population of walkers can be separated into two sets, positive walkers at $\mathbf{R}^+$ which satisfy $\Psi_T(\mathbf{R}^+) \geq 0$, and negative walkers at $\mathbf{R}^-$ with $\Psi_T(\mathbf{R}^-) < 0$. Postive and negative walkers are separated by the nodal surface $\mathbf{S}$ where $\Psi_T(\mathbf{S}) \equiv 0$, and it is enough to know the exact nodal surface. Note that the nodal surface $\mathbf{S}$ is a hypersphere in $Nd-1$ dimensions. For any sufficiently regular antisymmetric trial function, applying the Permutation operator to any positive configuration, we obtain a negative walker, and vice versa. It is therefore sufficient to sample only the positive space as long as we are only interested in physical observables which commutes with the permutation operator.

Fixed-node approximation. Everything above is fine, but we still do not know the nodal surface for almost all fermion problems. Since we have no further idea right now, let us search for the best approximation we can do. In the fixed-node approximation, we simple impose the nodes of a given trial wave function. Once started with positive walkers, our fixed-node DMC algorithm will converges to an eigenfunction of the Hamiltonian

$$H\Psi_{FN}(\mathbf{R}) = E_{FN}\Psi_{FN}(\mathbf{R}), \quad \text{for all } \mathbf{R} \text{ in } \mathbf{R}^+, \text{ the positive region with } \Psi_T(\mathbf{R}) \geq 0 \tag{56}$$

On the nodes $\mathbf{S}$ of Ψ_T, we also have $\Psi_{FN}(\mathbf{S}) = 0$, and we can continue the wave function to the negative regions, $\mathbf{R}^-$ using permutations $\Psi_{FN}(\mathbf{R}) = (-)^{|P|}\Psi_{FN}(P\mathbf{R})$, where the permutation P can be determined from solving $\Psi_T(\mathbf{R}) = (-)^{|P|}\Psi_T(P\mathbf{R})$ for P. As long as Ψ_T is a sufficiently regular fermionic trial wave function, we can reach all configuration space by this procedure, and the continued Ψ_{FN} is a continuous antisymmetric wave function.

Upper bound theorem for fixed-node wave functions. Unfortunately, the first derivatives of the constructed FN wavefunction with respect to $\mathbf{r}_i$ are in general not continuous at the nodal surface. Therefore, we cannot directly apply the variational principle as the underlying expansion in eigenfunctions of the true Hamiltonian is only complete for wavefunctions with continuous first derivatives. However, we can smear out our wave function at a distance ϵ close to the nodes to make them sufficiently smooth to apply the variational theorem, so that the smoothed function provides an upper bound for the energy. This smoothing will increase the absolute value of the curvature $\sim \epsilon^{-1}$ close to the node and the laplacian of the kinetic energy will produce large absolute values, $\sim \varepsilon^{-1}$ However, since the wave function vanishes as ϵ, the kinetic energy contribution of the smoothed wave function close to the nodal region $\sim \int_\epsilon \psi\nabla^2\psi \sim \epsilon$ vanishes. Therefore, the energy of our fixed-node wave function provides a true upper bound to the fermion ground state energy, E_F,

$$E_F \leq E_{FN} = \frac{\int d\mathbf{R}\Psi_{FN}(\mathbf{R})E_L(\mathbf{R})\Psi_{FN}(\mathbf{R})}{\int d\mathbf{R}\Psi_{FN}(\mathbf{R})\Psi_{FN}(\mathbf{R})} \tag{57}$$

For many-body fermion problems, the fixed-node energies are the most accurate values. The upper bound property further allows us to judge the quality of different trial wave function without relying on comparison with experiment.

Fixed-phase approximation. As a generalization of the fixed-node approach, the fixed-phase approximation is based on a complex trial wave function

$$\Psi_T(\mathbf{R}) = A(\mathbf{R})\exp[-i\varphi(\mathbf{R})], \quad \text{with positive ampltude } A(\mathbf{R}) \geq 0 \text{ and real phase } \varphi(\mathbf{R}) \tag{58}$$

For any given phase, we can then minimize the energy of the trial wave function for an explicitly given phase, $\varphi(\mathbf{R})$. However, since the phase is only well defined (and behaved) for non-vanishing amplitude, we also have to fix the nodes of the amplitude. An argument similar to that above shows that the fixed-phase wave function provides also an upper bound for the ground state energy in the same symmetry class as Ψ_T. Fixed-phase methods are needed for treating twisted boundary conditions, magnetic field effects, etc.