UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL ATOMIC ENERGY AGENCY
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

SMR.998a - 6

Research Workshop on Condensed Matter Physics
30 June - 22 August 1997
MINIWORKSHOP ON
QUANTUM MONTE CARLO SIMULATIONS OF LIQUIDS AND SOLIDS
30 JUNE - 11 JULY 1997
and
CONFERENCE ON
QUANTUM SOLIDS AND POLARIZED SYSTEMS
3 - 5 JULY 1997

"Variational Monte Carlo"

G. SENATORE
Dipartimento di Fisica Teorica
Universita' di Trieste
Strada Costiera 11
Trieste
Italy

These are preliminary lecture notes, intended only for distribution to participants.

Variational Monte Carlo

- If the wavefunction is given, MC provides a straightforward mean of evaluating state averages.

- Assume that we have a trial wavefunction $\Psi(R;\mathbf{a})$ depending on a set of parameters $\mathbf{a} = (a_1, a_2, \ldots, a_p)$. The variational theorem states that

$$E(\mathbf{a}) = \frac{\int dR\, \Psi(R;\mathbf{a}))^* H \Psi(R;\mathbf{a})}{\int dR |\Psi(R;\mathbf{a})|^2} \geq E_0,$$

  with E_0 the exact ground state energy.

- We know how to generate configurations (say M) distributed with the probability $\pi(R;\mathbf{a}) = |\Psi(R,\mathbf{a})|^2$. Hence we can estimate $E(\mathbf{a})$,

$$E(\mathbf{a}) = \frac{1}{M} \sum_{i=1}^{M} E_L(R_i;\mathbf{a}),$$

  with $E_L(R_i;\mathbf{a}) = \Psi^{-1}(R,\mathbf{a}) H \Psi(R,\mathbf{a})$.

- We can optimize $\Psi(R;\mathbf{a})$ by minimizing $E(\mathbf{a})$, to obtain the best upper bound to E_0.

Variance

The optimization of $\Psi(R;\mathbf{a})$ can be also achieved with other techniques.

- Another quantity enjoying a minimum property is the variance

$$\sigma^2(\mathbf{a}) = \frac{\int dR\, \Psi(R;\mathbf{a}))^* (H-E)^2 \Psi(R;\mathbf{a})}{\int dR |\Psi(R;\mathbf{a})|^2} \geq 0.$$

  The minimum value attainable by σ^2 is 0, which is achieved whenever $\Psi(R,\mathbf{a})$ coincides with an exact eigenstate of H, say ϕ_n with eigenvalue E_n, and E is set equal to the E_n.

- In principle one could judge on the quality of a minimization by looking at the size of σ^2. In practice, variance minimization has a number of bonuses:
  - It can be used to study excited states .
  - Being a sum of squares, σ^2 can be efficiently minimized using efficient algorithms like that of Levenberg and Marquand.
  - It requires a smaller number of configurations as compared with the energy minimization.
  - The only way σ^2 can be made small is by having $E_L(R;\mathbf{a})$ smooth and close to an eigenvalue, whereas the energy minimization can be biased by configurations with $E_L(R;\mathbf{a})$ to low.

– G. Senatore –

The Trial Wavefunction

A good trial wavefunction, apart from being flexible enough should satisfy a minimum number of basic requirements:

1. Ψ and $\nabla\Psi$ should be continuous for finite potential $V(R)$.

2. Not only $\int \Psi^2$ and $\int \Psi^* H \Psi$ should exists, but also $\int \Psi^* H^2 \Psi$, in order that the variance exists and statistical errors are finite.

3. $\Psi(R)$ should have the correct symmetry property, i.e., $\Psi(R) = (-)^P \Psi(PR)$ for Fermions and $\Psi(R) = \Psi(PR)$ for Bosons, for any particle permutation P.

4. The wavefunction should embody all the know exact behaviours, so as to make the local energy as smooth as possible.

Symmetry

Bosons

For an uniform Bose fluid the simplest trial function is of the Bijl–Jastrow type

$$\Psi_2(R) = \exp[-\sum_{i<j} u(r_{ij})] \equiv J(R).$$

More refined wavefunctions include three–body correlations to read

$$\Psi_3(R) = J(R) \cdot \exp[-\sum_{i<j<k} u_3(r_{ij}, r_{ik}, cos(\hat{r}_{ij} \cdot \hat{r}_{ik}))].$$

Fermions

A typical Fermion wavefunction is obtained augmenting the Bose function by a determinant ensuring antisymmetry, in the simplest case

$$\Psi(R) = J(R) \cdot det[\varphi_i(r_k, \sigma_k)],$$

with $\varphi_i(r,\sigma_k)$ the i-th spin orbital. One needs, for N particle N distinct spin orbitals to get a non-vanishing determinant.

Similarly to the case of Bosons, more refined wavefunctions are obtained with resorting to triplet *pseudopotentials*:

$$\Psi(R) = \Psi_3(R) \cdot det[\varphi_i(r_k, \sigma_k)].$$

Exact behaviours

Simple limiting behaviours of the wavefunction can be obtained imposing the smoothness of $E_L(R)$ for small or large r_{ij}, with the other coordinates kept fixed.

Small r behaviour of u(r)

One looks at the dominant terms in $E_L(R)$ when two particles come close.

- For Bosons

$$E_L(R) = v(r) + 2D\nabla^2 u(r) - 2D(\nabla u(r))^2 + \cdots,$$

with $v(r)$ the pair interaction. So for LJ interaction one is led to $u(r) \propto 1/r^5, \quad \to 0$.

- For Fermions one would chose the $\varphi_i(r)$ as exact solution of the independent particle problem; then with an analysis similar to the one above one would get
$u(r) = a_{\sigma,\sigma'} r, \quad \to 0$, with

  - $a_{\uparrow\downarrow} = e^2/(4D)$, and
  - $a_{\downarrow\downarrow} = e^2/(8D)$.

Large r behaviour of $u(r)$

The study of u for large r is most easily accomplished by rewriting the variational energy in reciprocal space in terms of the collective coordinates $\rho_{\mathbf{k}} = \sum_i exp(i\mathbf{k}\cdot\mathbf{r}_i)$, with $S(k) = (1/N)\langle\rho_{\mathbf{k}}\rho_{-\mathbf{k}}\rangle$ the static structure factor.

Using the RPA approximation one gets

$$E_V = E_F + \sum_{\mathbf{k}} \left(S(k)\cdot(Dk^2u^2(k) + \frac{1}{2}v(k)) - \frac{v(k)}{2} \right)$$

and

$$S(k) = \frac{S_0(k)}{1 + 2u(k)S_0(k)}.$$

Variation with respect to $u(k)$ immediately yields

$$2u(q) = -\frac{1}{S_0(k)} + \sqrt{\frac{1}{S_0^2(k)} + \frac{2v(k)}{Dk^2}},$$

which implies, for charged Fermions in 3 dimensions, $u(r) \propto 1/r, \quad \to \infty$.

For short ranged potential, like Helium, replacing $u(k)$ with a constant for small k one obtains $u(r) \propto 1/r^2, \quad \to \infty$.

Reweighting

Optimization techniques require taking derivatives of MC estimates with respect to the variational parameters.

In principle one would obtain a sample $\{R_i\}$ from $\pi(R;\mathbf{a})$ and a sample $\{R_i'\}$ from $\pi(R;\mathbf{a}')$ to calculate for example

$$E(\mathbf{a}) = \frac{1}{M}\sum_{i=1}^{M} E_L(R_i;\mathbf{a}),$$

and

$$E(\mathbf{a}') = \frac{1}{M}\sum_{i=1}^{M} E_L(R_i';\mathbf{a}'),$$

and from these the derivative of $E(\mathbf{a})$. This procedure, however, turns out to be unstable due to the independent statistical errors on the two estimates of the energy.

If the two parameters sets are close enough to each other a winning strategy is to use the same sample, say $\{R_i\}$ for both evaluations. In other words,

$$E(\mathbf{a}') = \frac{1}{M}\frac{\sum_{i=1}^{M} E_L(R_i';\mathbf{a}')w(R_i)}{\sum_i w(R_i)},$$

with

$$w(R) = |\Psi(R;\mathbf{a}')|^2/|\Psi(R;\mathbf{a})|^2.$$

Filtering (projection) techniques

In the VMC a systematic optimization of Ψ is attempted, which however is limited by one's ability to model the wavefunction. It would be highly desirable to have schemes that starting from a given trial function would project out the exact ground state, by filtering out the higher energy components.

For Bosons, in fact, such schemes are available (see below). For Fermions, one has to fix the nodal structure and so these schemes become variational, though superior to conventional VMC.

The strategy is very simple and invariably goes through some suitable *evolution* either in imaginary time (DMC, PIGS), or time integrated (GFMC).

- GFMC

$$\Phi_{n+1}(R) = (E_T + V_0)\int dR'\, g(R,R')\Phi_n(R'),$$

$$g(R,R') = \langle R|\frac{1}{H+V_0}|R'\rangle,$$

with $H + V_0$ positive definite and E_T close to the ground state energy E_0. Expanding Φ_0 in eigenfunctions of H

$\Phi_0 = \Psi = \sum_i c_i \phi_i$, one immediately gets

$$\Phi_n(R) = \sum_i \left[\frac{E_T + V_0}{E_i + V_0}\right]^n c_i \phi_i(R),$$

which for large n yield

$$\Phi_n \propto \left[\frac{E_T + V_0}{E_o + V_0}\right]^n c_o \phi_0,$$

provided that $c_0 \neq 0$.

- DMC

$$\Phi(R,(n+1)\tau) = \int dR'\, G(R,R';t)\Phi(R';n\tau),$$

$$G(R,R';\tau) = \langle R|\exp(-\tau(H - E_T))|R'\rangle,$$

and E_T close to the ground state. Again for large n one projects out the ground state according to

$$\Phi_n \propto \exp(-\tau(E_0 - E_T))c_0\phi_0.$$

Note that $g(R,R') = \int_0^\infty d\tau G(R,R';\tau)$ with $E_T = -V_0$.

- PIGS

$$\frac{\langle\Psi|e^{-\tau H/2}He^{-\tau H/2}|\Psi\rangle}{\langle\Psi|e^{-\tau H}|\Psi\rangle}$$

Both DMC and GFMC involve $d \cdot N$ dimensional integrations, which require MC technique and can be dealt with by resorting to random walks.

Evidently, one has to sample also the appropriate Green's functions and this can be arranged with a small time expansion for DMC (see Chin lectures) and other techniques for GFMC (Domain Green's functions, ...).

Importance sampling

The GFMC and DMC schemes described above, though formally legitimate for Bosons, in fact are not stable due to infinities in the potential energy.

The cure is simple and robust and was discovered by Kalos.

For GFMC, rather than evolving $\Phi_n(R)$ with the initial condition $\Phi_0(R) = \Psi(R)$ one evolves $f_n(R) = \Phi_n(R)\Psi(R)$ starting from $f_0(R) = \Psi^2(R)$. After a transient, for n large enough $f_n(R) \propto \Phi_0(R)\Psi(R)$.

Similarly, for DMC one evolves $f(R, n\tau) = \Phi(R; n\tau)\Psi$, with the initial condition $f(R; 0) = \Psi^2(R)$a, and converging to $f(R; n\tau) \propto \Phi_0(R)\Psi(R)$ for large n.

The evolution of f is dictated by equations similar to those given above, with the green's functions which are replaced by

$$G(R, R'; \tau) \to K(R, R'; \tau) = \Psi(R)G(R, R'; \tau)\Psi^{-1}(R'),$$

for the DMC and

$$g(R, R') \to k(R, R') = \Psi(R)g(R, R')\Psi^{-1}(R'),$$

for the GFMC.

A detailed discussion of this point will be given by Chin, I guess.

Fermion nodes and Fixed-Node approximation

The possibility of sampling $f(R)$ ($f(R; t)$) in stable manner is related to its positivity. (See however Kalos lectures).

The Fermion ground state, for more than 2 particles has nodes, i.e., regions with positive and negative sign, which however enjoy the tiling property. Given one of the equivalent nodal pocket $\mathcal{D}_1$ [a region of given sign] all the other pockets are obtained by applying permutations to $\mathcal{D}_1$. This resolves the full spatial domain without holes or superpositions.

Assuming that the nodes of the sought wavefunctions coincide with those of $\Psi(R)$ [Fixed–node approximation] one is back to a nonnegative f which can be treated as a probability. This yields a stable algorithm, which evidently is variational: one has restricted the search in the wavefunction space. Moreover, due to the tiling property one does not need to worry about which pockets have been populated

Clearly, if one new the nodes (as in 1 dimension) choosing a trial functions with the exact nodes would lead to the exact ground state.

• But: transient estimates and nodal relaxation, tough not stable.

Ewald sums and N extrapolation

- Properties in the long-wavelength limit depend on long–range behaviour of the trial function as well as on the modeling of the system under study.

- Especially for Coulomb systems it is crucial to consider periodic replicas of the simulation cell, and to sum interactions with all the replicas (Ewald sums). It is also important to Ewald sum the pseudopotential.

- Even allowing for Ewald sums Coulomb systems have residual size effects due to the size of the cell (finite number of particles).

- Important size effects are present in Fermion systems in the Fermi liquid regime due to levels shell structure.

- To study systems in the thermodynamic limit, one can:
  - Study the system at various N and then try to extrapolate to $N = \infty$. This is usually the case with VMC.
  - Assume that the N dependence does not depend e much on the details of the simulation and borrow the number dependence of VMC. This is usually done with DMC and GFMC.