Summary and problems with variational methods

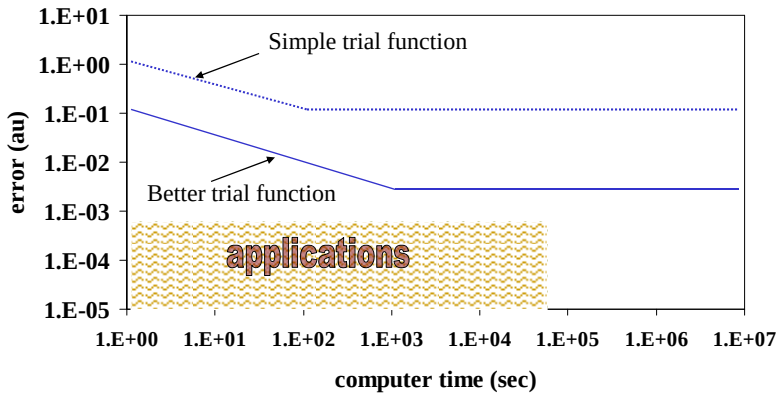
- Powerful method since you can use any trial function
- Scaling (computational effort vs. size) is almost classical
- Learn directly about what works in wavefunctions
- No sign problem

- Optimization is time consuming
- Energy is insensitive to order parameter
- Non-energetic properties are less accurate. O(1) vs. O(2) for energy.
- Difficult to find out how accurate results are.
- Favors simple states over more complicated states, e.g.
  - Solid over liquid
  - Polarized over unpolarized

What goes into the trial wave function comes out! "GIGO"

We need a more automatic method! Projector Monte Carlo

Summary of Variational (VMC)

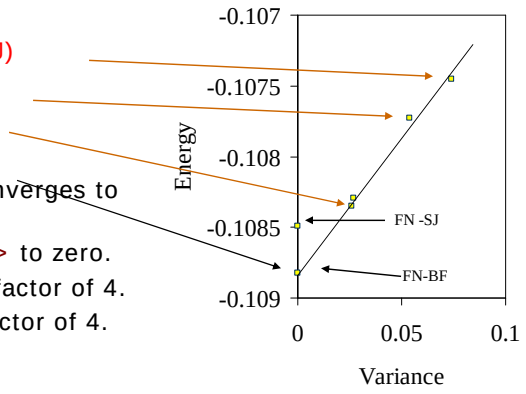


Dependence of energy on wavefunction

3d Electron fluid at a density $r_s=10$

Kwon, Ceperley, Martin, Phys. Rev. ***B58****,6800, 1998*

- Wavefunctions
  - Slater-Jastrow (SJ)
  - three-body (3)
  - backflow (BF)
  - fixed-node (FN)
- Energy $\langle\phi|H|\phi\rangle$ converges to ground state
- Variance $\langle\phi[H-E]^2\phi\rangle$ to zero.
- Using 3B-BF gains a factor of 4.
- Using DMC gains a factor of 4.



Projector Monte Carlo

•Originally suggested by Fermi and implemented in 1950 by Donsker and Kac for H atom.

•Practical methods and application developed by Kalos:

PHYSICAL REVIEW A VOLUME 9, NUMBER 5 MAY 1974

Helium at zero temperature with hard-sphere and other forces

M. H. Kalos*
Courant Institute of Mathematical Sciences, New York University, New York, New York 10012

D. Levesque and L. Verlet
Laboratoire de Physique Théorique et Hautes Energies, Orsay, France †
(Received 22 August 1973)

Various theoretical and numerical problems relating to heliumlike systems in their ground states are treated. New developments in the numerical solution of the Schrödinger equation permit the solution of 256-body systems with hard-sphere forces. Using periodic boundary conditions, fluid and crystal states can be described; results for the energy and radial-distribution functions are given. A new method of correcting for low-lying phonon excitations

Projector Monte Carlo

(variants: Green's function MC, Diffusion MC, Reptation MC)

- Project single state using the Hamiltonian

$$\phi(t) = e^{-(\mathrm{H}-\mathrm{E_T})\mathrm{t}}\phi(0)$$

- We show that this is a diffusion + branching operator. Maybe we can interpret as a probability. But is this a probability?
- **Yes!** for bosons since ground state can be made real and non-negative.
- **But** all excited states must have sign changes. This is the "sign problem."
- For efficiency we do "importance sampling."
- Avoid sign problem with the fixed-node method.

Diffusion Monte Carlo

- How do we analyze this operator? →

$$\psi(R,t) = e^{-(\mathrm{H}-\mathrm{E_T})\mathrm{t}}\psi(R,0)$$

- Expand into exact eigenstates of H. →

$$H\phi_\alpha = E_\alpha\phi_\alpha$$

$$\psi(R,0) = \sum_\alpha \phi_\alpha(R)\left\langle\phi_\alpha\middle|\psi(0)\right\rangle$$

- Then the evolution is simple in this basis.

$$\psi(R,t) = \sum_\alpha \phi_\alpha(R)e^{-t(E_\alpha - E_T)}\left\langle\phi_\alpha\middle|\psi(0)\right\rangle$$

- Long time limit is lowest energy state that overlaps with the initial state, usually the ground state.

$$\lim_{t\to\infty}\psi(R,t) = \phi_0(R)e^{-t(E_0-E_T)}\left\langle\phi_0\middle|\psi(0)\right\rangle$$

$$E_0 \approx E_T \Rightarrow \textit{normalization fixed}$$

- How to carry out on the computer?

Notation

- Individual coordinate of a particle r_i
- All 3N coordinates R= (r_1,r_2, …. r_N)
- R can depend on "imaginary time", "time slice" or "Trotter index" "t" or on iteration number: R_t.
- Total potential energy = V(R)
- Kinetic energy = $-\lambda\sum_{i=1}^{N}\nabla_i^2$ where $\lambda \equiv \frac{\hbar^2}{2m}$
- Hamiltonian $\hat{H} = \hat{T} + \hat{V}$

The Green's function

- Operator notation

$$\frac{d\hat{\rho}}{dt} = -\hat{H}\hat{\rho}$$

$$\hat{\rho} = e^{-\hat{H}t}$$

- We define the coordinate green's function (or density matrix by:

$$G(R \to R';t) = \left\langle R \left| e^{-t\hat{H}} \right| R' \right\rangle$$

Roughly the probability density of going from R_0 to R in "time" t. (but is it a probability?)

$$-\frac{\partial G(R_0 \to R;t)}{\partial t} = \hat{H}G(R' \to R;t)$$

- Properties:

$$G(R_0 \to R;0) = \delta(R_0 - R)$$

$$G(R_0 \to R;t) = \sum_{\alpha}\phi_\alpha^*(R_0)\phi_\alpha(R)e^{-tE_\alpha}$$

Froebinius Theorem

When can we consider the wavefunction as a probability?
First how about the Green's function?

$$G(R_0 \to R;0) = \delta(R_0 - R) \geq 0$$

Trotter's theorem implies it continues to be positive at all times.

$$G(R_0 \to R;t) \geq 0$$

But if we start with a non-negative function it will stay non-negative, and can be interpreted as a p.d.f.

Not true for all Hamiltonians (require off-diagonal matrix elements to be non-positive.) (not pseudopotentials, not magnetic fields.)

Only true for the bosonic ground state.

Monte Carlo process

- Now consider the variable "t" as a continuous time *(it is really imaginary time).*
- Take derivative with respect to time to get evolution.
- This is a diffusion + branching process.
- Justify in terms of Trotter's theorem.

$$-\frac{\partial \psi(R,t)}{\partial t} = (H - E_T)\psi(R,t)$$

$$H = -\sum_i \frac{\hbar^2}{2m_i}\nabla_i^2 + V(R)$$

$$\begin{cases} -\dfrac{\partial \psi(R,t)}{\partial t} = -\displaystyle\sum_i \frac{\hbar^2}{2m_i}\nabla_i^2 \psi(R,t) \\ -\dfrac{\partial \psi(R,t)}{\partial t} = (V(R) - E_T)\psi(R,t) \end{cases}$$

Requires interpretation of the wavefunction as a probability density.

But is it? Only in the boson ground state. Otherwise there are nodes. Come back to later.

Trotter's theorem

- How do we find the solution of: $\frac{d\hat{\rho}}{dt} = (A+B)\hat{\rho}$
- The operator solution is: $\hat{\rho} = e^{(A+B)t}$
- Trotter's theorem (1959): $\hat{\rho} = \lim_{n\to\infty}\left[e^{\frac{t}{n}\hat{A}}e^{\frac{t}{n}\hat{B}}\right]^n$
- Assumes that A,B and A+B are reasonable operators.

$$\left\langle R_0\left|\left[e^{\frac{t}{n}\hat{A}}e^{\frac{t}{n}\hat{B}}\right]^n\right|R_n\right\rangle = \left\langle R_0\left|e^{\frac{t}{n}\hat{A}}\right|R'_1\right\rangle\left\langle R'_1\left|e^{\frac{t}{n}\hat{B}}\right|R_1\right\rangle....\left\langle R_{n-1}\left|e^{\frac{t}{n}\hat{A}}\right|R'_n\right\rangle\left\langle R'_n\left|e^{\frac{t}{n}\hat{B}}\right|R_n\right\rangle$$

- This means we just have to figure out what each operator does independently and then alternate their effect. This is rigorous in the limit as n→∞.
- In the DMC case A is diffusion operator, B is a branching operator.
- Just like "molecular dynamics" At small time we evaluate each operator separately.

Evaluation of kinetic density matrix

$$\left\langle r\left|e^{-\tau\hat{T}}\right|r'\right\rangle = \sum_\alpha \phi^*_\alpha(r)\phi_\alpha(r')e^{-\tau T_\alpha}$$

In PBC eigenfunctions of $\hat{T} = \frac{1}{\sqrt{\Omega}}e^{-i\vec{k}\vec{r}}$

and eigenvalues are λk^2

$$\left\langle r\left|e^{-\tau\hat{T}}\right|r'\right\rangle = \sum_k \frac{1}{\Omega}e^{-i\vec{k}\vec{r}}e^{i\vec{k}\vec{r}'}e^{-\tau\lambda k^2}$$

convert to an integral

$$\left\langle r\left|e^{-\tau\hat{T}}\right|r'\right\rangle = \frac{1}{(2\pi)^3}\int dk e^{i\vec{k}(\vec{r}'-\vec{r})-\tau\lambda k^2} = (4\pi\lambda\tau)^{-3/2}e^{-(r-r')^2/4\lambda\tau}$$

Danger: makes assumption about boundaries and statistics.

This is a diffusion process.

Putting this together

- n is number of time slices.
- τ is the "time-step"

$$\hat{\rho} = e^{-\beta(\hat{T}+\hat{V})}$$

$$\hat{\rho} = \lim_{n\to\infty}\left[e^{-\tau\hat{T}}e^{-\tau\hat{V}}\right]^n$$

$$\tau = \beta/n$$

- V is "diagonal"

$$\left\langle r\left|e^{-\tau\hat{T}}\right|r'\right\rangle = \left(4\pi\lambda\tau\right)^{-3/2} e^{-(r-r')^2/4\lambda\tau}$$

$$\left\langle r\left|e^{-\tau\hat{V}}\right|r'\right\rangle = \delta(r-r')e^{-\tau V(r)}$$

$$\left\langle R_0 e^{-n\tau\hat{H}} R_n\right\rangle \sim \left\langle R_0\left|e^{-\tau\hat{T}}\right|R_1\right\rangle e^{-\tau V(R_1)} \ldots \left\langle R_{n-1}\left|e^{-\tau\hat{T}}\right|R_n\right\rangle e^{-\tau V(R_n)}$$

- Error at finite n comes from commutator is roughly: $e^{-\frac{\tau^2}{2}\left[\hat{T},\hat{V}\right]}$
- Diffusion preserves normalization but potential does not!

Basic DMC algorithm

- Construct an ensemble (population P(0)) sampled from the trial wavefunction. $\{R_1, R_{2,\ldots}, R_P\}$
- Go through ensemble and diffuse each one (timestep τ)

$$R'_k = R_k + \sqrt{2\lambda\tau}\zeta(t)$$ ← ndrn

- number of copies= $\underline{e^{-\tau(V(R)-E_T)} + u}$ ← uprn, floor function
- Trial energy E_T adjusted to keep population fixed.

$$E_0 = \lim_{t\to\infty}\frac{\int dRH\phi(R,t)}{\int dR\phi(R,t)} \approx \langle V(R)\rangle_{\phi(\infty)}$$

- Problems:
  - Branching is uncontrolled
  - Population unstable
  - What do we do about fermi statistics?

Sampling the normal distribution

- Inverse mapping is a little slow, also of infinite range.
- Trick: generate 2 at a time: r=(x,y)

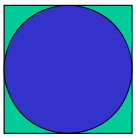
$$p(x,y)dxdy = (2\pi)^{-1}\exp(-\frac{r^2}{2}) = p(r)rdrd\theta$$

$$p(v)dv = \frac{1}{2}e^{-v/2} \text{ with } v = r^2$$

$$x = \sqrt{-2\ln(u_1)}\cos(2\pi u_2)$$

$$y = \sqrt{-2\ln(u_1)}\sin(2\pi u_2)$$

- Or sample angle using rejection technique:
  - Sample (x,y) in square
  - Accept if $x^2+y^2 < 1$
  - Normalize to get the correct r.



Code to sample normal distribution

Normal distribution $\langle x\rangle = x_0$ and $\langle (x-x_0)^2\rangle = \sigma^2$

```
1     x=sprng()-0.5
      y=sprng()-0.5
      r2=x*x+y*y
      if (r2>0.25) go to 1
      radius= sigma*sqrt (-2*ln(sprng())/r2)
      xnormal=x0+x*radius
      ynormal=y0+y*radius
```

- No trig functions, 1 log, 1 sqrt, 1 divide
- Mixes up regularity of random numbers
- Efficiency of angle generation is $4/\pi$.
- Gets 2 ndrn's each time.

Population Bias

- Having the right trial energy guarantees that population will on the average be stable, but fluctuations will always cause the population to either grow too large or too small.
- Various ways to control the population
- Suppose P_0 is the desired population and P(t) is the current population. How much do we have to adjust E_T to make $P(t+T)=P_0$?

$$P(t+T) = e^{-T(-\delta E_T)} P(t) = P_0$$

$$\delta E_T = \frac{\ln(P(t)/P_0)}{T}$$

- Feedback procedure: $\boxed{E_T = E_{T0} + \kappa \ln\left(P/P_0\right)}$

•There will be a (small) bias in the energy caused by a limited population.

Importance Sampling

Kalos 1970, Ceperley 1979

- Why should we sample the wavefunction? The physically correct pdf is $|\phi|^2$.
- Importance sample (multiply) by trial wave function.

$$f(R,t) \equiv \psi_T(R)\phi(R,t) \qquad \lim_{t\to\infty} f(R,t) \equiv \psi_T(R)\phi_0(R)$$

$$-\frac{\partial f(R,t)}{\partial t} = \psi_T(R) H\left[f(R,t)/\psi_T(R)\right]$$ Commute ψ through H

$$-\frac{\partial f(R,t)}{\partial t} = -\lambda\nabla^2 f - \lambda\nabla\left(2f\nabla\ln\psi_T(R)\right) + \left(\psi_T^{-1} H \psi_T\right) f(R,t)$$

Evolution = diffusion + drift + branching

- Use accept/reject step for more accurate evolution. make acceptance ratio>99% . Determines time step.
- We have three terms in the evolution equation. Trotter's theorem still applies.

Brownian Dynamics

Consider a big molecule in a solvent. In the high viscosity limit the "master equation" (Smoluchowski or Fokker-Planck eq.) is:

$$\frac{\partial \rho(R,t)}{\partial t} = D\nabla^2 \rho(R,t) - \beta D \nabla[F(R)\rho(R,t)]$$

$$R(t+\tau) = R(t) + \tau\beta D F(R(t)) + \eta(t)$$

$$\langle \eta(t) \rangle = 0 \qquad \langle \eta(t)^2 \rangle = 2\tau D$$

$$G(R \to R') = c \exp\left(-\frac{(R' - R - \beta D \tau F(R))^2}{2D\tau} \right)$$

Diffusion Quantum Monte Carlo without branching is the same as Brownian Dynamics. Use same techniques.

Green's function for a gradient

What is Green's function for the operator?

$$\vec{F}\vec{\nabla}$$

variables separate to 1D problems

Evolution equation for Green's function:

$$\frac{\partial G(x,t)}{\partial t} = -F \frac{\partial G(x,t)}{\partial x} \quad \text{solution} \quad G(x,t) = h(x - Ft)$$

This operator just causes probability distribution to drift in the direction of F.

Smoluchowski equation for Brownian motion it was the effect of gravitational field on the motion of colloids.

In practice, we limit the gradient so the walk is not pushed too far.

- To the pure diffusion algorithm we have added a drift step that pushes the random walk in directions of increasing trial function: $R' = R + 2\lambda\tau\nabla \ln \psi_T(R)$
- Branching is now controlled by the local energy

$$E_L(R) - E_T = \psi^{-1}(R)\hat{H}\psi(R) - E_T$$

- Because of zero variance principle, fluctuations are controlled.
- Cusp condition can limit infinities coming from singular potentials.
- We still determine E_T by keeping asymptotic population stable.

$$E_0 = \lim_{t\to\infty} \frac{\int dR\phi(R,t)H\psi_T(R)}{\int dRf(R,t)} \approx \left\langle E_\psi(R)\right\rangle_{f(\infty)}$$

- Must have accurate "time" evolution. Adding accept/reject step is a major improvement.

- Importanced sampled Green's function:

$$G(R \to R') = \frac{\psi(R')}{\psi(R)}\left\langle R\left|e^{-\tau H}\right|R'\right\rangle$$

- Exact property of DMC Green's function

$$\left|\Psi(R)\right|^2 G(R \to R') = \left|\Psi(R')\right|^2 G(R' \to R)$$

- We enforce detailed balance to decrease time step errors.

$$A(s \to s') = \min\left[1, \frac{G(s' \to s)\left|\psi(s')\right|^2}{G(s \to s')\left|\psi(s)\right|^2}\right]$$

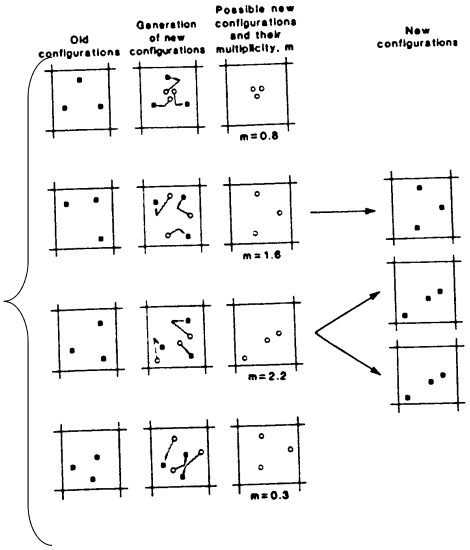
- VMC satisfies detailed balance.
- Typically we choose time step to have 99% acceptance ratio.
- Method gives exact result if either time step is zero or trial function is exact.

Schematic of DMC

Ensemble evolves according to

- Diffusion
- Drift
- branching

ensemble



Ceperley Projector Monte Carlo

DIFFUSION MONTE CARLO CODE

```
call initstate(s_old)                                  Initialize the ensemble of states
psi_old = psi(s_old)                                   Evaluate psi_trial
d_old = drift(s_old)                                   Evaluate grad psi_trial
LOOP {                                                 Loop over steps
   LOOP {                                              Loop over walkers
      call sample(s_old,s_new,T_new,d_old,1)           Sample new state from drifted Gaussian
      psi_new = psi (s_new)                            Evaluate psi_trial
      if (psi_new * psi_old < 0 ) {                    Check node crossing
         weight = 0                                    Kill walker if it crosses a node of psi_trial
      } else {
         d_old = drift (s_old)                         Evaluate grad psi_trial
         call sample(s_new,s_old,T_old,d_new,0)        Find transition prob. for going backward
         A = (p_new/T_new)/(p_old/T_old)               Acceptance prob.
         if(A > rand () ) {                            Accept the move
            s_old=s_new
            p_old=p_new
            naccept = naccept +1} }
      weight *= exp(- tau * local_energy (s_old) ) }   Update weight
   call reweight (s_old)                               Reweight ensemble
   call averages(s_old)                                Collect statistics
```

Ceperley Projector Monte Carlo

Local Markov process

- Let us make a course approximation to a molecular dynamics of a large molecule in the presence of a solvent that we want to eliminate.
- Particle will follow a continuous trajectory R(t).
- Let us assume that the movement is Markovian and local.
- Describe in terms of the Green's function $\rho(r,r',t)$
  This is probability density of a particle at (r,t) starting at (r',0).
- The master equation describes its time evolution:

Evolution $$\frac{d\rho}{dt} = \hat{O}\rho(r,r';t)$$

Initial condition $$\rho(r,r';0) = \delta(r-r')$$

General form of evolution

- What could we have for O?
- Not hard to prove that O must be a sum of first and second derivative operators. The most general form is:

$$\frac{\partial\rho(R,R';t)}{\partial t} = \vec{\nabla}\vec{D}(R)\left(\vec{\nabla}\rho(R,R';t) - \beta\vec{F}(R)\rho(R,R';t)\right)$$

- Clearly probability is conserved.
- Steady state solution is

$$\int dR\rho(R,R';t) = 1$$

$$\frac{d\rho(R,R';t)}{dt} = 0 \Rightarrow \vec{\nabla}\rho(R,R';t) = \beta\vec{F}(R)\rho(R,R';t)$$

$$\lim_{t\to\infty}\rho(R,R';t) = \frac{e^{-\beta V(R)}}{Z}$$

- How do we prove dynamics is unique?
- If it is a continuous walk-it has to be made of little random steps.
- Central limit theorem says that the only things that survive adding together many little steps are:
  - Mean value: this is the drift term
  - The variance: this is the diffusion term
- Hence the smart MC gaussian is the most general Markovian form, if we allow a general diffusion matrix.
- If we choose it to be isotropic and independent of position then D(R) is a constant.
  - it could vary in some media both in position and direction
  - also it could couple particles far apart, hydrodynamical effects.

Moment Expansion

$$\frac{\partial \rho(R,R';t)}{\partial t} = \vec{\nabla}\vec{D}(R)\left(\vec{\nabla}\rho(R,R';t) - \beta\vec{F}(R)\rho(R,R';t)\right)$$

$$\left\langle \frac{dR}{dt} \right\rangle = \int dR R \frac{d\,\rho(R,R';t)}{dt} =$$

$$= \int dR R\vec{\nabla}\vec{D}(R)\left(\vec{\nabla}\rho(R,R';t) - \beta\vec{F}(R)\rho(R,R';t)\right) =$$

$$= -\int dR\vec{D}(R)\left(\vec{\nabla}\rho(R,R';t) - \beta\vec{F}(R)\rho(R,R';t)\right) =$$

$$= -\int dR\left(-\vec{\nabla}\vec{D}(R) - \beta\vec{F}(R)\right)\rho(R,R';t)$$

$$\left\langle \frac{dR}{dt} \right\rangle\Bigg|_{t=0} = \vec{\nabla}\vec{D}(R') + \beta\vec{F}(R')$$

For the variance:

$$\frac{d\left\langle \overrightarrow{(R-R')}\overrightarrow{(R-R')} \right\rangle}{dt}\Bigg|_{t=0} = 2\overleftrightarrow{D}(R')$$

Mixed estimators

- Problem is that PMC samples the wrong distribution.
- OK for the energy
- Linear extrapolation helps correct this systematic error

$$\langle A\rangle_M \equiv \frac{\int dR\psi^*(R)A\phi(R)}{\int dR\psi^*(R)\phi(R)}$$

$$\langle A\rangle_o \equiv \frac{\int dR\phi^*(R)A\phi(R)}{\int dR\phi^*(R)\phi(R)}$$

$$\langle A\rangle_V \equiv \frac{\int dR\psi^*(R)A\psi(R)}{\int dR\psi^*(R)\psi(R)}$$

$$\langle A\rangle_0 \simeq 2\langle A\rangle_M - \langle A\rangle_V + O\left((\phi-\psi)^2\right)$$

$$\langle A\rangle_0 \simeq \frac{\langle A\rangle_M^2}{\langle A\rangle_V} + O\left((\phi-\psi)^2\right) \text{ for the density}$$

- Other solutions:
  - Maximum overlap
  - Forward walking
  - Reptation/path integrals

$$\langle A\rangle_M = \langle A\rangle_V \Rightarrow \int dR(\phi-\psi)^2 \text{ minimized wrt A}$$

Forward Walking

Kalos et al. 1974.

- Let's calculate the average population resulting from DMC starting from a single point R_0 after a time `t'.

$$P(R_0;t) = \int dR \frac{\psi(R)}{\psi(R_0)} \left\langle R \left| e^{-t(H-E_T)} \right| R_0 \right\rangle$$

expand the density matrix in terms of exact eigenstates

$$P(R_0;t) = \int dR \frac{\psi(R)}{\psi(R_0)} \sum_\alpha \phi_\alpha(R)\phi_\alpha(R_0) e^{-t(H-E_T)}$$

$$\lim_{t\to\infty} P(R_0;t) = \frac{\phi_0(R_0)}{\psi(R_0)} \langle \psi\phi_0 \rangle$$

- We can estimate the correction to the mixed estimator by weighting with the number of descendants of a given configuration.

$$\langle A\rangle_0 = \lim_{t\to\infty} \frac{1}{M} \sum_i P(R_i;t) A(R_i)$$

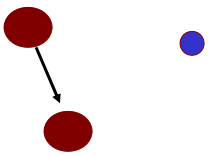
- Problem: the fluctuations in the weights eventually diverge. Don't make 't' too large.

Fusion sticking coefficient

Phys. Rev. ***A 31****, 1999 (1985).*

- Consider the 3 body system (μ d t)
- For the sticking coefficient, we need the exact wavefunction at the point where 2 nuclei are at the same position. (this is a singular point)

$$\psi\left(r_1 = r_2, r_3\right)$$



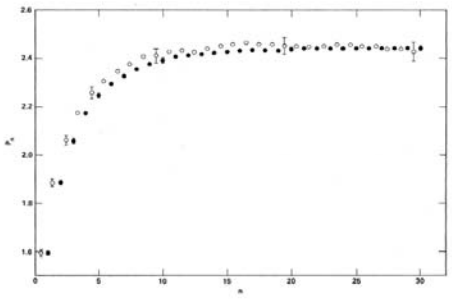
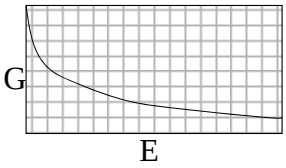


FIG. 1. The growth of the population vs the number of Monte Carlo generations using the population estimator (○) and the local energy estimator (●) at the triplet coalescence point (all three particles starting at the origin). The ○'s are shifted one-half generation to the left for clarity.

Ceperley Projector Monte Carlo

Other projector functions can be used

$$G(E) = \begin{cases} e^{-\tau(E-E_T)} & \text{Diffusion MC} \\ \left[1+\tau\left(E-E_T\right)\right]^{-1} & \text{Green's Function MC} \\ \left[1-\tau\left(E-E_T\right)\right] & \text{Power MC} \end{cases}$$



$G(E_T) = 1 \Rightarrow$ ground state remains after many iterations

$$\tau = -\left.\frac{dG}{dE}\right|_0 = \text{time step}$$

for all 3 cases: $\lim_{n\to\infty} G(E)^n = e^{-n\tau(E-E_T)}$

- Common effect on long-time (iteration) limit.
- 3rd choice generates a Krylov sequence. Only works for bounded spectra such as a lattice model.

Ceperley Projector Monte Carlo

Green's Function Monte Carlo

Kalos, Levesque, Verlet Phys. Rev. A9, 2178 (1974).

- It is possible to make a zero time-step-error method
- Works with the integral formulation of DMC

$$G(R,R') = \left\langle R \left| \left[1+\tau\left(H-E_T\right)\right]^{-1} \right| R' \right\rangle = \int_0^\infty \frac{d\beta}{\tau} e^{-\beta\left(\frac{1}{\tau}+H-E_T\right)}$$

- Sample time-step from Poisson distribution
- Express operator in a series expansion and sample the terms stochastically.

$$G(R,R') = H(R,R') + \int dR'' G(R,R'') K(R'',R')$$

- Recent Revival: "Continuous time Monte Carlo" for lattice models.

Fermions?

- How can we do fermion simulations? The initial condition can be made real but not positive (for more than 1 electron in the same spin state)
- In *transient estimate* or *released-node* methods one carries along the sign as a weight and samples the modulus.

$$\phi(t) = e^{-(\hat{H}-E_T)t} \operatorname{sign}(\phi(R,0)) \, |\phi(R,0)|$$

- Do not forbid crossing of the nodes, but carry along sign when walks cross.
- What's wrong with node release:
  - Because walks don't die at the nodes, the computational effort increases (bosonic noise)
  - The signal is in the cancellation which dominates

Monte Carlo can add but not subtract

Transient Estimate Approach

$$\Psi(\beta) = e^{-\frac{\beta}{2}H}\Psi$$

$$Z(\beta) = \langle \Psi(\beta)\Psi(\beta)\rangle = \langle \Psi e^{-\beta H}\Psi\rangle = \int dR_0 ... dR_p \Psi(R_0)\langle R_0 e^{-\tau H} R_1\rangle .... \langle R_{p-1} e^{-\tau H} R_p\rangle \Psi(R_p)$$

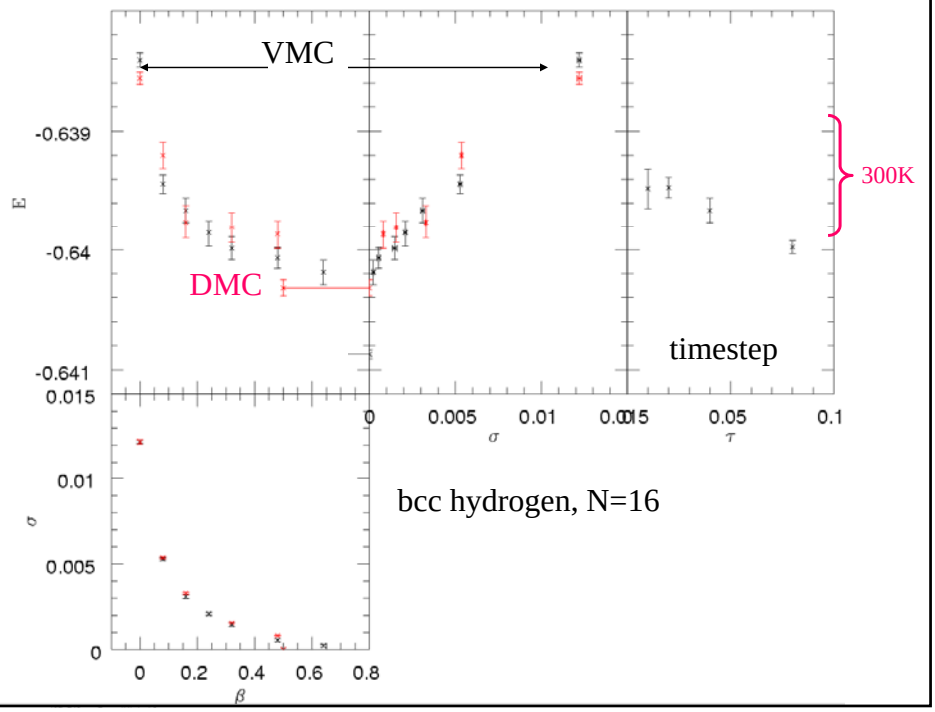
$$E(\beta) = \frac{\langle \Psi(\beta) H \Psi(\beta)\rangle}{\langle \Psi(\beta)\Psi(\beta)\rangle} = \langle E_L(R_0)\rangle_\beta \qquad \tau = \frac{\beta}{p}$$

- $\psi(\beta)$ converges to the exact ground state
- E is an upper bound converging to the exact answer monotonically

$$Z(\beta) = \int dR_0 ... dR_p \left|\Psi(R_0)\right| \langle R_0 e^{-\tau H} R_1\rangle .... \langle R_{p-1} e^{-\tau H} R_p\rangle \left|\Psi(R_p)\right| \sigma(R_0)\sigma(R_P)$$

$$\frac{Z_{\text{fermi}}}{Z_{\text{bose}}} = \langle \sigma(R_0)\sigma(R_P)\rangle$$

Ceperley Projector Monte Carlo

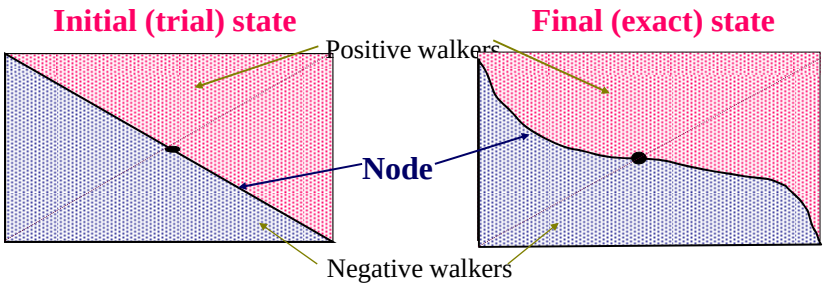


bcc hydrogen, N=16

Model fermion problem: Particle in a box

Symmetric potential: $V(\mathbf{r}) = V(-\mathbf{r})$

Antisymmetric state: $\phi(\mathbf{r}) = -\phi(-\mathbf{r})$

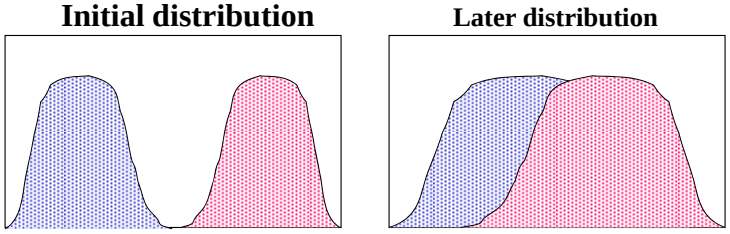


Sign of walkers fixed by initial position. They are allowed to diffuse freely. f(r)= number of positive-negative walkers. Node is dynamically established by diffusion process. (cancellation of positive and negative walkers.)

$$\langle E(t) \rangle = \frac{\sum \sigma(0)\sigma(t)E(t)}{\sum \sigma(0)\sigma(t)}$$

Ceperley Projector Monte Carlo

Scaling in Released-Node



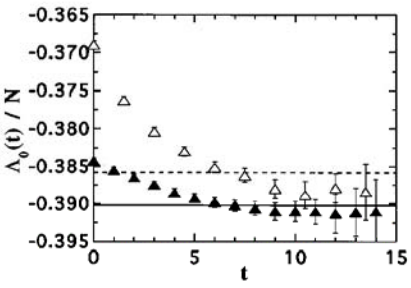
- At any point, positive and negative walkers will tend to cancel so the signal is drown out by the fluctuations.
- Signal/noise ratio is : $e^{-t[E_F - E_B]}$ **t**=projection time
  E_F and E_B are Fermion, Bose energy (proportional to **N**)
- Converges but at a slower rate. Higher accuracy, larger **t**.
- For general excited states:
  Exponential complexity! CPUtime $\propto \varepsilon^{-2(1+\frac{E_F}{E_g})} \approx \varepsilon^{-2N\frac{e_F}{E_g}}$
- Not a fermion problem but an excited state problem.
- Cancellation is difficult in high dimensions.

Ceperley Projector Monte Carlo

Exact fermion calculations

- Possible for the electron gas for up to 60 electrons.
- 2DEG at rs=1 N=26
- Transient estimate calculation with SJ and BF-3B trial functions.

$$\left\langle \Psi_T \left| e^{-tH} \right| \Psi_T \right\rangle$$



Ceperley Projector Monte Carlo

General statement of the "fermion problem"

- Given a system with **N** fermions and a known Hamiltonian and a property **O.** (usually the energy).
- How much time **T** will it take to estimate **O** to an accuracy **ε**? How does **T** scale with **N** and **ε**?
- If you can map the quantum system onto an equivalent problem in classical statistical mechanics then:

$$T \propto N^{\alpha} \varepsilon^{-2} \qquad \text{With } 0 < \alpha < 4$$

This would be a "solved" quantum problem!

- All approximations must be controlled!
- Algebraic scaling in N!

e.g. properties of Boltzmann or Bose systems in equilibrium.

Ceperley Projector Monte Carlo

"Solved Problems"

- 1-D problem. (simply forbid exchanges)
- Bosons and Boltzmanons at any temperature
- Some lattice models: Heisenberg model, 1/2 filled Hubbard model on bipartite lattice (Hirsch)
- Spin symmetric systems with purely attractive interactions: u<0 Hubbard model, nuclear Gaussian model.
- Harmonic oscillators or systems with many symmetries.
- Any problem with $\langle i|H|j\rangle \leq 0$
- Fermions in special boxes
- Other lattice models

- Kalos and coworkers have invented a pairing method but it is not clear whether it is approximation free and stable.

The sign problem

- The fermion problem is intellectually and technologically very important.
- Progress is possible but ***danger***-the problem maybe more subtle than you first might think. New ideas are needed.
- No fermion methods are perfect but QMC is competitive with other methods and more general.
- The fermion problem is one of a group of related problems in quantum mechanics (**e.g** dynamics).
- Feynman argues that general many-body quantum simulation is exponentially slow on a classical computer.
- Maybe we have to "solve" quantum problems using "analog" quantum computers: programmable quantum computers that can emulate any quantum system.

Fixed-node method

- Initial distribution is a pdf.
  It comes from a VMC simulation. $f(R,0) = |\psi_T(R)|^2$
- Drift term pushes walks away from the nodes.
- Impose the condition: $\phi(R) = 0$ when $\psi_T(R) = 0.$
- This is the fixed-node BC
- Will give an upper bound to the exact energy, the best upper bound consistent with the FNBC. $E_{FN} \geq E_0$

  $E_{FN} = E_0$ if $\phi_0(R)\psi(R) \geq 0$ all R

•f(R,t) has a discontinuous gradient at the nodal location.

•Accurate method because Bose correlations are done exactly.

•Scales well, like the VMC method, as N^3. Classical complexity.

•Can be generalized from the continuum to lattice finite temperature, magnetic fields, ...

•One needs trial functions with accurate nodes.

Proof of fixed-node theorem

- Suppose we solve S.E. in a subvolume V determined by the nodes of an antisymetric trial function.

$\hat{H}\phi_{FN} = E_{FN}\phi_{FN}$ inside V

Extend the solution to all space with the permutation operator.

$$\hat{\phi}_{FN}(R) \equiv \frac{1}{N!}\sum_P (-1)^P \phi_{FN}(PR)$$

Inside a given sub-volume only permutations of a given sign (±) contribute.

Hence the extend solution is non-zero.

Evaluate the variational energy the extended trial function.

$$E_0 \leq \frac{\sum_{PP'} (-1)^{P+P'} \int dR \phi_{FN}^*(PR)\hat{H}\phi_{FN}(P'R)}{\sum_{PP'} (-1)^{P+P'} \int dR \phi_{FN}^*(PR)\phi_{FN}(P'R)} = E_{FN} \leq E_{VMC}$$

Edges of volumes do not contribute to the integral since the extend solution vanishes there.

Nodal Properties

If we know the sign of the exact wavefunction (the nodes), we can solve the fermion problem with the fixed-node method.

- If $\phi(R)$ is real, nodes are $\phi(R)=0$ where R is the 3N dimensional vector.
- Nodes are a 3N-1 dimensional surface. (Do not confuse with single particle orbital nodes!)
- Coincidence points $\mathbf{r}_i = \mathbf{r}_j$ are 3N-3 dimensional hyper-planes
- In 1 spatial dimension these "points" exhaust the nodes: *fermion problem is easy to solve in 1D* with the "no crossing rule."
- Coincidence points (and other symmetries) only constrain nodes in higher dimensions, they do not determine them.
- The nodal surfaces define nodal volumes. How many nodal volumes are there? **Conjecture**: there are typically *only* 2 different volumes (+ and -) except in 1D. (but only demonstrated for free particles.)

Ceperley Projector Monte Carlo

Nodal Picture: 2d slice thru 322d space

- Free electron
- Other electrons

- Nodes pass thru their positions
- Divides space into 2 regions
- Wavelength given by interparticle spacing

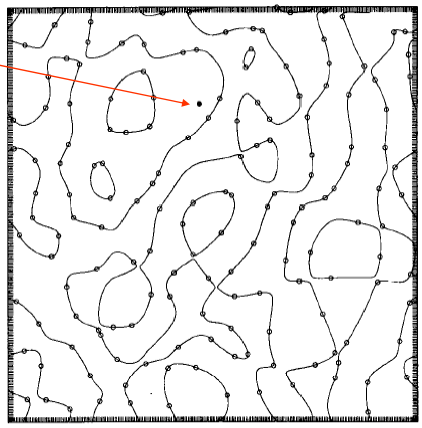


Fig. 3. A 2D cross section of the ground-state wave function of 161 free (polarized) fermions in a periodic square. All 161 particle positions were sampled using variational Monte Carlo from $\phi(R)^2$. The filled circle indicates the original position of the first particle. The other 160 particles are fixed at positions indicated by the open circles, and nodes of the wave function as a function of the position of the first particle are plotted. The resolution of the contouring program is approximately half of the fine scale shown around the border of the plot.

Ceperley Projector Monte Carlo

SPIN?

- How do we treat spin in QMC?
- For extended systems we use the S_z representation.
- We have a fixed number of up and down electrons and we antisymmetrize among electrons with the same spin.
- This leads to 2 Slater determinants.
- For a given trial function, its real part is also a trial function (but it may have different symmetries), for example momentum

$$\left(e^{ikr}, e^{-ikr}\right) \quad \text{or} \quad \left(\cos(kr), \sin(kr)\right)$$

- For the ground state, without magnetic fields, spin-orbit interaction we can always work with real functions.
- However, in some cases it may be better to work with complex functions.

Fixed-Phase method

Ortiz, Martin, DMC 1993

- Generalize the FN method to complex trial functions: $\Psi\left(R\right) = e^{-U(R)}$
- Since the Hamiltonian is Hermitian, the variational energy is real:

$$E_V = \frac{\int dR\, e^{-2\Re U(R)} \left[V(R) + \lambda \nabla^2 U(R) - \lambda \left[\Re \nabla U(R)\right]^2 + \lambda \left[\Im \nabla U(R)\right]^2 \right]}{\int dR\, e^{-2\Re U(R)}}$$

- We see only one place where the energy depends on the phase of the wavefunction.
- If we fix the phase, then we add this term to the potential energy. In a magnetic field we get also the vector potential.

$$\text{effective potential} = V(R) + \sum_i \lambda_i \left[A(r_i) + \Im \nabla_i U(R) \right]^2$$

- We can now do VMC or DMC and get upper bounds as before.
- The imaginary part of the local energy will not be zero unless the right phase is used.
- Used for twisted boundary conditions, magnetic fields, vortices, phonons, spin states, ...

Problem with core electrons

- Bad scaling in both VMC and DMC
- In VMC, energy fluctuations from core dominate the calculation
- In DMC, time step will be controlled by core dynamics
- Solution is to eliminate core states by a pseudopotential

- Conventional solution: semi-local form

$$\left\langle r\left|\hat{v}_{e-core}\right|r'\right\rangle = v_{local}(r)\delta(r-r') + \sum_l v_l(r)P_l(\cos(r\cdot r'))$$

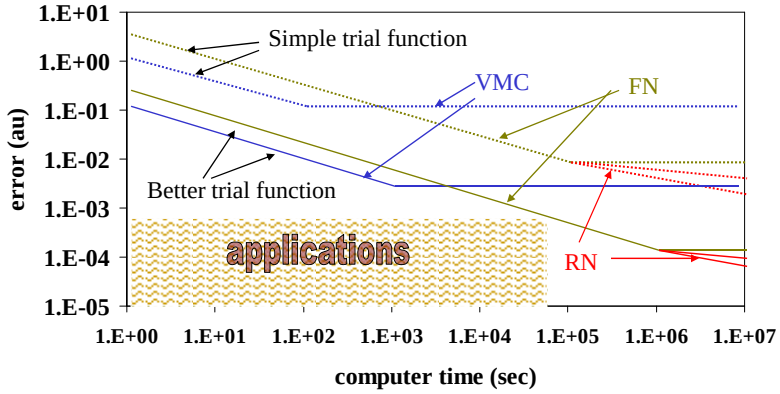
- Ensures that valence electrons go into well defined valence states with the wavefunction and energy for each angular momentum state prescribed.
- PP is non-local: OK for VMC. Leads to an extra MC integral. But DMC uses a locality approximation and good trial functions. Extra approximation.

Solid state / chemical applications

- Use LDA derived pseudopotentials
- Take orbitals from other methods:
  - "Gaussian-xx" give orbitals for molecules
  - DFT-PW codes give orbitals for extended systems
- HF is slightly better because of self-interaction effects within DFT
- Multiply by a Jastrow function (electron gas or otherwise). Can include higher order e-e-n terms
- Must add a compensating e-n term in order to cancel out purely repulsive character of e-e correlation.
- Assuming LDA density is correct, this can be done by making sure VMC electron density=LDA electron density. (Fahy correction).

Summary of T=0 methods:

Variational(VMC), Fixed-node(FN), Released-node(RN)



Ceperley Projector Monte Carlo

Problems with projector methods

- Fixed-node is a super-variational method
- DMC dynamics is determined by Hamiltonian
- Zero-variance principle allows very accurate calculation of ground state energy if trial function is good.
- Projector methods need a trial wavefunction for accuracy. They are essentially methods that perturb from the trial function to the exact function. (Note: if you don't use a trial function, you are perturbing from the ideal gas)
- Difficulty calculating properties other than energy. We must use "extrapolated estimators" or "forward walking".

$$f(R,\infty) = \phi_0(R)\psi_T(R) \quad \text{not} \quad |\phi_0(R)|^2$$

- Bad for phase transitions and finite temperature, complex systems.
- Path Integral MC solves some of these problems.

Ceperley Projector Monte Carlo