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

Problems with Variational MC

- 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

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. →

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

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

- Then the evolution is simple in this basis.

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

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

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

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

- How to carry out on the computer?

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:

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

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

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.

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.

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

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

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

Trotter's formula

- 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 formula (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.

Putting this together

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

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

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

$$\tau = \beta/n$$

- V is "diagonal"

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

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

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

- 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)

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

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

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

- 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)

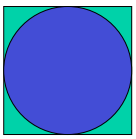
$$p(x,y)dxdy = (2\pi)^{-1}\exp(-\frac{r^2}{2}) = p(r)rdrd\theta$$

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

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

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

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



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.

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.

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

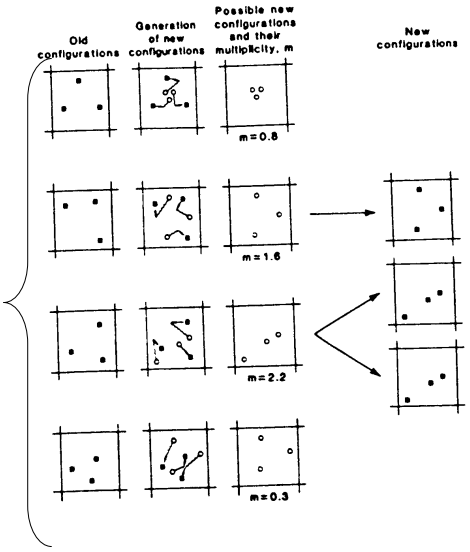
- VMC satisfies detailed balance.
- Typically we choose time step to have 99% acceptance ratio.
- Method gives exact result if either time step is zero or trial function is exact.

Schematic of DMC

Ensemble evolves according to

- Diffusion
- Drift
- branching

ensemble



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)
         if(A > rand () ) {                             Acceptance prob.
            s_old=s_new
            p_old=p_new                                 Accept the move
            naccept = naccept +1} }
      weight *= exp(- tau * local_energy (s_old) ) }    Update weight
   call  reweight (s_old)                               Reweight ensemble
   call  averages(s_old)                                Collect statistics
```


Mixed estimators

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

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

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

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

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

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

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

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

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.

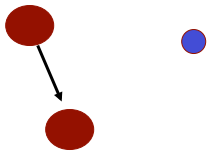
Don't make 't' too large.

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)

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



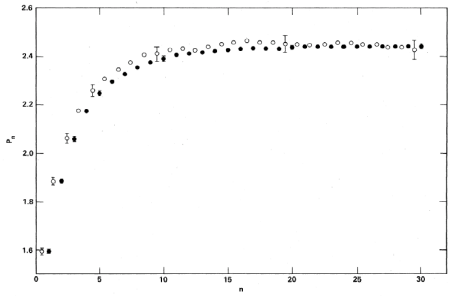
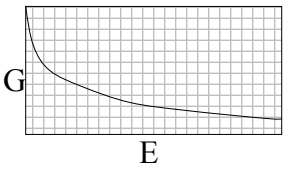


FIG. 1. The growth of the population vs the number of Monte Carlo generations using the population estimator (○) and the local energy estimator (●) at the triplet coalescence point (all three particles starting at the origin). The ○'s are shifted one-half generation to the left for clarity.

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.

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 and DMFT solver.