Quantum Monte Carlo Methods

Manish Lohani
Freie Universität Berlin

Top 10 algorithms of 20th century

- Metropolis Algorithm for Monte Carlo
- Simplex Method for Linear Programming
- Krylov Subspace Iteration Methods
- The Decompositional Approach to Matrix Computations
- The Fortran Optimizing Compiler
- QR Algorithm for Computing Eigenvalues
- Quicksort Algorithm for Sorting
- Fast Fourier Transform
- Integer Relation Detection
- Fast Multipole Method

https://www.computer.org/csdl/magazine/cs/2000/01/c1022/13rRUxBJhBm

from *SIAM News*, Volume 33, Number 4

The Best of the 20th Century: Editors Name Top 10 Algorithms

By Barry A. Cipra

Algos is the Greek word for pain. *Algor* is Latin, to be cold. Neither is the root for *algorithm*, which stems instead from al-Khwarizmi, the name of the ninth-century Arab scholar whose book *al-jabr wa'l muqabalah* devolved into today's high school algebra textbooks. Al-Khwarizmi stressed the importance of methodical procedures for solving problems. Were he around today, he'd no doubt be impressed by the advances in his eponymous approach.

Some of the very best algorithms of the computer age are highlighted in the January/February 2000 issue of *Computing in Science & Engineering*, a joint publication of the American Institute of Physics and the IEEE Computer Society. Guest editors Jack Don-garra of the University of Tennessee and Oak Ridge National Laboratory and Fran-cis Sullivan of the Center for Comput-ing Sciences at the Institute for Defense Analyses put togeth-er a list they call the "Top Ten Algorithms of the Century."

"We tried to assemble the 10 al-gorithms with the greatest influence on the development and practice of science and engineering in the 20th century," Dongarra and Sullivan write. As with any top-10 list, their selections—and non-selections—are bound to be controversial, they acknowledge. When it comes to picking the algorithmic best, there seems to be no best algorithm.

Without further ado, here's the CiSE top-10 list, in chronological order. (Dates and names associated with the algorithms should be read as first-order approximations. Most algorithms take shape over time, with many contributors.)

1946: John von Neumann, Stan Ulam, and Nick Metropolis, all at the Los Alamos Scientific Laboratory, cook up the Metropolis algorithm, also known as the **Monte Carlo method**.

The Metropolis algorithm aims to obtain approximate solutions to numerical problems with unmanageably many degrees of freedom and to combinatorial problems of factorial size, by mimicking a random process. Given the digital computer's reputation for deterministic calculation, it's fitting that one of its earliest applications was the generation of random numbers.

1947: George Dantzig, at the RAND Corporation, creates the **simplex method for linear programming**.

In terms of widespread application, Dantzig's algorithm is one of the most successful of all time: Linear programming dominates the world of industry, where economic survival depends on the ability to optimize within budgetary and other constraints. (Of course, the "real" problems of industry are often nonlinear; the use of linear programming is sometimes dictated by the computational budget.) The simplex method is an elegant way of arriving at optimal answers. Although theoretically susceptible to exponential delays, the algorithm

http://www.uta.edu/faculty/rcli/TopTen/topten.pdf

Motivation: Expectation Value of Observable

Expectation Value of Observable f(x) from a Probability Distribution ρ(x)

$$E_\rho(f) = \frac{\int_{\mathbb{R}} \rho(x) f(x) dx}{\int_{\mathbb{R}} \rho(x) dx}$$

Examples:

- Canonical ensemble

$$\rho(q,p) = \frac{e^{-\beta E(q,p)}}{\int e^{-\beta E(q,p)} d^{3N}q \; d^{3N}p}$$

- Quantum system

$$< \Psi_{trial} | \widehat{H} | \Psi_{trial} > \; = \frac{\int_{\Omega^N} \Psi^*_{trial} \widehat{H} \, \Psi_{trial} d^N r}{\int_{\Omega^N} \Psi^*_{trial} \Psi_{trial} d^N r}$$

To solve these, we need some Numerical Integration techniques

Numerical Integration and Limitation with Higher dimension

Numerical Integration Techniques

- Simpsons Rule
- Trapezoidal Rule

Simpson's Rule with M points per dimension

- In one dimension, Error is $O(M^{-4})$
- In d-dimension, $N = M^d$ points and Error is $O(N^{-4/d})$

So,how to deal with higher dimensions ?

Monte carlo Methods

Brief History

18th Century: Buffon's needle problem
1940's: Ulam, Neumann and Metropolis to study diffusion of neutrons, Used in the Manhattan Project
1950s : at Los Alamos for development of the hydrogen bomb

Computational methods to approximate quantities in a way that involves randomness(stochastic)

Applications

- Numerical integration(high dimensional)
- sampling from high dimensional probability distribution
- Optimization problems(gives flexibility to jump from local minimum)

Simple Monte Carlo: Mathematical way

- Consider the following expectation of f(x)

$$E_\rho(f) = \int_{\mathbb{R}} \rho(x) f(x) dx \Big/ \int_{\mathbb{R}} \rho(x) dx$$

ρ(x) : Probability distribution
x$\in \mathbb{R}$: High-dimensional state vector
f(x) : Function of state / observable

Algorithm:

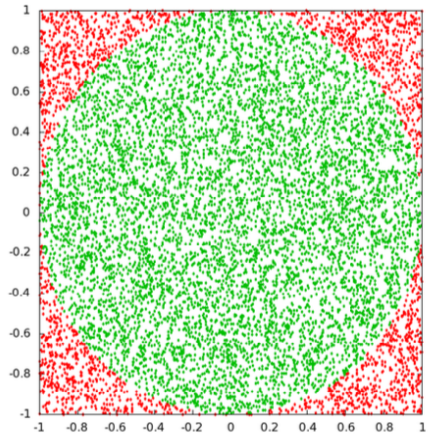
1. Sample N times randomly from $\rho(x)$
2. Estimate $E[f] \approx N^{-1} \sum_i f(x_i)$, if N is very large

- **Statistical Error:**

$$\Delta = \sqrt{\frac{Var(f)}{N}} \propto N^{-1/2}$$

In d $>$ 8 dimensions Monte Carlo works better than Simpsons

Simple Monte Carlo : Calculation of π



$$\pi = 4\frac{A_{\text{circle}}}{A_{\text{square}}}$$

Steps:

1. generate a uniformly distributed sample of N points in the surrounding square
2. count the number of points that end up in the circle

Simple Monte Carlo : Calculation of π

Mathematically,

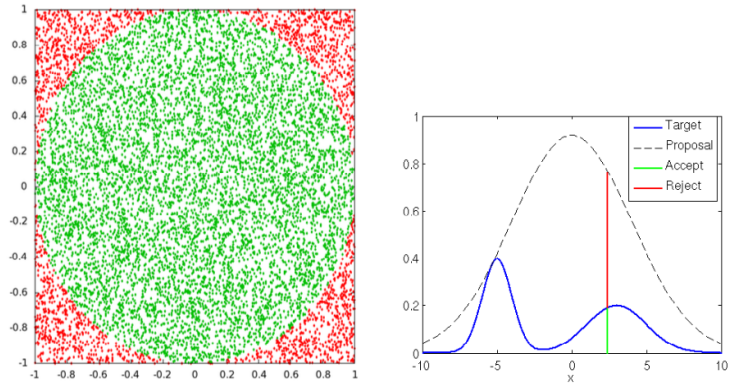
$$f(x) = \begin{cases} 1 \text{ or accept}, & \text{if inside of circle}, \\ 0 \text{ or reject}, & \text{if outside} \end{cases}$$

if N is very large then

$$\frac{A_{\text{circle}}}{A_{\text{square}}} \approx N^{-1} \sum_i f(x)_i = \frac{N_c}{N}$$

So,How we can generalise this concept ?

Rejection Sampling



- Generate samples from desired probability distribution $\rho(x)$
- Take proposal probability distribution g(x) (an upper bound) that is easy to sample such that $\rho(x) \leq Mg(x)$

Mathematically,

1. S is empty
2. For i = 1..., N
   (a) Sample x_i from g(x)
   (b) accept x with $p_{\text{acc}}(x) = \rho(x)/Mg(x)$ and add to S
   otherwise Reject

Rejection Sampling

But calculating M itself is an intractable problem in high-dimensional spaces

And, what if, the distribution has finite value in very small regions and near zero in other regions : **Very Low Efficient**

Importance Sampling

- Normalised PDF g(x) is selected such that the 'important' regions in desired probability space has a higher selection probability.
- Variance Reduction technique

$$\int_{\mathbb{R}} \rho(x)dx = \int_{\mathbb{R}} g(x)(\rho(x)/g(x))dx$$
$$= E[\rho(x)/g(x)]$$

Estimation :

$$E[f] = \frac{\sum_i f(x_i)\frac{\rho(x_i)}{g(x_i)}}{\sum_i \frac{\rho(x_)i}{g(x)_i}}$$

Problems of Monte Carlo

Importance Sampling

- Large Variance(if Under-Sampling)
- To find g(x) for high dimension PDF

In Monte Carlo methods, we need to calculate Normalisation Constant
e.g. statistical Mechanics

$$\rho(q,p) = \frac{e^{-\beta E(q,p)}}{\int e^{-\beta E(q,p)} d^{3N}q \; d^{3N}p}$$

So, how we can sample from High dimensional probability space efficiently?

Markov Chains and Metropolis Algorithm

Markov Chain or random walk

- **Conditional probabilistic model** Consist of sequence of events x_1, x_2, x_3..., x_N such that

$$P(X_{new}) = f(X_{new}, X_{last})$$

OR

$$P(X_t|X_1, X_2, ..., X_{t-1}) = P(X_t|X_{t-1})$$

- f defines propagation logic of "walk"

Properties of Markov Chain to generate sample points :

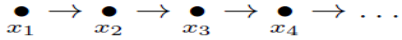
- $\rho(x) \neq 0$ (Accessibility)
- Ergodic
- Non-periodic

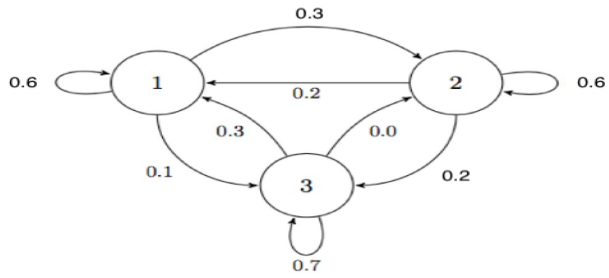
Note: Length of chain is common source of error

Now, how to implement this model?

Markov Chain or random walk

Markov chain





Stochastic Matrix

$$\mathbf{P} = \begin{pmatrix} P_{11} = P(1 \to 1|1) & \dots & P(1 \to N|1) \\ \vdots & \ddots & \vdots \\ P(N \to 1|N) & \dots & P(N \to N|N) \end{pmatrix}$$

$P(x_i \to x_{i+1} | x_i)$ is conditional probability of $x_i \to x_{i+1}$

Properties of Stochastic Matrix

1. $P_{ij} \leq 1$
2. $\sum_j P(x_{ij}) = 1$

Markov chain Propagation

ρ^0 :initial state probability vector $\begin{pmatrix} \rho_1^0 \\ \rho_2^0 \\ . \\ . \\ \rho_n^0 \end{pmatrix}$

ρ_i^0 denotes initial possibility of finding the system in state i

Markov chain step:

Propagated Distribution: ${\rho^{(k+1)}}^T = {\rho^{(k)}}^T \ \mathrm{P}$

${\rho^{(1)}}^T = {\rho^{(0)}}^T \mathrm{P}$

${\rho^{(2)}}^T = {\rho^{(1)}}^T \mathrm{P}$

.

.

${\rho^{(k)}}^T = {\rho^{(k-1)}}^T \mathrm{P}$

Markov chain Propagation

Asymptotic Behavior

- $\rho = \lim_{k \to \infty} \rho_0^T \mathrm{P}^k$ follows $\rho^T = \rho^T \mathrm{P}$
- also called as Stationary Distribution
- ρ_0 can be arbitrarily picked.

Now, how to use this concept in practical sampling and calculations ?

Detailed Balance and its connection with Stochastic Matrix

- **Detailed Balance:** "At equilibrium, ensemble should be invariant with respect to transformation"

$$\rho_i P_{i \to j} = \rho_j P_{j \to i}$$

OR

$$\rho_i P_{ij} = \rho_j P_{ji}$$

- **For $\rho(x)$ to be stationary distribution of Markov chain:**

P must be eigen vector of ρ

$$\rho^{(j)^T} = \rho^{(i)^T} P$$

OR

$$\sum_{i=1}^{N} \rho_i P_{ij} = \rho_j$$

Detailed Balance and its connection with Stochastic Matrix

- Detailed balance verify the above condition

$$\sum_{i=1}^{N} \rho_i P_{ij} = \sum_{i=1}^{N} \rho_j P_{ji} = \rho_j \sum_{i=1}^{N} P_{ji} = \rho_j$$

- So we need to construct Markov chain that obeys detailed balance.
- Using Detailed Balance:

$$\frac{\rho_j}{\rho_i} = \frac{P_{ij}}{P_{ji}}$$

- **Non-dependence on Normalisation Constant**

Metropolis Algorithm(1953)

THE JOURNAL OF CHEMICAL PHYSICS VOLUME 21, NUMBER 6 JUNE, 1953

Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

AND

EDWARD TELLER,* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.

Metropolis Procedure

- Fix the desired sampling Distribution $\rho(x)$
- Decide the proposal Criteria t_{ij} for $x_i \rightarrow x_j$
- Initialize with state $x_i = x_0$
- (a) propose x_1
  (b) calculate

$$A = \frac{\rho(x_1)}{\rho(x_0)}$$

- Acceptance-Rejection Step:
  If $A \geq 1$:
  Accept the move
  Else Generate $\eta \sim \mathcal{N}[0,1]$:
  1. if $\eta < A$, Accept the move
  2. if $\eta \geq A$, Reject and set $x_j = x_0$
     Repeat with new proposal x_1 until accepted

  If Accepted, set $x_j = x_1$ and Repeat the procedure for next move

Further Advancement : Hastings, W.K. (1970)

Decomposition of Transition Process

$$\text{Transition} = \text{Proposal} + \text{Acceptance}$$

$$P_{ij} = t_{ij} * a_{ij}$$

t_{ij} : Proposal conditional probability of a transition i $\rightarrow$j given i
a_{ij} : Acceptance probability of choice of t_{ij}

Using Detailed Balance :

$$\frac{\rho_j}{\rho_i} = \frac{P_{ij}}{P_{ji}} = \frac{t_{ij} * a_{ij}}{t_{ji} * a_{ji}}$$

OR

$$\frac{a_{ij}}{a_{ji}} = \frac{\rho_j t_{ij}}{\rho_i t_{ji}}$$

The ratio $\frac{a_{ij}}{a_{ji}}$ is called "Acceptance ratio of move"

Decomposition of Transition Process

- **Only need to calculate ratio** $\frac{\rho_j}{\rho_i}$
- **Non-dependence on Normalisation Constant**
- choose a function a_{ij} that satisfies the detailed balance
- We can choose different functions here
- The most common choice is Metropolis criteria

$$a_{ij}\left(x_i \rightarrow x_{i+1}\right) = min\left(1, \frac{\rho_j t_{ij}}{\rho_i t_{ji}}\right)$$

Generalised Metropolis Procedure

- Fix the desired sampling Distribution $\rho(x)$
- Decide the proposal Criteria t_{ij} for $x_i \rightarrow x_j$
- Initialize with state $x_i = x_0$
- (a) propose x_1
  (b) calculate

$$A = \frac{\rho(x_1)}{\rho(x_0)} \frac{t(x_1 \rightarrow x_0)}{t(x_0 \rightarrow x_1)}$$

- Acceptance-Rejection Step:
  If $A \geq 1$:
  Accept the move and set $x_j = x_1$
  Else Generate $\eta \sim \mathcal{N}[0,1]$:
  1. if $\eta < A$, Accept and set $x_j = x_1$
  2. if $\eta \geq A$, Reject and set $x_j = x_0$

  Repeat with new proposal x_1 until accepted

  If Accepted and Repeat the procedure

Some points to remember

- Accepted chain of States x_0, x_1, x_2..., x_T approach the desired Distribution.
- Accuracy of results depends on:
  1. similarity of proposal distribution with Desired distribution
  2. the length of chain (T)
- examples of proposed Criteria:
  1. uniform distribution with a volume centered at x_i
  2. Gaussian distribution centered on x_i

Quantum Monte Carlo

Variational Quantum Monte Carlo

PHYSICAL REVIEW VOLUME 138, NUMBER 2A 19 APRIL 1965

Ground State of Liquid He⁴†

W. L. McMillan*
Department of Physics, University of Illinois, Urbana, Illinois
(Received 16 November 1964)

The properties of the ground state of liquid He⁴ are studied using a variational wave function of the form $\Pi_{i<j} f(r_{ij})$. The Lennard-Jones 12-6 potential is used with parameters determined from the gas data by deBoer and Michiels. The configuration space integrals are performed by a Monte Carlo technique for 32 and 108 atoms in a cube with periodic boundary conditions. With $f(r) = \exp[-(2.6\text{ Å}/r)^5]$, the ground-state energy is found to be -0.78×10^{-15} ergs/atom, which is 20% above the experimental value. The liquid structure factor and the two-particle correlation function are in reasonably good agreement with the x-ray and neutron scattering experiments.

INTRODUCTION

THE interacting Bose gas[1–5] has been the subject of intensive theoretical investigation as a microscopic model for the behavior of liquid He⁴. Successful calculations have been performed for the Bose gas with weak repulsive interactions[1] and for the low-density Bose gas with short-range repulsive interactions.[2] This work provides a beautiful model of the superfluid behavior of the interacting Bose gas and of the existence of phonons and quantized vortices, but it does not permit one to perform quantitative calculations for

Estimates of the potential have been made theoretically from the atomic structure of helium and empirically from the gas kinetic data. Theoretically, it consists of two parts: the attractive Van der Waals interaction at[10,11] large r and the repulsive Coulomb and exchange interactions[12] for small r, where the wave functions of the two atoms overlap appreciably. deBoer and Michiels[13] have obtained an empirical interaction by fitting the parameters of the Lennard-Jones 12-6 potential to the experimental values of the second virial coefficient above 60°K. They found

- First QMC calculation
- PhD Thesis of Mcmillan (1964)
- Intuitive Generalisation of classical MC technique
- VMC Calculation of ground state of liquid Helium 4 (Bosons)
- Zero temprature method

Variational Quantum Monte Carlo

PHYSICAL REVIEW B VOLUME 16, NUMBER 7 1 OCTOBER 1977

Monte Carlo simulation of a many-fermion study*

D. Ceperley and G. V. Chester
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853

M. H. Kalos
Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012

(Received 15 December 1976)

The Metropolis Monte Carlo method is used to sample the square of an antisymmetric wave function composed of a product of a Jastrow wave function and a number of Slater determinants. We calculate variational energies for ^{3}He and several models of neutron matter. The first-order Wu-Feenberg expansion is shown always to underestimate the energy, sometimes seriously. The phase diagram for ground-state Yukawa matter is determined. There is a class of Yukawa potentials which do not lead to a crystal phase at any density.

I. INTRODUCTION

In recent years, there have been many Monte Carlo simulations of Bose systems. The simplest of these uses the Metropolis method from classical statistical mechanics to sample a variational

Other ground-state properties are also computed; their values are not stationary but are expected to be representative of the ground state if the trial wave function is good.

Direct sampling of the N-particle probability yields a rigorous upper bound to the energy; all

- Generalised to Fermions (Helium 3)
- D. Ceperley , G. V. Chester and M. H. Kalos (1976)
- Added Slater Determinants in Trial wave function

Variational Quantum Monte Carlo

Goal: To solve the Many-body Hamiltonian to get exact ground state properties

$$\widehat{H}\Psi(R) = E\Psi(R)$$

R : Quantum degree of freedom(Position, Spin)

- Generalised version of Classical MCMC
- Uses Metropolis MC sampling
- Variational Theorm is the basis of VQMC

Variational Theorm

Expectation value of Hamiltonian of a trial wave function is always an upper bound of ground state energy

$$E_v = \frac{\int_{\Omega^N} \Psi_{trial}(R)^* \widehat{H} \Psi_{trial}(R) d^N r}{\int_{\Omega^N} \Psi_{trial}(R)^* \Psi_{trial}(R) d^N r} \geq E_0$$

E_0 : Ground state Energy

Note: Ψ_{trial} taken as not normalised

Now, How can we formulate this into Monte Carlo sampling: Mcmillan Intution

VQMC: Mcmillan Intution

We can write E_v into below format :

$$E_v = \int \frac{|\psi(R)|^2 \frac{\widehat{H}\psi(R)}{\psi(R)}}{\int |\psi(R)|^2} = \int_R \rho(R) E_L(R) dR$$

$\rho(R) = \frac{|\psi(R)|^2}{\int |\psi(R)|^2}$

$\mathrm{E}_L(\mathrm{R}) = \frac{\widehat{H}\psi(R)}{\psi(R)}$ **[Local Energy]**

Note: Here $\Psi(R)$ is $\Psi trial(R)$

This formulation of E_v gives exact mapping of Expectation value

VQMC :Mcmillan Intution

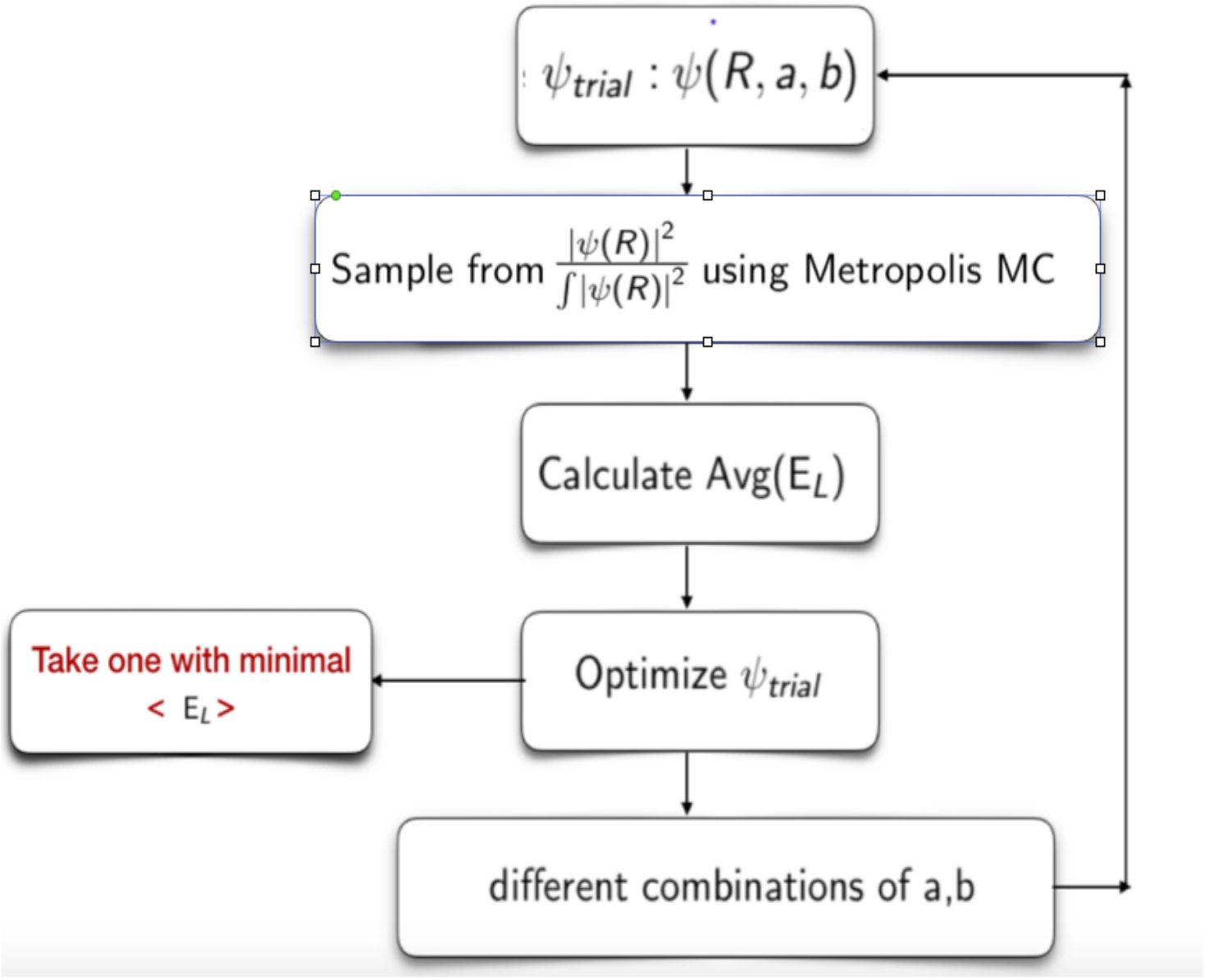
Comparing with classical MC:

$$< g(x) >= \int_{\rho(x)} \rho(x) g(x) dx = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} g(x_i)$$

we can write

$$< E_L(R) >= \int \frac{|\psi(R)|^2 \frac{\widehat{H}\psi(R)}{\psi(R)}}{\int |\psi(R)|^2} = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \frac{\widehat{H}\psi(R_i)}{\psi(R_i)}$$

VQMC: Algorithm



VQMC: Choice of ψ

Properties:

- Normalizable
- Finite Variance
- Lower the variance ,better the wave function

$$\sigma^2 = \int_R \rho(R)(E_v - E_L)^2 dR$$

- Generally taken as product of pair functions

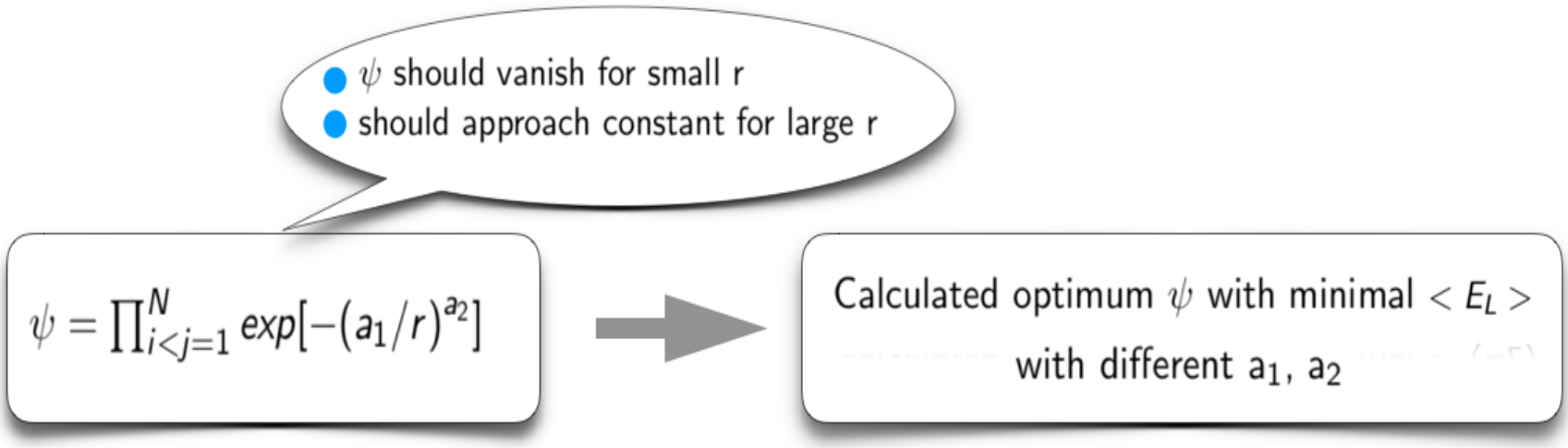
$$\psi = \prod_{i<j=1}^{N} f(r_{ij})$$

Choice of ψ: Liquid He4: Mcmillan Approach

Jastrow Wave function (1950's):

$$\psi = \prod_{i<j=1}^{N} exp[-u(r_{ij})]$$

$u(r_{ij})$ = **Correlation function**



Choice of ψ: Liquid He4(Boson): Results

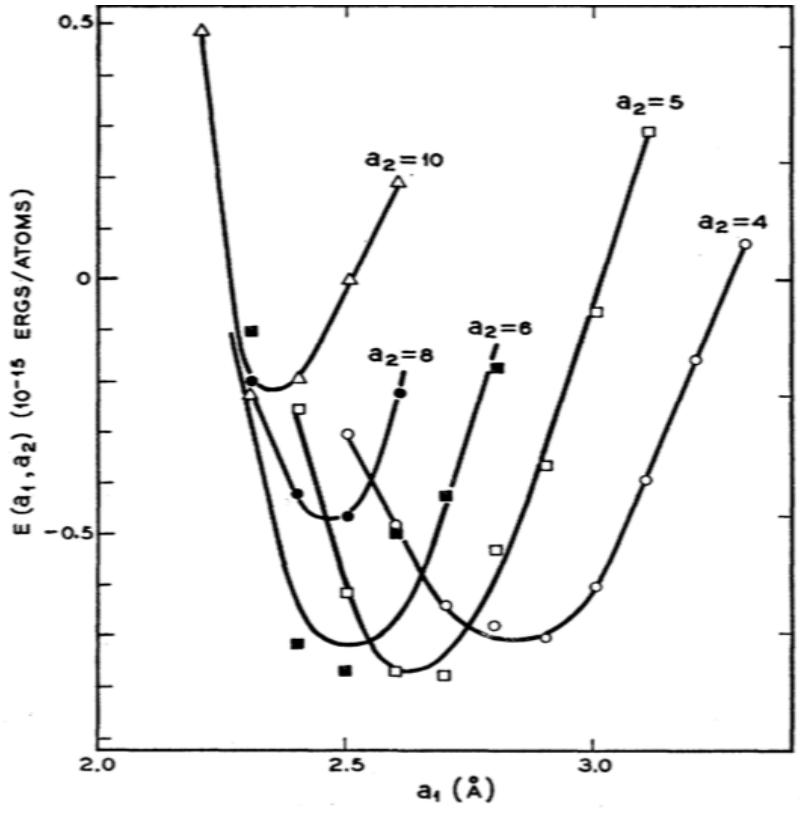


FIG. 1. The ground-state energy versus the two variational parameters, a_1 and a_2, for the density $\rho = 2.2 \times 10^{-22}$ atoms/cc as found from the preliminary calculations.

a1 = 2.51 and a2 = 5

Standard Deviation of Ground state energy : 2%

Choice of ψ: Fermions (Ceperley,Chester,Kalos)

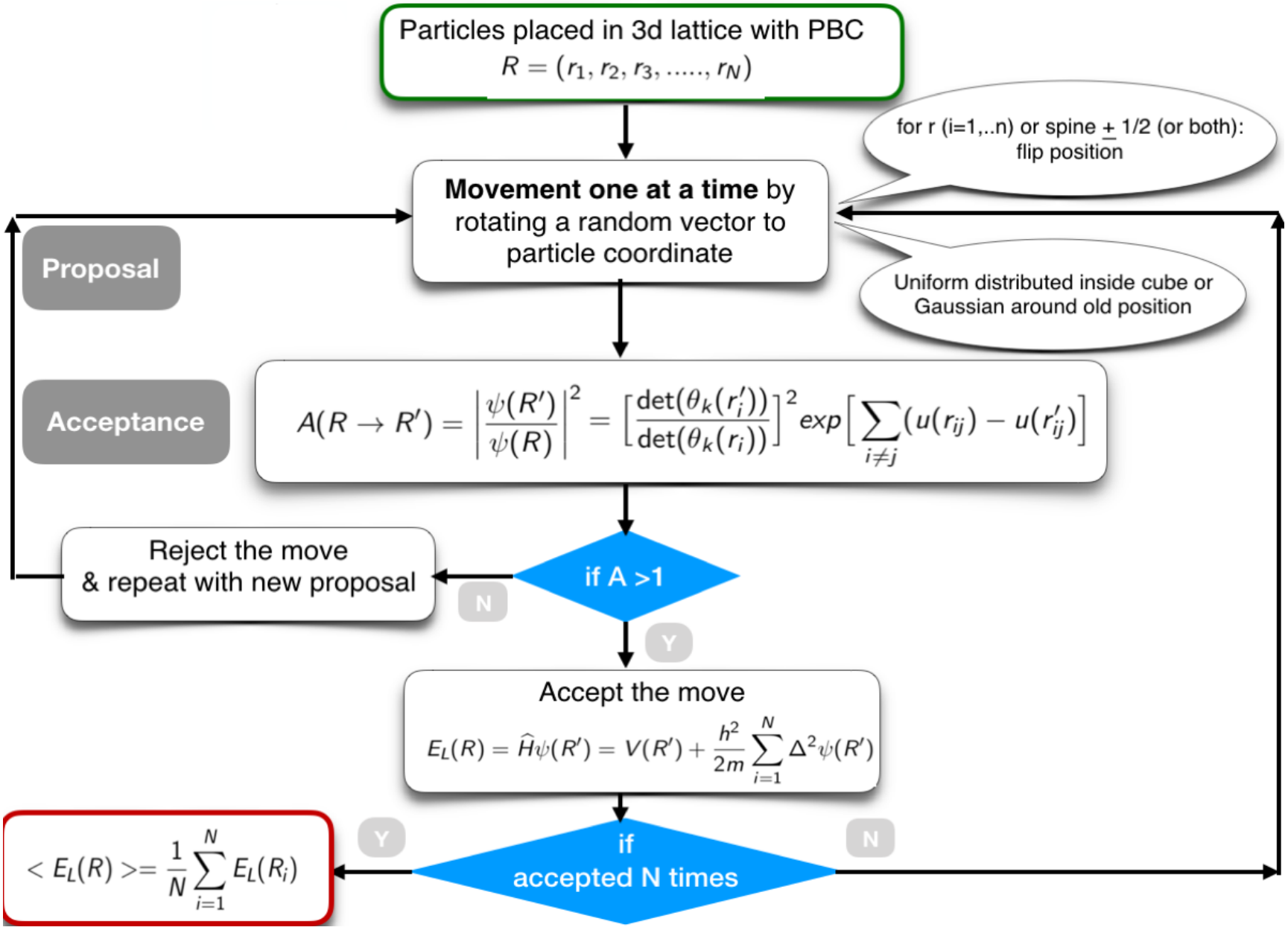
Fermions:

$$\psi = \det(\theta_k(r_i)) \; exp(-\sum_{i \neq j} u(r_{ij}))$$

Slater Determinant (No correlation)

Jastrow Factor (Gives correlation)

Implementation



Optimisation of ψ to ψ_0:

Examples:

1. Define Jastrow factor as

$$exp[\sum_{i<j=1}^{N} -u(r_{ij,a,b})]$$

where

$$u(r_{ij,a,b}) = \frac{|r_i - r_j|^a}{b}$$

2. Take a linear combination of Slater orbitals with adjustable coefficients

References:

1. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E.Teller, J.Chem. Phys. 21, 1087 (1953)
2. W. L. McMillan, Phys. Rev. 138, A442 (1965)
3. D. Ceperley, G. V. Chester, and M. H. Kalos. Phys. Rev. B 16, 3081(1977)

Thank You