

Introduction to quantum Monte Carlo methods: Lectures I and II

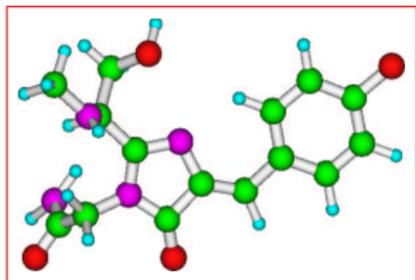
Claudia Filippi

Instituut-Lorentz, Universiteit Leiden, The Netherlands

Summer School: QMC from Minerals and Materials to Molecules
July 9-19, 2007, University of Illinois at Urbana-Champaign

A quick reminder: what is electronic structure theory?

A quantum mechanical and first-principle approach



→ Collection of ions + electrons



Only input: Z_α, N_α

Work in the Born-Oppenheimer approximation

Solve the Schrödinger equation for the electrons in the ionic field

$$\mathcal{H} = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_i v_{\text{ext}}(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

Solving the many-electron Schrödinger equation

$$\mathcal{H} = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_i v_{\text{ext}}(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

What do we want to compute?

Fermionic ground state and low-lying excited states

Evaluate expectation values $\frac{\langle \Psi_n | \mathcal{O} | \Psi_n \rangle}{\langle \Psi_n | \Psi_n \rangle}$

Where is the difficulty?

Electron-electron interaction \rightarrow non-separable

Is there an optimal theoretical approach?

- Density functional theory methods
Large systems but **approximate** exchange/correlation
- Quantum chemistry post-Hartree-Fock methods

↓
CI

↓
MCSCF

↓
CC ...

Very accurate on **small** systems

- **Quantum Monte Carlo techniques**

Fully-correlated calculations

Stochastic solution of Schrödinger equation

Most **accurate** benchmarks for **medium-large** systems

An analogy

Density functional theory



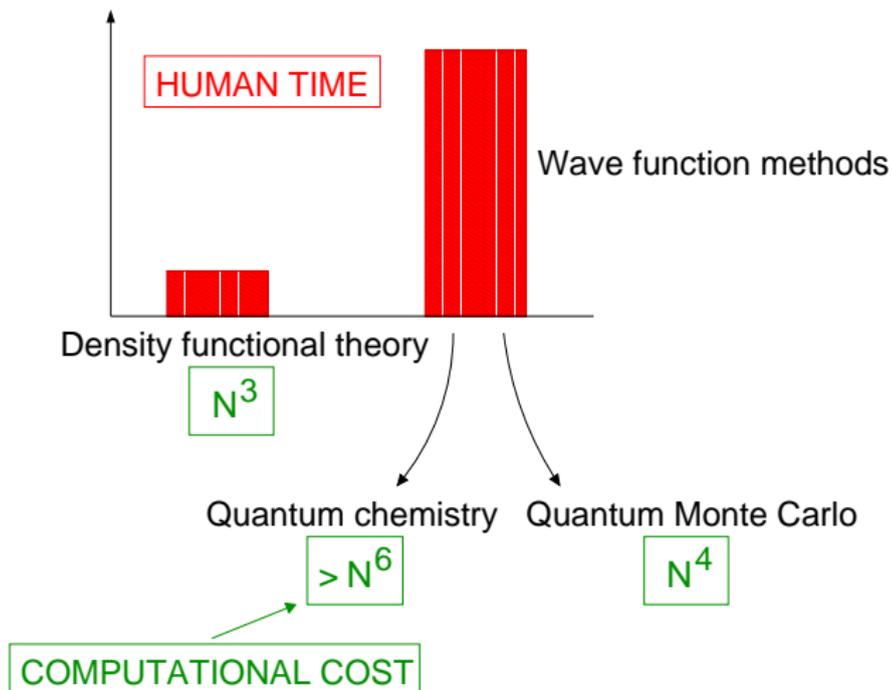
Quantum chemistry



Quantum Monte Carlo



If you can, use density functional theory!

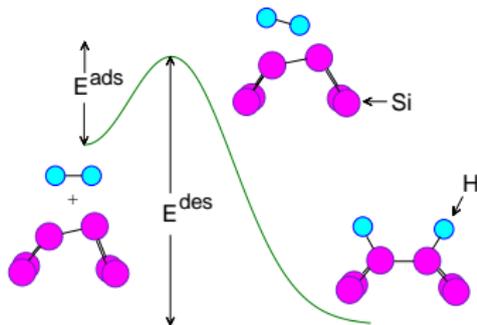


All is relative ...

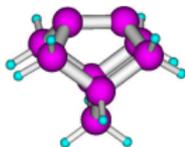
We think of density functional theory as cheap and painless!

... but density functional theory does not always work

A “classical” example: Adsorption/desorption of H_2 on Si(001)



For a small model cluster



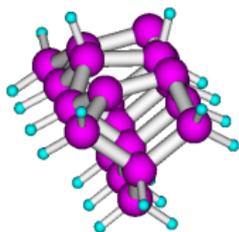
	E_a^{ads}	E_a^{des}	E_{rxn}
DFT	0.69	2.86	2.17
QMC	1.01(6)	3.65(6)	2.64(6)

eV

DFT error persists for larger models!

Favorable scaling of QMC with system size

QMC possible for realistic clusters with 2, 3, 4 ... surface dimers



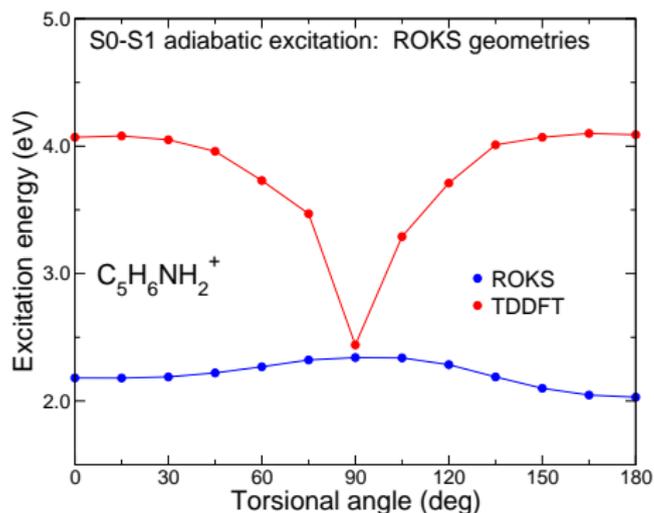
Accurate QMC calculations doable from small to large scales

Error of DFT is large \rightarrow 0.8 eV on desorption barrier !

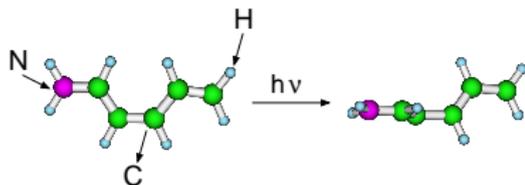
Healy, Filippi *et al.* PRL (2001); Filippi *et al.* PRL (2002)

What about DFT and excited states?

- Restricted open-shell Kohn-Sham method (DFT-ROKS)
- Time-dependent density functional theory (TDDFT)



Minimal model of rhodopsin



Comparison with QMC → Neither approach is reliable

When DFT has problems \rightarrow Wave function based methods

Wave function $\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N)$ where $\mathbf{x} = (\mathbf{r}, \sigma)$ and $\sigma = \pm 1$

How do we compute expectation values?

Many-body wave functions in traditional quantum chemistry

Interacting $\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) \leftrightarrow$ One-particle basis $\psi(\mathbf{x})$

Ψ expanded in determinants of single-particle orbitals $\psi(\mathbf{x})$

Single-particle orbitals expanded on Gaussian basis

\Rightarrow All integrals can be computed analytically

Many-body wave functions in traditional quantum chemistry

A jungle of acronyms: CI, CASSCF, MRCI, CASPT2 ...

Expansion in linear combination of determinants

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) \longrightarrow D_{\text{HF}} = \begin{vmatrix} \psi_1(\mathbf{x}_1) & \dots & \psi_1(\mathbf{x}_N) \\ \vdots & & \vdots \\ \psi_N(\mathbf{x}_1) & \dots & \psi_N(\mathbf{x}_N) \end{vmatrix}$$

$c_0 D_{\text{HF}} + c_1 D_1 + c_2 D_2 + \dots$ millions of determinants

$$\begin{vmatrix} \psi_1(\mathbf{x}_1) & \dots & \psi_1(\mathbf{x}_N) \\ \vdots & & \vdots \\ \psi_{N+1}(\mathbf{x}_1) & \dots & \psi_{N+1}(\mathbf{x}_N) \end{vmatrix}$$

Integrals computed analytically but slowly converging expansion

Can we use a more compact Ψ ?

We want to construct an accurate and more compact Ψ

Explicit dependence on the inter-electronic distances r_{ij}

How do we compute expectation values if no single-electron basis?

A different way of writing the expectation values

Consider the expectation value of the Hamiltonian on Ψ

$$\begin{aligned} E_V &= \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\int d\mathbf{R} \Psi^*(\mathbf{R}) \mathcal{H} \Psi(\mathbf{R})}{\int d\mathbf{R} \Psi^*(\mathbf{R}) \Psi(\mathbf{R})} \geq E_0 \\ &= \int d\mathbf{R} \frac{\mathcal{H} \Psi(\mathbf{R})}{\Psi(\mathbf{R})} \frac{|\Psi(\mathbf{R})|^2}{\int d\mathbf{R} |\Psi(\mathbf{R})|^2} \\ &= \int d\mathbf{R} E_L(\mathbf{R}) \rho(\mathbf{R}) = \langle E_L(\mathbf{R}) \rangle_\rho \end{aligned}$$

ρ is a distribution function and $E_L(\mathbf{R}) = \frac{\mathcal{H} \Psi(\mathbf{R})}{\Psi(\mathbf{R})}$ the local energy

Variational Monte Carlo: a random walk of the electrons

Use Monte Carlo integration to compute expectation values

- ▷ Sample \mathbf{R} from $\rho(\mathbf{R})$ using Metropolis algorithm
- ▷ Average local energy $E_L(\mathbf{R}) = \frac{\mathcal{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$ to obtain E_V as

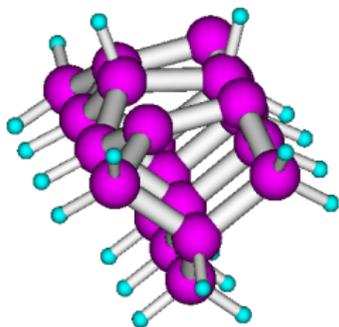
$$E_V = \langle E_L(\mathbf{R}) \rangle_\rho \approx \frac{1}{M} \sum_{i=1}^M E_L(\mathbf{R}_i)$$



Random walk in $3N$ dimensions, $\mathbf{R} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$

Just a **trick** to evaluate integrals in many dimensions

Is it really “just” a trick?



$\text{Si}_{21}\text{H}_{22}$

Number of electrons $4 \times 21 + 22 = 106$

Number of dimensions $3 \times 106 = 318$

Integral on a grid with 10 points/dimension $\rightarrow 10^{318}$ points!

MC is a powerful trick \Rightarrow Freedom in form of the wave function Ψ

Are there any conditions on many-body Ψ to be used in VMC?

Within VMC, we can use any “computable” wave function if

- ▷ Continuous, normalizable, proper symmetry
- ▷ **Finite variance**

$$\sigma^2 = \frac{\langle \Psi | (\mathcal{H} - E_V)^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle (E_L(\mathbf{R}) - E_V)^2 \rangle_\rho$$

since the Monte Carlo error goes as $\text{err}(E_V) \sim \frac{\sigma}{\sqrt{M}}$

Zero variance principle: if $\Psi \rightarrow \Psi_0$, $E_L(\mathbf{R})$ does not fluctuate

Variational Monte Carlo and the generalized Metropolis algorithm

How do we sample distribution function $\rho(\mathbf{R}) = \frac{|\Psi(\mathbf{R})|^2}{\int d\mathbf{R} |\Psi(\mathbf{R})|^2}$?

Aim → Obtain a set of $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_M\}$ distributed as $\rho(\mathbf{R})$

Let us generate a **Markov chain**

- ▷ Start from arbitrary initial state \mathbf{R}_i
- ▷ Use stochastic transition matrix $M(\mathbf{R}_f|\mathbf{R}_i)$

$$M(\mathbf{R}_f|\mathbf{R}_i) \geq 0 \quad \sum_{\mathbf{R}_f} M(\mathbf{R}_f|\mathbf{R}_i) = 1.$$

as probability of making transition $\mathbf{R}_i \rightarrow \mathbf{R}_f$

- ▷ Evolve the system by repeated application of M

Stationarity condition

To sample ρ , use M which satisfies **stationarity condition**:

$$\sum_i M(\mathbf{R}_f | \mathbf{R}_i) \rho(\mathbf{R}_i) = \rho(\mathbf{R}_f) \quad \forall \mathbf{R}_f$$

▷ Stationarity condition

⇒ If we start with ρ , we continue to sample ρ

▷ Stationarity condition + stochastic property of M + ergodicity

⇒ Any initial distribution will evolve to ρ

More stringent condition

In practice, we impose detailed balance condition

$$M(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i) = M(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f)$$

Stationarity condition can be obtained by summing over \mathbf{R}_i

$$\sum_i M(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i) = \sum_i M(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f) = \rho(\mathbf{R}_f)$$

Detailed balance is a sufficient but not necessary condition

How do we construct the transition matrix M in practice?

Write transition matrix M as proposal $T \times$ acceptance A

$$M(\mathbf{R}_f|\mathbf{R}_i) = A(\mathbf{R}_f|\mathbf{R}_i) T(\mathbf{R}_f|\mathbf{R}_i)$$

M and T are stochastic matrices but A is not

Rewriting detailed balance condition

$$M(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i) = M(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f)$$

$$A(\mathbf{R}_f|\mathbf{R}_i) T(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i) = A(\mathbf{R}_i|\mathbf{R}_f) T(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f)$$

$$\text{or } \frac{A(\mathbf{R}_f|\mathbf{R}_i)}{A(\mathbf{R}_i|\mathbf{R}_f)} = \frac{T(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i)}$$

Choice of acceptance matrix A

(1)

Detailed balance condition is

$$\frac{A(\mathbf{R}_f|\mathbf{R}_i)}{A(\mathbf{R}_i|\mathbf{R}_f)} = \frac{T(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i)}$$

For a given choice of T , infinite choices of A satisfy this equation

Any function $A(\mathbf{R}_f|\mathbf{R}_i) = F\left(\frac{T(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i)}\right)$ with

$$\frac{F(x)}{F(1/x)} = x$$

will do the job!

Choice of acceptance matrix A

(2)

Original choice by Metropolis *et al.* maximizes the acceptance

$$A(\mathbf{R}_f|\mathbf{R}_i) = \min \left\{ 1, \frac{T(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i)} \right\}$$

Note: $\rho(\mathbf{R})$ does not have to be normalized

Original Metropolis method

$$\text{Symmetric } T(\mathbf{R}_f|\mathbf{R}_i) = 1/\Delta^{3N} \Rightarrow A(\mathbf{R}_f|\mathbf{R}_i) = \min \left\{ 1, \frac{\rho(\mathbf{R}_f)}{\rho(\mathbf{R}_i)} \right\}$$

Original Metropolis method

Aim → Obtain a set of $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_M\}$ distributed as $\rho(\mathbf{R})$

Operationally, simple algorithm:

1. Pick a starting \mathbf{R} and evaluate $\rho(\mathbf{R})$
2. Choose \mathbf{R}' at random in a box centered at \mathbf{R}
3. If $\rho(\mathbf{R}') \geq \rho(\mathbf{R})$, move accepted → put \mathbf{R}' in the set
4. If $\rho(\mathbf{R}') < \rho(\mathbf{R})$, move accepted with $p = \frac{\rho(\mathbf{R}')}{\rho(\mathbf{R})}$

To do this, pick a random number $\chi \in [0, 1]$:

- a) If $\chi < p$, move accepted → put \mathbf{R}' in the set
- b) If $\chi > p$, move rejected → put another entry of \mathbf{R} in the set

Choice of proposal matrix T

(1)

Is the original choice of T by Metropolis the best possible choice ?

Walk sequentially correlated $\Rightarrow M_{\text{eff}} < M$ independent observations

$$M_{\text{eff}} = \frac{M}{T_{\text{corr}}} \text{ with } T_{\text{corr}} \text{ autocorrelation time of desired observable}$$

Aim is to achieve fast evolution of the system and reduce T_{corr}

Use **freedom** in choice of T to have high acceptance

$$\frac{T(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i)} \approx 1 \Rightarrow A(\mathbf{R}_f|\mathbf{R}_i) \approx 1$$

and small T_{corr} of desired observable

Limitation: we need to be able to sample T directly!

Choice of proposal matrix T

(2)

If Δ is the linear dimension of domain around \mathbf{R}_i

$$\frac{A(\mathbf{R}_f|\mathbf{R}_i)}{A(\mathbf{R}_i|\mathbf{R}_f)} = \frac{T(\mathbf{R}_i|\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i)} \frac{\rho(\mathbf{R}_f)}{\rho(\mathbf{R}_i)} \approx 1 - \mathcal{O}(\Delta^m)$$

- ▷ T symmetric as in original Metropolis algorithm gives $m = 1$
- ▷ A choice motivated by diffusion Monte Carlo with $m = 2$ is

$$T(\mathbf{R}_f|\mathbf{R}_i) = \mathcal{N} \exp \left[-\frac{(\mathbf{R}_f - \mathbf{R}_i - \mathbf{V}(\mathbf{R}_i)\tau)^2}{2\tau} \right] \quad \text{with} \quad \mathbf{V}(\mathbf{R}_i) = \frac{\nabla \Psi(\mathbf{R}_i)}{\Psi(\mathbf{R}_i)}$$

- ▷ Other (better) choices of T are possible

Acceptance and T_{corr} for the total energy E_V

Example: All-electron Be atom with simple wave function

Simple Metropolis

Δ	T_{corr}	\bar{A}
1.00	41	0.17
0.75	21	0.28
0.50	17	0.46
0.20	45	0.75

Drift-diffusion transition

τ	T_{corr}	\bar{A}
0.100	13	0.42
0.050	7	0.66
0.020	8	0.87
0.010	14	0.94

Generalized Metropolis algorithm

1. Choose distribution $\rho(\mathbf{R})$ and transition probability $T(\mathbf{R}_f|\mathbf{R}_i)$
2. Initialize the configuration \mathbf{R}_i
3. Advance the configuration from \mathbf{R}_i to \mathbf{R}'
 - a) Sample \mathbf{R}' from $T(\mathbf{R}'|\mathbf{R}_i)$.
 - b) Calculate the ratio $q = \frac{T(\mathbf{R}_i|\mathbf{R}') \rho(\mathbf{R}')}{T(\mathbf{R}'|\mathbf{R}_i) \rho(\mathbf{R}_i)}$
 - c) Accept or reject with probability q

Pick a uniformly distributed random number $\chi \in [0, 1]$

if $\chi < q$, move accepted \rightarrow set $\mathbf{R}_f = \mathbf{R}'$

if $\chi > q$, move rejected \rightarrow set $\mathbf{R}_f = \mathbf{R}$
4. Throw away first κ configurations of equilibration time
5. Collect the averages and block them to obtain the error bars

Improvements on simple and drift-diffusion algorithms

- ▷ For all-electron and pseudopotential systems:

Move one electron at the time → Decorrelate faster

Does total matrix $M = \prod_{i=1}^N M_i$ satisfy stationarity condition?

Yes if matrices M_1, M_2, \dots, M_n satisfy stationarity condition

- ▷ For all-electron systems (Umrigar PRL 1993)

- Core electrons set the length scales

- T must distinguish between core and valence electrons

- Do not use cartesian coordinates

- Derivative discontinuity of Ψ at nuclei

Better algorithms can achieve $T_{\text{corr}} = 1 - 2$

We compute the expectation value of the Hamiltonian \mathcal{H} as

$$\begin{aligned} E_V &= \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \\ &= \int d\mathbf{R} \frac{\mathcal{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \frac{|\Psi(\mathbf{R})|^2}{\int d\mathbf{R} |\Psi(\mathbf{R})|^2} \\ &= \int d\mathbf{R} E_L(\mathbf{R}) \rho(\mathbf{R}) \\ &= \langle E_L(\mathbf{R}) \rangle_\rho \approx \frac{1}{M} \sum_{i=1}^M E_L(\mathbf{R}_i) \end{aligned}$$

Note: a) Metropolis method: ρ does not have to be normalized

→ For complex Ψ we do not know the normalization!

b) If $\Psi \rightarrow$ eigenfunction, $E_L(\mathbf{R})$ does not fluctuate

Expectation values in variational Monte Carlo

(2)

The energy is computed by averaging the local energy

$$E_V = \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle E_L(\mathbf{R}) \rangle_\rho$$

The variance of the local energy is given by

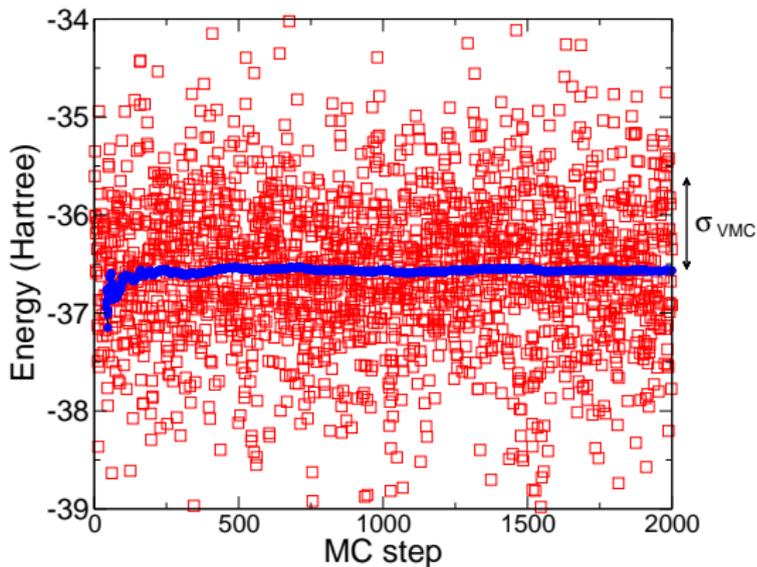
$$\sigma^2 = \frac{\langle \Psi | (\mathcal{H} - E_V)^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle (E_L(\mathbf{R}) - E_V)^2 \rangle_\rho$$

The statistical Monte Carlo error goes as $\text{err}(E_V) \sim \frac{\sigma}{\sqrt{M}}$

Note: For other operators, substitute \mathcal{H} with \mathcal{X}

Typical VMC run

Example: Local energy and average energy of acetone (C_3H_6O)



$$E_{VMC} = \langle E_L(\mathbf{R}) \rangle_{\rho} = -36.542 \pm 0.001 \text{ Hartree (40} \times \text{20000 steps)}$$

$$\sigma_{VMC} = \langle (E_L(\mathbf{R}) - E_{VMC})^2 \rangle_{\rho} = 0.90 \text{ Hartree}$$

Variational Monte Carlo \rightarrow Freedom in choice of Ψ

Monte Carlo integration allows the use of complex and accurate Ψ

\Rightarrow More compact representation of Ψ than in quantum chemistry

\Rightarrow Beyond $c_0 D_{\text{HF}} + c_1 D_1 + c_2 D_2 + \dots$ millions of determinants

Jastrow-Slater wave function

(1)

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathcal{J}(\mathbf{r}_1, \dots, \mathbf{r}_N) \sum_k d_k D_k^\uparrow(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow}) D_k^\downarrow(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N)$$

\mathcal{J} \longrightarrow Jastrow correlation factor

- Positive function of inter-particle distances
- Explicit dependence on electron-electron distances r_{ij}
- Takes care of divergences in potential

Jastrow-Slater wave function

(2)

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathcal{J}(\mathbf{r}_1, \dots, \mathbf{r}_N) \sum_k d_k D_k^\uparrow(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow}) D_k^\downarrow(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N)$$

$\sum_k d_k D_k^\uparrow D_k^\downarrow$ \longrightarrow Determinants of single-particle orbitals

- **Few** and not millions of determinants as in quantum chemistry
- Slater basis to expand orbitals in all-electron calculations

$$\phi(\mathbf{r}) = \sum_{\alpha}^{\text{Nuclei}} \sum_{k_{\alpha}} c_{k_{\alpha}} r_{\alpha}^{n_{k_{\alpha}}-1} \exp(-\zeta_{k_{\alpha}} r_{\alpha}) Y_{l_{k_{\alpha}} m_{k_{\alpha}}}(\hat{\mathbf{r}}_{\alpha})$$

Gaussian atomic basis used in pseudopotential calculations

- Slater component determines the nodal surface

What is strange with the Jastrow-Slater wave function?

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathcal{J}(\mathbf{r}_1, \dots, \mathbf{r}_N) \sum_k d_k D_k^\uparrow(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow}) D_k^\downarrow(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N)$$

▷ Why is Ψ not depending on the spin variables σ ?

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \Psi(\mathbf{r}_1, \sigma_1, \dots, \mathbf{r}_N, \sigma_N) \quad \text{with} \quad \sigma_i = \pm 1$$

▷ Why is Ψ not totally antisymmetric?

Why can we factorize $D_k^\uparrow D_k^\downarrow$?

Consider N electrons with $N = N_\uparrow + N_\downarrow$ and $S_z = (N_\uparrow - N_\downarrow)/2$

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \Psi(\mathbf{r}_1, \sigma_1, \dots, \mathbf{r}_N, \sigma_N) \quad \text{with} \quad \sigma_i = \pm 1$$

Define a spin function ζ_1

$$\zeta_1(\sigma_1, \dots, \sigma_N) = \chi_\uparrow(\sigma_1) \dots \chi_\uparrow(\sigma_{N_\uparrow}) \chi_\downarrow(\sigma_{N_\uparrow+1}) \dots \chi_\downarrow(\sigma_N)$$

Generate $K = N!/N_\uparrow!N_\downarrow!$ functions ζ_i by permuting indices in ζ_1

The functions ζ_i form a complete, orthonormal set in spin space

$$\sum_{\sigma_1 \dots \sigma_N} \zeta_i(\sigma_1, \dots, \sigma_N) \zeta_j(\sigma_1, \dots, \sigma_N) = \delta_{ij}$$

Wave function with space and spin variables

Expand the wave function Ψ in terms of its spin components

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \sum_{i=1}^K F_i(\mathbf{r}_1, \dots, \mathbf{r}_N) \zeta_i(\sigma_1, \dots, \sigma_N)$$

Ψ is totally antisymmetric \Rightarrow

- ▷ $F_i = -F_j$ for interchange of like-spin
- ▷ $F_i = \pm$ permutation of F_1

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \mathcal{A} \{ F_1(\mathbf{r}_1, \dots, \mathbf{r}_N) \zeta_1(\sigma_1, \dots, \sigma_N) \}$$

Can we get rid of spin variables? Spin-assigned wave functions

Note that if \mathcal{O} is a spin-independent operator

$$\langle \Psi | \mathcal{O} | \Psi \rangle = \langle F_1 | \mathcal{O} | F_1 \rangle$$

since the functions ζ_j form an orthonormal set

More convenient to use F_1 instead of full wave function Ψ

To obtain F_1 , assign the spin-variables of particles:

Particle	1	2	...	N_{\uparrow}	$N_{\uparrow+1}$...	N
σ	1	1	...	1	-1	...	-1

$$F_1(\mathbf{r}_1, \dots, \mathbf{r}_N) = \Psi(\mathbf{r}_1, 1, \dots, \mathbf{r}_{N_{\uparrow}}, 1, \mathbf{r}_{N_{\uparrow+1}}, -1, \dots, \mathbf{r}_N, -1)$$

Spin assignment: a simple wave function for the Be atom

(1)

Be atom, $1s^2 2s^2 \Rightarrow N_{\uparrow} = N_{\downarrow} = 2, S_z = 0$

Determinant of spin-orbitals $\phi_{1s} \chi_{\uparrow}, \phi_{2s} \chi_{\uparrow}, \phi_{1s} \chi_{\downarrow}, \phi_{2s} \chi_{\downarrow}$

$$D = \frac{1}{\sqrt{4!}} \begin{vmatrix} \phi_{1s}(\mathbf{r}_1)\chi_{\uparrow}(\sigma_1) & \dots & \phi_{1s}(\mathbf{r}_4)\chi_{\uparrow}(\sigma_4) \\ \phi_{2s}(\mathbf{r}_1)\chi_{\uparrow}(\sigma_1) & \dots & \phi_{2s}(\mathbf{r}_4)\chi_{\uparrow}(\sigma_4) \\ \phi_{1s}(\mathbf{r}_1)\chi_{\downarrow}(\sigma_1) & \dots & \phi_{1s}(\mathbf{r}_4)\chi_{\downarrow}(\sigma_4) \\ \phi_{2s}(\mathbf{r}_1)\chi_{\downarrow}(\sigma_1) & \dots & \phi_{2s}(\mathbf{r}_4)\chi_{\downarrow}(\sigma_4) \end{vmatrix}$$

Spin-assigned $F_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = D(\mathbf{r}_1, +1, \mathbf{r}_2, +1, \mathbf{r}_3, -1, \mathbf{r}_4, -1)$

$$F_1 = \frac{1}{\sqrt{4!}} \begin{vmatrix} \phi_{1s}(\mathbf{r}_1) & \phi_{1s}(\mathbf{r}_2) & 0 & 0 \\ \phi_{2s}(\mathbf{r}_1) & \phi_{2s}(\mathbf{r}_2) & 0 & 0 \\ 0 & 0 & \phi_{1s}(\mathbf{r}_3) & \phi_{1s}(\mathbf{r}_4) \\ 0 & 0 & \phi_{2s}(\mathbf{r}_3) & \phi_{2s}(\mathbf{r}_4) \end{vmatrix}$$

Spin assignment: a simple wave function for the Be atom

(2)

Be atom, $1s^2 2s^2 \Rightarrow N_{\uparrow} = N_{\downarrow} = 2, S_z = 0$

$$F_1 = \frac{1}{\sqrt{4!}} \begin{vmatrix} \phi_{1s}(\mathbf{r}_1) & \phi_{1s}(\mathbf{r}_2) & 0 & 0 \\ \phi_{2s}(\mathbf{r}_1) & \phi_{2s}(\mathbf{r}_2) & 0 & 0 \\ 0 & 0 & \phi_{1s}(\mathbf{r}_3) & \phi_{1s}(\mathbf{r}_4) \\ 0 & 0 & \phi_{2s}(\mathbf{r}_3) & \phi_{2s}(\mathbf{r}_4) \end{vmatrix}$$
$$= \frac{1}{\sqrt{4!}} \begin{vmatrix} \phi_{1s}(\mathbf{r}_1) & \phi_{1s}(\mathbf{r}_2) \\ \phi_{2s}(\mathbf{r}_1) & \phi_{2s}(\mathbf{r}_2) \end{vmatrix} \times \begin{vmatrix} \phi_{1s}(\mathbf{r}_3) & \phi_{1s}(\mathbf{r}_4) \\ \phi_{2s}(\mathbf{r}_3) & \phi_{2s}(\mathbf{r}_4) \end{vmatrix}$$

$$D(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) \rightarrow D^{\uparrow}(\mathbf{r}_1, \mathbf{r}_2) \times D^{\downarrow}(\mathbf{r}_3, \mathbf{r}_4)$$

Jastrow-Slater spin-assigned wave function

To obtain spin-assigned Jastrow-Slater wave functions, impose

Particle	1	2	...	N_{\uparrow}	$N_{\uparrow+1}$...	N
σ	1	1	...	1	-1	...	-1

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = F_1(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

$$= \mathcal{J}(\mathbf{r}_1, \dots, \mathbf{r}_N) \sum_k d_k D_k^{\uparrow}(\mathbf{r}_1, \dots, \mathbf{r}_{N_{\uparrow}}) D_k^{\downarrow}(\mathbf{r}_{N_{\uparrow}+1}, \dots, \mathbf{r}_N)$$

How do we impose space and spin symmetry on Jastrow-Slater Ψ ?

$\sum_k d_k D_k$ is constructed to have the proper space/spin symmetry

▷ Spatial symmetry

Often, $\mathcal{J} = \mathcal{J}(\{r_{ij}\}, \{r_{i\alpha}\})$ with i, j electrons and α nuclei

$\Rightarrow \mathcal{J}$ invariant under rotations, no effect on spatial symmetry of Ψ

▷ Spin symmetry

If \mathcal{J} is symmetric

\rightarrow for interchange of like-spin electrons $\Rightarrow \Psi$ eigenstate of S_z

\rightarrow for interchange of spatial variables $\Rightarrow \Psi$ eigenstate of S^2

Jastrow factor and divergences in the potential

At interparticle coalescence points, the potential diverges as

$$-\frac{Z}{r_{i\alpha}} \quad \text{for the electron-nucleus potential}$$

$$\frac{1}{r_{ij}} \quad \text{for the electron-electron potential}$$

Local energy $\frac{\mathcal{H}\Psi}{\Psi} = -\frac{1}{2} \sum_i \frac{\nabla_i^2 \Psi}{\Psi} + \mathcal{V}$ must be **finite**

\Rightarrow Kinetic energy must have opposite divergence to the potential \mathcal{V}

Divergence in potential and behavior of the local energy

Consider two particles of masses m_i, m_j and charges q_i, q_j

Assume $r_{ij} \rightarrow 0$ while all other particles are well separated

Keep only diverging terms in $\frac{\mathcal{H}\Psi}{\Psi}$ and go to relative coordinates close to $\mathbf{r} = \mathbf{r}_{ij} = 0$

$$\begin{aligned} -\frac{1}{2\mu_{ij}} \frac{\nabla^2 \Psi}{\Psi} + \mathcal{V}(r) &\sim -\frac{1}{2\mu_{ij}} \frac{\Psi''}{\Psi} - \frac{1}{\mu_{ij}} \frac{1}{r} \frac{\Psi'}{\Psi} + \mathcal{V}(r) \\ &\sim \boxed{-\frac{1}{\mu_{ij}} \frac{1}{r} \frac{\Psi'}{\Psi} + \mathcal{V}(r)} \end{aligned}$$

where $\mu_{ij} = m_i m_j / (m_i + m_j)$

Divergence in potential and cusp conditions

Diverging terms in the local energy

$$-\frac{1}{\mu_{ij}} \frac{1}{r} \frac{\Psi'}{\Psi} + \mathcal{V}(r) = -\frac{1}{\mu_{ij}} \frac{1}{r} \frac{\Psi'}{\Psi} + \frac{q_i q_j}{r} = \text{finite}$$

$\Rightarrow \Psi$ must satisfy Kato's cusp conditions:

$$\left. \frac{\partial \hat{\Psi}}{\partial r_{ij}} \right|_{r_{ij}=0} = \mu_{ij} q_i q_j \Psi(r_{ij} = 0)$$

where $\hat{\Psi}$ is a spherical average

Note: We assumed $\Psi(r_{ij} = 0) \neq 0$

Cusp conditions: example

The condition for the local energy to be finite at $r = 0$ is

$$\frac{\psi'}{\psi} = \mu_{ij} q_i q_j$$

- Electron-nucleus: $\mu = 1, q_i = 1, q_j = -Z \Rightarrow$

$$\left. \frac{\psi'}{\psi} \right|_{r=0} = -Z$$

- Electron-electron: $\mu = \frac{1}{2}, q_i = 1, q_j = 1 \Rightarrow$

$$\left. \frac{\psi'}{\psi} \right|_{r=0} = 1/2$$

Generalized cusp conditions

What about two electrons in a triplet state?

Or more generally two like-spin electrons (D^\uparrow or $D^\downarrow \rightarrow 0$)?

$$\Psi(r = r_{ij} = 0) = 0 \quad ?!?$$

Near $\mathbf{r} = \mathbf{r}_{ij} = 0$,

$$\Psi = \sum_{l=l_0}^{\infty} \sum_{m=-l}^l f_{lm}(r) r^l Y_{lm}(\theta, \phi)$$

Local energy is finite if

$$f_{lm}(r) = f_{lm}^{(0)} \left[1 + \frac{\gamma}{(l+1)} r + O(r^2) \right]$$

where $\gamma = q_i q_j \mu_{ij}$

R. T. Pack and W. Byers Brown, JCP **45**, 556 (1966)

Generalized cusp conditions: like-spin electrons

- Electron-electron singlet: $l_0 = 0 \Rightarrow \Psi \sim \left(1 + \frac{1}{2}r\right) \Rightarrow \frac{\Psi'}{\Psi} = \frac{1}{2}$
- Electron-electron triplet: $l_0 = 1 \Rightarrow \Psi \sim \left(1 + \frac{1}{4}r\right) r$

$\sigma = +1$ for first N_{\uparrow} electrons, $\sigma = -1$ for the others

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathcal{J}(\mathbf{r}_1, \dots, \mathbf{r}_{N_{\uparrow}}) \sum_k d_k D_k^{\uparrow}(\mathbf{r}_1, \dots, \mathbf{r}_{N_{\uparrow}}) D_k^{\downarrow}(\mathbf{r}_{N_{\uparrow}+1}, \dots, \mathbf{r}_N)$$

▷ Anti-parallel spins: $r_{ij} \rightarrow 0$ for $i \leq N_{\uparrow}, j \geq N_{\uparrow} + 1$

Usually, determinantal part $\neq 0$

$$\Rightarrow \mathcal{J}(r_{ij}) \sim \left(1 + \frac{1}{2} r_{ij}\right) \Leftrightarrow \boxed{\left. \frac{\mathcal{J}'}{\mathcal{J}} \right|_{r_{ij}=0} = \frac{1}{2}}$$

▷ Parallel spins: $r_{ij} \rightarrow 0$ for $i, j \leq N_{\uparrow}$ or $i, j \geq N_{\uparrow} + 1$

Determinantal part $\rightarrow 0$

$$\Rightarrow \mathcal{J}(r_{ij}) \sim \left(1 + \frac{1}{4} r_{ij}\right) \Leftrightarrow \boxed{\left. \frac{\mathcal{J}'}{\mathcal{J}} \right|_{r_{ij}=0} = \frac{1}{4}}$$

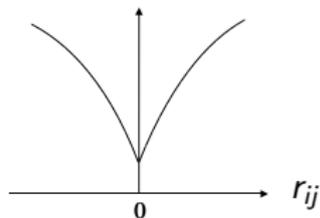
- ▷ Electron-electron cusps imposed through the Jastrow factor

Example: Simple Jastrow factor

$$\mathcal{J}(r_{ij}) = \prod_{i < j} \exp \left\{ b_0 \frac{r_{ij}}{1 + b r_{ij}} \right\}$$

with $b_0^{\uparrow\downarrow} = \frac{1}{2}$ or $b_0^{\uparrow\uparrow} = b_0^{\downarrow\downarrow} = \frac{1}{4}$

Imposes cusp conditions
+
keeps electrons apart



Cusp conditions and QMC wave functions

(3)

- ▶ Electron-nucleus cusps imposed through the determinantal part

Assume that the nucleus is at the origin and $\Psi(r_i = 0) \neq 0$

If each orbital satisfies the cusp conditions

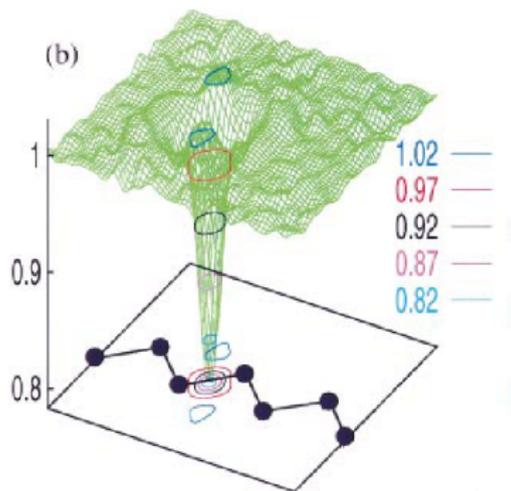
$$\begin{aligned} \left. \frac{\partial \hat{\phi}_j}{\partial r} \right|_{r=0} &= -Z \hat{\phi}_j(r=0) \\ \Rightarrow \left. \frac{\partial \sum_k d_k \hat{D}_k}{\partial r} \right|_{r=0} &= -Z \sum_k d_k \hat{D}_k(r=0) \end{aligned}$$

Note: Slater basis best suited for all-electron systems

No electron-nucleus cusp with pseudopotential

The effect of the Jastrow factor

Pair correlation function for $\uparrow\downarrow$ electrons in the (110) plane of Si
 $g_{\uparrow\downarrow}(\mathbf{r}, \mathbf{r}')$ with one electron is at the bond center



Hood *et al.* Phys. Rev. Lett. **78**, 3350 (1997)

Simple wave function for the Be atom

Be atom, $1s^2 2s^2 \Rightarrow N_{\uparrow} = N_{\downarrow} = 2, S_z = 0$

Spin-assigned $\Psi(\mathbf{r}_1, +1, \mathbf{r}_2, +1, \mathbf{r}_3, -1, \mathbf{r}_4, -1) = \mathcal{J} D$

▷ Factorized determinant

$$D = D^{\uparrow} \times D^{\downarrow} = \begin{vmatrix} \phi_{1s}(\mathbf{r}_1) & \phi_{1s}(\mathbf{r}_2) \\ \phi_{2s}(\mathbf{r}_1) & \phi_{2s}(\mathbf{r}_2) \end{vmatrix} \times \begin{vmatrix} \phi_{1s}(\mathbf{r}_3) & \phi_{1s}(\mathbf{r}_4) \\ \phi_{2s}(\mathbf{r}_3) & \phi_{2s}(\mathbf{r}_4) \end{vmatrix}$$

▷ Simple Jastrow factor

$$\mathcal{J} = \prod_{ij=13,14,23,24} \exp \left\{ \frac{1}{2} \frac{r_{ij}}{1 + b r_{ij}} \right\} \times \prod_{ij=12,34} \exp \left\{ \frac{1}{4} \frac{r_{ij}}{1 + b r_{ij}} \right\}$$

Jastrow factor for atoms and molecules: Beyond the simple form

Boys and Handy's form

$$\mathcal{J}(r_i, r_j, r_{ij}) = \prod_{\alpha, i < j} \exp \left\{ \sum c_{mnk}^{\alpha} (\bar{r}_{i\alpha}^m \bar{r}_{j\alpha}^n + \bar{r}_{i\alpha}^n \bar{r}_{j\alpha}^m) \bar{r}_{ij}^k \right\}$$

with $\bar{r}_{i\alpha} = \frac{a r_{i\alpha}}{1 + a r_{i\alpha}}$ and $\bar{r}_{ij} = \frac{d r_{ij}}{1 + d r_{ij}}$

Cusp conditions imposed by requiring:

For electron-electron cusps: $m = n = 0$ if $k = 1$

For electron-nucleus cusps: No $n = 1$ or $m = 1$, D satisfies cusps

More general form: Lift constraints and allow all values of n, m, k

Impose the cusp conditions via linear dependencies among c_{mnk}^{α}

Other scaling functions are possible, e.g. $(1 - e^{-ar})/a$

More general Jastrow form with e-n, e-e and e-e-n terms

$$\prod_{\alpha,i} \exp \{A(r_{i\alpha})\} \prod_{i<j} \exp \{B(r_{ij})\} \prod_{\alpha,i<j} \exp \{C(r_{i\alpha}, r_{j\alpha}, r_{ij})\}$$

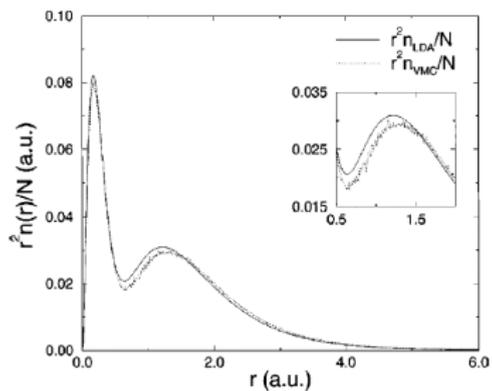
- ▷ Polynomials of scaled variables, e.g. $\bar{r} = r/(1 + ar)$
- ▷ $\mathcal{J} > 0$ and becomes constant for large r_i , r_j and r_{ij}
- ▷ Electron-electron terms B
 - Imposes the cusp conditions and keeps electrons apart
 - More general than simple $\mathcal{J}(r_{ij})$ gives small improvements
- ▷ Electron-nucleus terms A

Should be included if determinantal part (DFT or HF) is not reoptimized: e-e terms alter the single-particle density

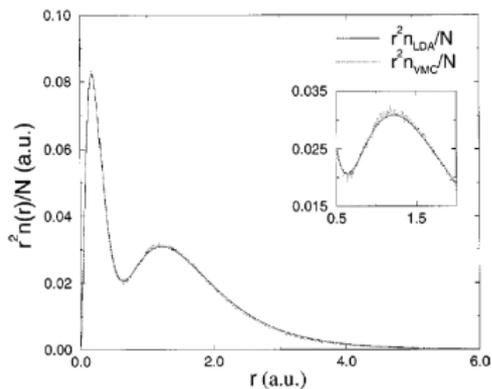
Role of the electron-nucleus terms

Example: Density of all-electron Carbon atom

DFT determinant + e-e \mathcal{J}



+ e-n \mathcal{J}



Foulkes *et al.* Rev. Mod. Phys. **73**, 33 (2001)

Some comments on Jastrow factor

(2)

▷ Electron-electron-nucleus terms C

If the order of the polynomial in the e-e-n terms is infinite, Ψ can exactly describe a two-electron atom or ion in an S state

For these systems, a 5th-order polynomial recovers more than 99.99% of the correlation energy, $E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}$

▷ Is this Jastrow factor adequate for multi-electron systems?

The e-e-n terms are the most important: due to the exclusion principle, it is rare for 3 or more electrons to be close, since at least 2 electrons must necessarily have the same spin

Jastrow factor with e-e, e-e-n and e-e-e-n terms

	\mathcal{J}	E_{VMC}	$E_{\text{VMC}}^{\text{corr}} (\%)$	σ_{VMC}	
Li	E_{HF}		-7.43273	0	
		e-e	-7.47427(4)	91.6	0.240
		+ e-e-n	-7.47788(1)	99.6	0.037
		+ e-e-e-n	-7.47797(1)	99.8	0.028
	E_{exact}		-7.47806	100	0
Ne	E_{HF}		-128.5471	0	
		e-e	-128.713(2)	42.5	1.90
		+ e-e-n	-128.9008(1)	90.6	0.90
		+ e-e-e-n	-128.9029(3)	91.1	0.88
	E_{exact}		-128.9376	100	0

Huang, Umrigar, Nightingale, J. Chem. Phys. **107**, 3007 (1997)

Dynamic and static correlation

$\Psi = \text{Jastrow} \times \text{Determinants} \rightarrow \text{Two types of correlation}$

▷ Dynamic correlation

Described by Jastrow factor

Due to inter-electron repulsion

Always present

▷ Static correlation

Described by a linear combination of determinants

Due to near-degeneracy of occupied and unoccupied orbitals

Not always present

Static correlation

(1)

Example: Be atom and $2s$ - $2p$ near-degeneracy

HF ground state configuration

$$1s^2 2s^2$$

Additional important configuration

$$1s^2 2p^2$$

Ground state has 1S symmetry \Rightarrow 4 determinants

$$D = (1s^\uparrow, 2s^\uparrow, 1s^\downarrow, 2s^\downarrow) + c \left[(1s^\uparrow, 2p_x^\uparrow, 1s^\downarrow, 2p_x^\downarrow) \right. \\ \left. + (1s^\uparrow, 2p_y^\uparrow, 1s^\downarrow, 2p_y^\downarrow) \right. \\ \left. + (1s^\uparrow, 2p_z^\uparrow, 1s^\downarrow, 2p_z^\downarrow) \right]$$

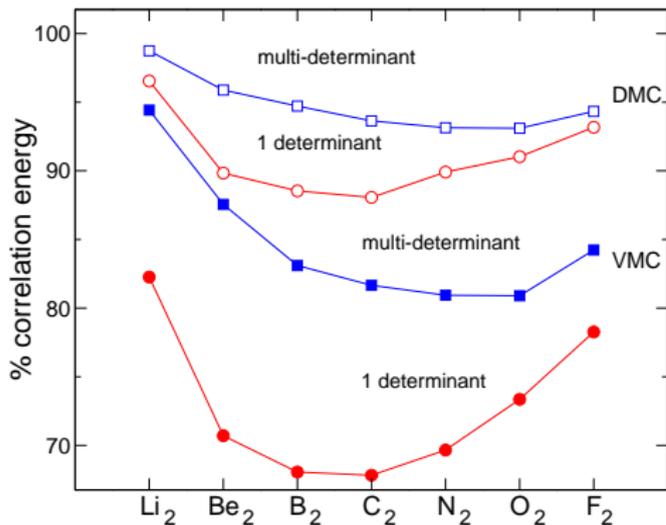
$$1s^2 2s^2 \quad \times \mathcal{J}(r_{ij}) \quad \rightarrow E_{\text{VMC}}^{\text{corr}} = 61\% \\ 1s^2 2s^2 \oplus 1s^2 2p^2 \quad \times \mathcal{J}(r_{ij}) \quad \rightarrow E_{\text{VMC}}^{\text{corr}} = 93\%$$

Example: $E_{\text{VMC}}^{\text{corr}}$ and $E_{\text{DMC}}^{\text{corr}}$ for 1st-row dimers

MO orbitals with atomic s - p Slater basis (all-electron)

Active MO orbitals are $2\sigma_g, 2\sigma_u, 3\sigma_g, 3\sigma_u, 1\pi_u, 1\pi_g$

5th-order polynomial \mathcal{J} (e-n, e-e, e-e-n)



Filippi and Umrigar, J. Chem. Phys. **105**, 213 (1996)

Determinant versus Jastrow factor

Determinantal part yields the nodes (zeros) of wave function

⇒ Quality of the fixed-node DMC solution

Why bother with the Jastrow factor?

Implications of using a good Jastrow factor for DMC:

- ▷ Efficiency: Smaller σ and time-step error ⇒ Gain in CPU time
- ▷ Expectation values other than energy ⇒ Mixed estimator
- ▷ Non-local pseudopotentials and localization error
⇒ Jastrow factor does affect fixed-node energy

Why should $\Psi_{\text{QMC}} = \mathcal{J}D$ work?

Full wave-function

Ψ

↓

Full Hamiltonian

\mathcal{H}

$$\mathcal{H}\Psi = E\Psi$$

→

Factorized wave-function

$\mathcal{J}\Phi$

↓

Effective Hamiltonian

\mathcal{H}_{eff}

$$\mathcal{H}\mathcal{J}\Phi = E\mathcal{J}\Phi \rightarrow \frac{\mathcal{H}\mathcal{J}}{\mathcal{J}}\Phi = E\Phi$$

$$\mathcal{H}_{\text{eff}}\Phi = E\Phi$$

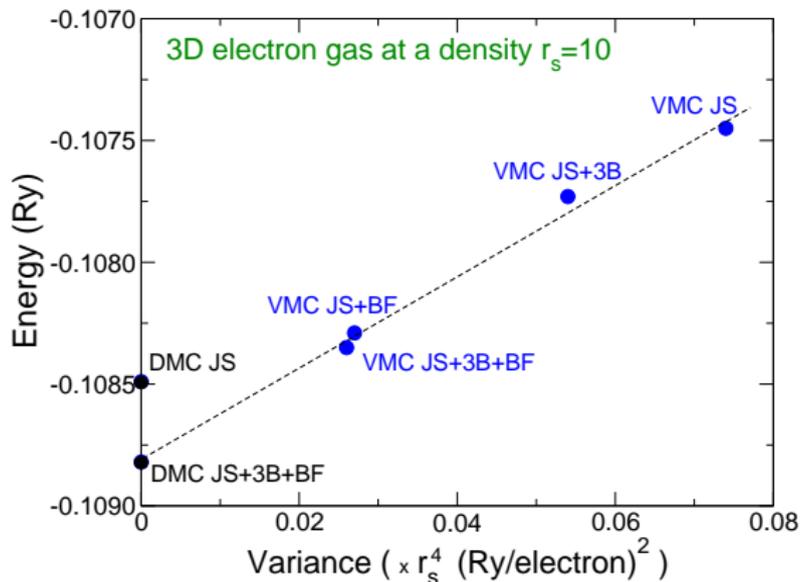
\mathcal{H}_{eff} weaker Hamiltonian than \mathcal{H}

⇒ $\Phi \approx$ non-interacting wave function D

⇒ Quantum Monte Carlo wave function $\Psi = \mathcal{J}D$

Why going beyond VMC?

Dependence of VMC from wave function Ψ



Kwon, Ceperley, Martin, Phys. Rev. B **58**, 6800 (1998)

Why going beyond VMC?

- ▷ Dependence on wave function: What goes in, comes out!
- ▷ No automatic way of constructing wave function Ψ
Choices must be made about functional form (human time)
- ▷ Hard to ensure good error cancelation on energy differences
e.g. easier to construct good Ψ for closed than open shells

Can we remove wave function bias?

Projector Monte Carlo methods

- ▷ Construct an operator which inverts spectrum of \mathcal{H}
- ▷ Use it to stochastically project the ground state of \mathcal{H}

Diffusion Monte Carlo

$$\exp[-\tau(\mathcal{H} - E_T)]$$

Green's function Monte Carlo

$$1/(\mathcal{H} - E_T)$$

Power Monte Carlo

$$E_T - \mathcal{H}$$

Diffusion Monte Carlo

Consider initial guess $\Psi^{(0)}$ and repeatedly apply projection operator

$$\Psi^{(n)} = e^{-\tau(\mathcal{H}-E_T)}\Psi^{(n-1)}$$

Expand $\Psi^{(0)}$ on the eigenstates Ψ_i with energies E_i of \mathcal{H}

$$\Psi^{(n)} = e^{-n\tau(\mathcal{H}-E_T)}\Psi^{(0)} = \sum_i \Psi_i \langle \Psi^{(0)} | \Psi_i \rangle e^{-n\tau(E_i-E_T)}$$

and obtain in the limit of $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \Psi^{(n)} = \Psi_0 \langle \Psi^{(0)} | \Psi_0 \rangle e^{-n\tau(E_0-E_T)}$$

If we choose $E_T \approx E_0$, we obtain

$$\lim_{n \rightarrow \infty} \Psi^{(n)} = \Psi_0$$

How do we perform the projection?

Rewrite projection equation in integral form

$$\Psi^{(n)}(\mathbf{R}', t + \tau) = \int d\mathbf{R} G(\mathbf{R}', \mathbf{R}, \tau) \Psi^{(n-1)}(\mathbf{R}, t)$$

where $G(\mathbf{R}', \mathbf{R}, \tau) = \langle \mathbf{R}' | e^{-\tau(\mathcal{H} - E_T)} | \mathbf{R} \rangle$

▷ Can we sample the wave function?

For the moment, assume we are dealing with **bosons**, so $\Psi > 0$

▷ Can we interpret $G(\mathbf{R}', \mathbf{R}, \tau)$ as a transition probability?

If yes, we can perform this integral by Monte Carlo integration

VMC and DMC as power methods

VMC Distribution function is given $\rho(\mathbf{R}) = \frac{|\Psi(\mathbf{R})|^2}{\int d\mathbf{R} |\Psi(\mathbf{R})|^2}$

Construct M which satisfies stationarity condition $M\rho = \rho$

→ ρ is eigenvector of M with eigenvalue 1

→ ρ is the dominant eigenvector $\Rightarrow \lim_{n \rightarrow \infty} M^n \rho_{\text{initial}} = \rho$

DMC Opposite procedure!

The matrix M is given $\rightarrow M = \langle \mathbf{R}' | e^{-\tau(\mathcal{H} - E_T)} | \mathbf{R} \rangle$

We want to find the dominant eigenvector $\rho = \Psi_0$

What can we say about the Green's function?

$$G(\mathbf{R}', \mathbf{R}, \tau) = \langle \mathbf{R}' | e^{-\tau(\mathcal{H} - E_T)} | \mathbf{R} \rangle$$

$G(\mathbf{R}', \mathbf{R}, \tau)$ satisfies the imaginary-time Schrödinger equation

$$(\mathcal{H} - E_T)G(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial G(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}$$

with $G(\mathbf{R}', \mathbf{R}, 0) = \delta(\mathbf{R}' - \mathbf{R})$

Can we interpret $G(\mathbf{R}', \mathbf{R}, \tau)$ as a transition probability?

(1)

$$\mathcal{H} = \mathcal{T}$$

Imaginary-time Schrödinger equation is a diffusion equation

$$-\frac{1}{2}\nabla^2 G(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial G(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}$$

The Green's function is given by a Gaussian

$$G(\mathbf{R}', \mathbf{R}, \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R})^2}{2\tau}\right]$$

Positive and can be sampled

Can we interpret $G(\mathbf{R}', \mathbf{R}, \tau)$ as a transition probability? (2)

$$\mathcal{H} = \mathcal{V}$$

$$(\mathcal{V}(\mathbf{R}) - E_T)G(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial G(\mathbf{R}, \mathbf{R}_0, t)}{\partial t},$$

The Green's function is given by

$$G(\mathbf{R}', \mathbf{R}, \tau) = \exp[-\tau(\mathcal{V}(\mathbf{R}) - E_T)] \delta(\mathbf{R} - \mathbf{R}'),$$

Positive but does not preserve the normalization

It is a factor by which we multiply the distribution $\Psi(\mathbf{R}, t)$

$\mathcal{H} = \mathcal{T} + \mathcal{V}$ and a combination of diffusion and branching

Trotter's theorem \rightarrow $e^{(A+B)\tau} = e^{A\tau} e^{B\tau} + \mathcal{O}(\tau^2)$

$$\begin{aligned}\langle \mathbf{R}' | e^{-\mathcal{H}\tau} | \mathbf{R}_0 \rangle &\approx \langle \mathbf{R}' | e^{-T\tau} e^{-\mathcal{V}\tau} | \mathbf{R}_0 \rangle \\ &= \int d\mathbf{R}'' \langle \mathbf{R}' | e^{-T\tau} | \mathbf{R}'' \rangle \langle \mathbf{R}'' | e^{-\mathcal{V}\tau} | \mathbf{R}_0 \rangle \\ &= \langle \mathbf{R}' | e^{-T\tau} | \mathbf{R}_0 \rangle e^{-\mathcal{V}(\mathbf{R}_0)\tau}\end{aligned}$$

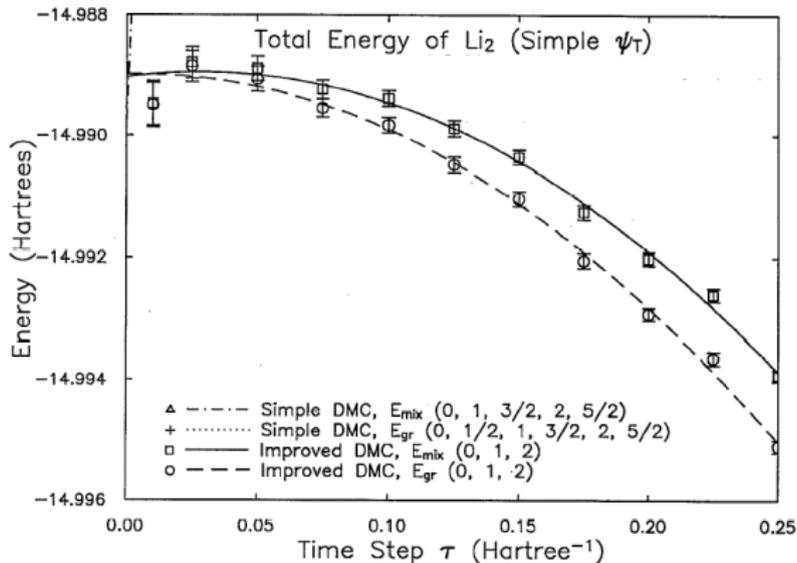
The Green's function in the $\text{short-time approximation}$ to $\mathcal{O}(\tau^2)$ is

$$G(\mathbf{R}', \mathbf{R}, \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R})^2}{2\tau}\right] \exp[-\tau(\mathcal{V}(\mathbf{R}) - E_T)]$$

DMC results must be extrapolated at short time-steps ($\tau \rightarrow 0$)

Time-step extrapolation

Example: Energy of Li_2 versus time-step τ



Umrigar, Nightingale, Runge, J. Chem. Phys. **94**, 2865 (1993)

The basic DMC algorithm is rather simple:

1. Sample $\Psi^{(0)}(\mathbf{R})$ with the Metropolis algorithm

Generate M_0 walkers $\mathbf{R}_1, \dots, \mathbf{R}_{M_0}$ (zeroth generation)

2. Diffuse each walker as $\mathbf{R}' = \mathbf{R} + \xi$

where ξ is sampled from $g(\xi) = (2\pi\tau)^{-3N/2} \exp(-\xi^2/2\tau)$

3. For each walker, compute the factor

$$p = \exp[-\tau(\mathcal{V}(\mathbf{R}) - E_T)]$$

Branch the walker with p the probability to survive

Continue →

4. Branch the walker with p the probability to survive

▷ If $p < 1$, the walker survives with probability p

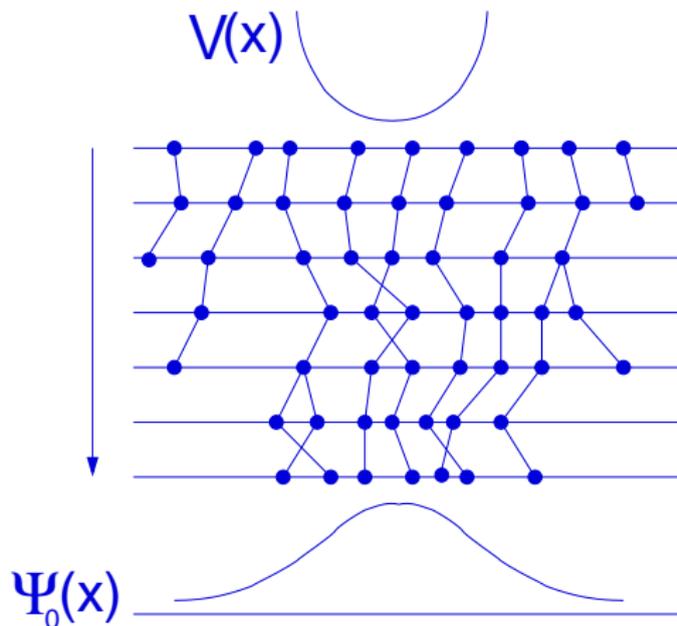
▷ If $p > 1$, the walker continues and new walkers with the same coordinates are created with probability $p - 1$

⇒ Number of copies of the current walker equal to $\text{int}(p + \eta)$
where η is a random number between $(0,1)$

5. Adjust E_T so that population fluctuates around target M_0

→ After many iterations, walkers distributed as $\Psi_0(\mathbf{R})$

Diffusion and branching in a harmonic potential



Walkers proliferate/die in regions of lower/higher potential than E_T

Some comments on the simple DMC algorithm

- ▷ E_T is adjusted to keep population stable

IF $M(t)$ is the current and M_0 the desired population

$$M(t + T) = M(t) e^{-T(-\delta E_T)} = M_0 \Rightarrow \delta E_T = \frac{1}{T} \ln \left[\frac{M_0}{M(t)} \right]$$

If $E_{\text{est}}(t)$ is current best estimate of the ground state

$$E_T(t + \tau) = E_{\text{est}}(t) + \frac{1}{g\tau} \ln [M_0/M(t)]$$

⇒ Feedback on E_T introduces population control bias

- ▷ Symmetric branching $\exp[-\tau(\mathcal{V}(\mathbf{R}) + \mathcal{V}(\mathbf{R}'))/2]$ starting from

$$e^{(A+B)\tau} = e^{A\tau/2} e^{B\tau} e^{A\tau/2} + \mathcal{O}(\tau^3)$$

Problems with simple algorithm

The simple algorithm is inefficient and unstable

- ▷ Potential can vary a lot and be unbounded
e.g. electron-nucleus interaction → Exploding population
- ▷ Branching factor grows with system size

Importance sampling

Start from integral equation

$$\psi^{(n)}(\mathbf{R}', t + \tau) = \int d\mathbf{R} G(\mathbf{R}', \mathbf{R}, \tau) \psi^{(n-1)}(\mathbf{R}, t)$$

Multiply each side by trial Ψ and define $f^{(n)}(\mathbf{R}) = \Psi(\mathbf{R})\psi^{(n)}(\mathbf{R})$

$$f^{(n)}(\mathbf{R}', t + \tau) = \int d\mathbf{R} \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) f^{(n-1)}(\mathbf{R}, t)$$

where the importance sampled Green's function is

$$\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = \Psi(\mathbf{R}') \langle \mathbf{R}' | e^{-\tau(\mathcal{H} - E_T)} | \mathbf{R} \rangle / \Psi(\mathbf{R})$$

We obtain $\lim_{n \rightarrow \infty} f^{(n)}(\mathbf{R}) = \Psi(\mathbf{R})\psi_0(\mathbf{R})$

Importance sampled Green's function

The importance sampled $\tilde{G}(\mathbf{R}, \mathbf{R}_0, \tau)$ satisfies

$$-\frac{1}{2}\nabla^2\tilde{G} + \nabla \cdot [\tilde{G}\mathbf{V}(\mathbf{R})] + [E_L(\mathbf{R}) - E_T]\tilde{G} = -\frac{\partial\tilde{G}}{\partial\tau}$$

with the quantum velocity $\mathbf{V}(\mathbf{R}) = \frac{\nabla\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$

We now have **drift** in addition to diffusion and branching terms

Trotter's theorem \Rightarrow Consider them separately for small enough τ

The drift-branching components: Reminder

Diffusion term

$$-\frac{1}{2}\nabla^2\tilde{G}(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial\tilde{G}(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}$$

$$\Rightarrow \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R})^2}{2\tau}\right]$$

Branching term

$$(E_L(\mathbf{R}) - E_T)\tilde{G}(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial\tilde{G}(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}$$

$$\Rightarrow \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = \exp[-\tau(E_L(\mathbf{R}) - E_T)] \delta(\mathbf{R} - \mathbf{R}')$$

The drift-diffusion-branching Green's function

$$-\frac{1}{2}\nabla^2\tilde{G} + \nabla \cdot [\tilde{G}\mathbf{V}(\mathbf{R})] + [E_L(\mathbf{R}) - E_T]\tilde{G} = -\frac{\partial\tilde{G}}{\partial\tau}$$

Drift term

Assume $\mathbf{V}(\mathbf{R}) = \frac{\nabla\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$ constant over the move (true as $\tau \rightarrow 0$)

The drift operator becomes $\mathbf{V} \cdot \nabla + \nabla \cdot \mathbf{V} \approx \mathbf{V} \cdot \nabla$ so that

$$\mathbf{V} \cdot \nabla \tilde{G}(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial\tilde{G}(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}$$

with solution $\tilde{G}(\mathbf{R}, \mathbf{R}_0, t) = \delta(\mathbf{R} - \mathbf{R}_0 - \mathbf{V}t)$

The drift-diffusion-branching Green's function

Drift-diffusion-branching short-time Green's function is

$$\begin{aligned} \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) &= (2\pi\tau)^{-3N/2} \exp \left[-\frac{(\mathbf{R}' - \mathbf{R} - \tau\mathbf{V}(\mathbf{R}))^2}{2\tau} \right] \times \\ &\times \exp \left\{ -\tau \left[(E_L(\mathbf{R}) + E_L(\mathbf{R}'))/2 - E_T \right] \right\} + \mathcal{O}(\tau^2) \end{aligned}$$

What is new in the drift-diffusion-branching expression?

- ▷ $\mathbf{V}(\mathbf{R})$ pushes walkers where Ψ is large
- ▷ $E_L(\mathbf{R})$ is better behaved than the potential $\mathcal{V}(\mathbf{R})$
Cusp conditions \Rightarrow No divergences when particles approach
As $\Psi \rightarrow \Psi_0$, $E_L \rightarrow E_0$ and branching factor is smaller

DMC algorithm with importance sampling

1. Sample initial walkers from $|\Psi(\mathbf{R})|^2$
2. Drift and diffuse the walkers as $\mathbf{R}' = \mathbf{R} + \xi + \tau\mathbf{V}(\mathbf{R})$
where ξ is sampled from $g(\xi) = (2\pi\tau)^{-3N/2} \exp(-\xi^2/2\tau)$
3. Branching step as in the simple algorithm but with the factor

$$\rho = \exp\{-\tau[(E_L(\mathbf{R}) + E_L(\mathbf{R}'))/2 - E_T]\}$$

4. Adjust the trial energy to keep the population stable

→ After many iterations, walkers distributed as $\Psi(\mathbf{R})\Psi_0(\mathbf{R})$

An important and simple improvement

If $\Psi = \Psi_0$, $E_L(\mathbf{R}) = E_0 \rightarrow$ No branching term \rightarrow Sample Ψ^2

Due to time-step approximation, we only sample Ψ^2 as $\tau \rightarrow 0$!

Solution Introduce accept/reject step like in Metropolis algorithm

$$\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) \approx \underbrace{\mathcal{N} \exp \left[-\frac{(\mathbf{R}' - \mathbf{R} - \mathbf{V}(\mathbf{R})\tau)^2}{2\tau} \right]}_{T(\mathbf{R}', \mathbf{R}, \tau)} \exp \left[-(E_L(\mathbf{R}) + E_L(\mathbf{R}')) \frac{\tau}{2} \right]$$

Walker drifts, diffuses and the move is accepted with probability

$$p = \min \left\{ 1, \frac{|\Psi(\mathbf{R}')|^2 T(\mathbf{R}, \mathbf{R}', \tau)}{|\Psi(\mathbf{R})|^2 T(\mathbf{R}', \mathbf{R}, \tau)} \right\}$$

\rightarrow Improved algorithm with smaller time-step error

Electrons are fermions!

We assumed that $\Psi_0 > 0$ and that we are dealing with bosons

Fermions $\rightarrow \Psi$ is antisymmetric and changes sign!

How can we impose antisymmetry in simple DMC method?

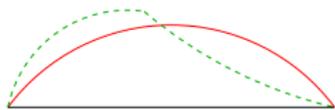
Idea Rewrite initial distribution $\Psi^{(0)}$ as

$$\Psi^{(0)} = \Psi_+^{(0)} - \Psi_-^{(0)}$$

and evolve $\Psi_+^{(0)}$ and $\Psi_-^{(0)}$ separately. Will this idea work?

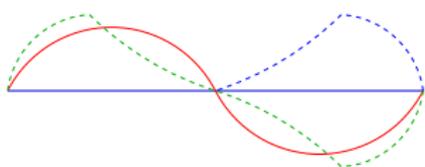
Particle in a box and the fermionic problem

Ground state $\Psi^{(0)}(\mathbf{R}) \rightarrow \Psi_0(\mathbf{R})$



Excited state $\Psi_1(\mathbf{R})$ changes sign!

Let us try our trick $\rightarrow \Psi^{(0)}(\mathbf{R}) = \Psi_+^{(0)}(\mathbf{R}) - \Psi_-^{(0)}(\mathbf{R})$



$\Psi_-^{(0)}(\mathbf{R}), \Psi_+^{(0)}(\mathbf{R}) \rightarrow \Psi_0(\mathbf{R})$

Is a trick possible for DMC with importance sampling?

Does it help to work with $f(\mathbf{R}) = \Psi(\mathbf{R})\Psi_0(\mathbf{R})$?

▷ Initial distribution $\Psi(\mathbf{R})^2 > 0$ poses no problems. Good start!

▷ Iterate as $f^{(n)}(\mathbf{R}', t + \tau) = \int d\mathbf{R} \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) f^{(n-1)}(\mathbf{R}, t)$

If move $\mathbf{R} \rightarrow \mathbf{R}'$ changes sign of Ψ so that $\Psi(\mathbf{R}')/\Psi(\mathbf{R}) < 0$

$\Rightarrow \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = \Psi(\mathbf{R}') \langle \mathbf{R}' | e^{-\tau(\mathcal{H} - E_T)} | \mathbf{R} \rangle / \Psi(\mathbf{R})$ changes sign!

We have no luck !?!

See next lecture by Lubos Mitás