Introduction to quantum Monte Carlo methods

Part 2

Claudia Filippi

MESA+ Institute for Nanotechnology, Universiteit Twente, The Netherlands

Han-Sur-Lesse Winter School, Nov 27-Dec 1, 2023

Recap: where we started (1)

Solve the Schrödinger equation for the electrons in the ionic field

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

Solve for the wave function of the interacting electron system

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

Possible "simple" approach

$\rightarrow$ Start from one-particle $\{\psi_i(\mathbf{x})\}$

$\rightarrow$ Build non-interacting many-body basis of determinants

Example: Configuration interaction

$$\Psi_{\mathrm{CI}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_k c_k D_k$$

$$\Updownarrow$$

$$\Psi_{\mathrm{CI}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \quad \cdots$$

Build solution as expansion over determinants

Advantage of a linear expansion over determinants?

Find optimal coefficients by solving a generalized secular problem

$$\Psi_{\mathrm{CI}} = \sum_{i=1}^{K} c_i D_i \;\; \Rightarrow \;\; \boxed{\sum_{j=1}^{K} \langle D_i | \mathcal{H} | D_j \rangle c_j^{(k)} = E_{\mathrm{CI}}^{(k)} \sum_{j=1}^{K} \langle D_i | D_j \rangle c_j^{(k)}}$$

If Gaussian basis $\rightarrow$ All integrals computed analytically !

but slow convergence . . .

What about a wave function depending on $\boxed{r_{ij}}$? Use Monte Carlo

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

$$
\begin{aligned}
E_V \quad &= \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \\
&= \int \mathrm{d}\mathbf{R} \, \frac{\mathcal{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \, \frac{|\Psi(\mathbf{R})|^2}{\int \mathrm{d}\mathbf{R} |\Psi(\mathbf{R})|^2} \\
&= \int \mathrm{d}\mathbf{R} \, E_{\mathrm{L}}(\mathbf{R}) \, P(\mathbf{R}) \\
&= \langle E_{\mathrm{L}}(\mathbf{R}) \rangle_P
\end{aligned}
$$

Obtain M samples distributed as $P(\mathbf{R})$ and estimate E_V as

$$
\bar{E}_V = \frac{1}{M} \sum_{i=1}^{M} E_{\mathrm{L}}(\mathbf{R}_i)
$$

The energy and variance can be rewritten as

$$
\begin{aligned}
E_V &= \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle E_{\mathrm{L}}(\mathbf{R}) \rangle_P \\
\sigma_E^2 &= \frac{\langle \Psi | (\mathcal{H} - E_V)^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle (E_{\mathrm{L}}(\mathbf{R}) - E_V)^2 \rangle_P
\end{aligned}
$$

and estimated as averages

$$
\begin{aligned}
\bar{E}_V &= \frac{1}{M} \sum_{i=1}^{M} E_{\mathrm{L}}(\mathbf{R}_i) \\
\bar{\sigma}_E^2 &= \frac{1}{M} \sum_{i=1}^{M} (E_{\mathrm{L}}(\mathbf{R}_i) - \bar{E}_V)^2
\end{aligned}
$$

The statistical Monte Carlo error goes as $\mathrm{err}(E_V) \sim \dfrac{\sigma_E}{\sqrt{M}}$

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

To obtain M samples distributed as $P(\mathbf{R})$, generate a Markov chain

$$\ldots \xrightarrow{M} \mathbf{R} \xrightarrow{M} \mathbf{R}' \xrightarrow{M} \mathbf{R}'' \xrightarrow{M} \ldots$$
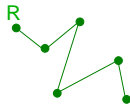
using $M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ probability for transition $\mathbf{R}_\mathrm{i} \to \mathbf{R}_\mathrm{f}$

If $M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) \geq 0$ and $\displaystyle\int d\mathbf{R}_\mathrm{f} M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = 1$

and if M satisfies stationarity condition:

$$\int \mathrm{d}\mathbf{R}_\mathrm{i}\, M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\, P(\mathbf{R}_\mathrm{i}) = P(\mathbf{R}_\mathrm{f}) \quad \forall\, \mathbf{R}_\mathrm{f}$$

$\Rightarrow$ Any initial distribution will evolve to P

In practice, we impose detailed balance condition

$$M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\, P(\mathbf{R}_\mathrm{i}) = M(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\, P(\mathbf{R}_\mathrm{f})$$
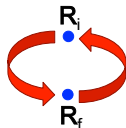
We write M as proposal $T \times$ acceptance A

$$M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\, T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$$
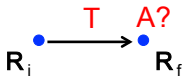
and choose $\quad A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \min\left\{1, \dfrac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\, P(\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\, P(\mathbf{R}_\mathrm{i})}\right\}$

so that detailed balance condition is satisfied $\rightarrow P(\mathbf{R})$ is sampled

Some notes about Metropolis algorithm

a) $P(\mathbf{R}) = \dfrac{|\Psi(\mathbf{R})|^2}{\int d\mathbf{R} |\Psi(\mathbf{R})|^2}$ does not have to be normalized

   $\rightarrow$ For complex Ψ we do not know the normalization!

b) Data generated in a Monte Carlo run are correlated
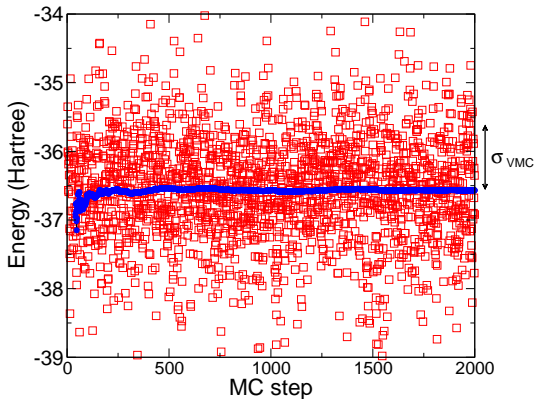
   $\rightarrow$ Choose a good T !

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

   $\rightarrow$ Fewer Monte Carlo steps needed to estimate E_V

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

$E_{\mathrm{VMC}} = \langle E_{\mathrm{L}}(\mathbf{R}) \rangle_P = -36.542 \pm 0.001$ Hartree (40×20000 steps)

$\sigma_{\mathrm{VMC}} = \langle (E_{\mathrm{L}}(\mathbf{R}) - E_{\mathrm{VMC}})^2 \rangle_P = 0.90$ Hartree

Variational Monte Carlo $\rightarrow$ Freedom in choice of Ψ

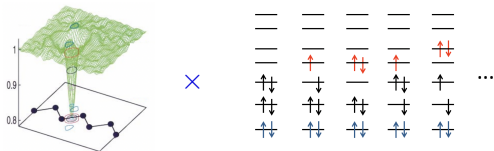Monte Carlo integration allows the use of complex and accurate Ψ

$\Rightarrow$ More compact representation of Ψ than in quantum chemistry

$\Rightarrow$ Beyond $\quad c_0 D_{\mathrm{HF}} + c_1 D_1 + c_2 D_2 + \ldots$ millions of determinants

Commonly employed compact Jastrow-Slater wave functions

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \mathcal{J}(\mathbf{r}_1, \ldots, \mathbf{r}_N) \times \sum_i c_i \, D_i(\mathbf{r}_1, \ldots, \mathbf{r}_N)$$

$\boxed{\mathcal{J}} \longrightarrow$ Jastrow correlation factor

− Explicit dependence on electron-electron distances $\boxed{r_{ij}}$

$\boxed{\sum c_i \, D_i} \longrightarrow$ Determinants of single-particle orbitals

− $\boxed{\text{Few}}$ and not millions of determinants

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 $\boxed{\dfrac{\mathcal{H}\Psi}{\Psi} = -\dfrac{1}{2}\sum_i \dfrac{\nabla_i^2 \Psi}{\Psi} + \mathcal{V}}$ must be $\boxed{\text{finite}}$

$\Rightarrow$ 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 $r_{ij} \to 0$ while all other particles are well separated

Keep only diverging terms in $\dfrac{\mathcal{H}\Psi}{\Psi}$ for particles i and j

and go to relative coordinates close to $\mathbf{r} = \mathbf{r}_{ij} = 0$

$$-\frac{1}{2\mu_{ij}}\frac{\nabla^2\Psi}{\Psi} + \mathcal{V}(r) \sim -\frac{1}{2\mu_{ij}}\frac{1}{\Psi}\left(\frac{\partial^2\Psi}{\partial r^2} + \frac{2}{r}\frac{\partial\Psi}{\partial r}\right) + \mathcal{V}(r) + \dots$$

$$\sim \boxed{-\frac{1}{\mu_{ij}}\frac{1}{r}\frac{\Psi'}{\Psi} + \mathcal{V}(r)} + \dots$$

where $\mu_{ij} = m_i\,m_j/(m_i + m_j)$ and we assumed $\Psi(r = r_{ij} = 0) \neq 0$

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 <u>Kato's cusp conditions</u>:

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

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

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

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

$$\frac{\Psi'}{\Psi} = \mu_{ij} q_i\, q_j$$

- Electron-nucleus: $\quad \mu = 1,\ q_i = 1,\ q_j = -Z \quad \Rightarrow \quad \boxed{\left.\frac{\Psi'}{\Psi}\right|_{r=0} = -Z}$

- Electron-electron: $\quad \mu = \dfrac{1}{2},\ q_i = 1,\ q_j = 1 \quad \Rightarrow \quad \boxed{\left.\frac{\Psi'}{\Psi}\right|_{r=0} = 1/2}$

▷ 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 $\quad b_0^{\uparrow\downarrow} = \frac{1}{2} \quad$ or $\quad b_0^{\uparrow\uparrow} = b_0^{\downarrow\downarrow} = \frac{1}{4}$
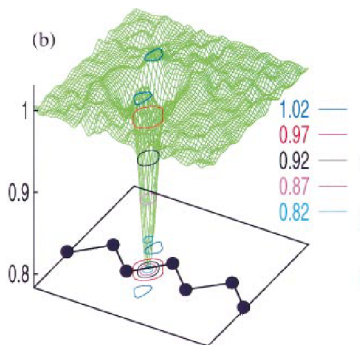
Imposes cusp conditions
$+$
keeps electrons apart
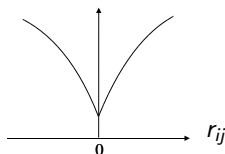
▷ Electron-nucleus cusps imposed through the determinantal part

Pair correlation function for ↑↓ 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)

Dynamic and static correlation

$\Psi = $ Jastrow $\times$ Determinants $\rightarrow$ 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

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

HF ground state configuration $\boxed{1s^2 2s^2}$

Additional important configuration $\boxed{1s^2 2p^2}$
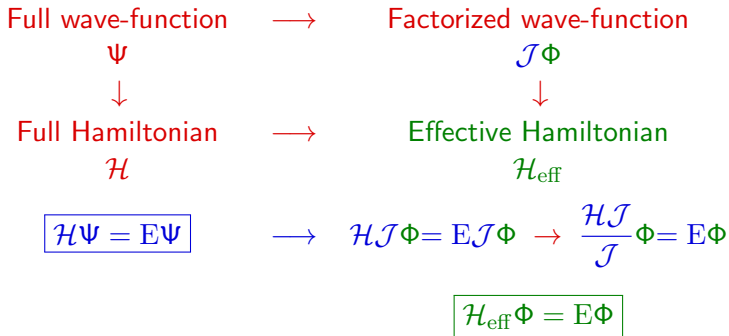
$$
\begin{array}{lll}
1s^2 2s^2 & \times\ \mathcal{J}(r_{ij}) & \to E_{\mathrm{VMC}}^{\mathrm{corr}} = 61\% \\
1s^2 2s^2 \oplus 1s^2 2p^2 & \times\ \mathcal{J}(r_{ij}) & \to E_{\mathrm{VMC}}^{\mathrm{corr}} = 93\%
\end{array}
$$

where $E^{\mathrm{corr}} = E_{\mathrm{exact}} - E_{\mathrm{HF}}$

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

Full wave-function $\longrightarrow$ Factorized wave-function
Ψ $\mathcal{J}\Phi$
$\downarrow$ $\downarrow$
Full Hamiltonian $\longrightarrow$ Effective Hamiltonian
$\mathcal{H}$ $\mathcal{H}_{\mathrm{eff}}$

$$\mathcal{H}\Psi = \mathrm{E}\Psi \qquad \longrightarrow \quad \mathcal{H}\mathcal{J}\Phi = \mathrm{E}\mathcal{J}\Phi \ \rightarrow \ \frac{\mathcal{H}\mathcal{J}}{\mathcal{J}}\Phi = \mathrm{E}\Phi$$

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

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

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

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

Optimize by energy minimization $\rightarrow$ Tricky?

Statistical error: Both a blessing and a curse!

$\Psi_{\{\alpha_k\}} \rightarrow$ Energy and its derivatives wrt parameters $\{\alpha_k\}$

$$
E_V = \int d\mathbf{R} \, \frac{\mathcal{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \frac{|\Psi(\mathbf{R})|^2}{\int d\mathbf{R}|\Psi(\mathbf{R})|^2} = \langle E_L \rangle_{\Psi^2}
$$

$$
\partial_k E_V = \left\langle \frac{\partial_k \Psi}{\Psi} E_L + \frac{\mathcal{H}\partial_k \Psi}{\Psi} - 2E_V \frac{\partial_k \Psi}{\Psi} \right\rangle_{\Psi^2}
$$

$$
= \boxed{2 \left\langle \frac{\partial_k \Psi}{\Psi}(E_L - E_V) \right\rangle_{\Psi^2}}
$$

The last expression is obtained using Hermiticity of $\mathcal{H}$

Use expressions with smaller fluctuations

Two mathematically equivalent expressions of the energy gradient

$$\partial_k E_V = \left\langle \frac{\partial_k \Psi}{\Psi} E_{\mathrm{L}} + \frac{\mathcal{H} \partial_k \Psi}{\Psi} - 2E_V \frac{\partial_k \Psi}{\Psi} \right\rangle_{\Psi^2} = \boxed{2 \left\langle \frac{\partial_k \Psi}{\Psi} (E_{\mathrm{L}} - E_V) \right\rangle_{\Psi^2}}$$

Why using the last expression?

Lower fluctuations $\rightarrow 0$ as $\Psi \rightarrow \Psi_0$
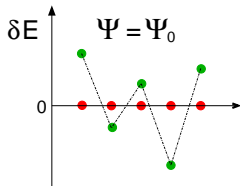
Play similar tricks with the Hessian or its approximations

$\rightarrow$ Toolbox of optimization approaches

Umrigar, Toulouse, Sorella, Filippi, Scemama (2005-2007)

Important developments → Efficient computation of derivatives!

Computing Ψ, E_L, $\partial\Psi$, ∂E_L for many parameters + determinants

Example: Polyenes C_nH_{n+2} → from C_4H_6 to $\boxed{C_{60}H_{62}}$

CPU$_{vmc+derivs}$ / CPU$_{vmc}$

wave function optimization

number of parameters

CPU$_{vmc+forces}$ / CPU$_{vmc}$

C_4H_6
C_8H_{10}
$C_{16}H_{18}$
$C_{28}H_{30}$
$C_{60}H_{62}$

geometry optimization

number of determinants
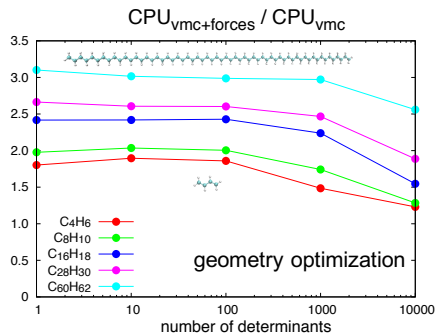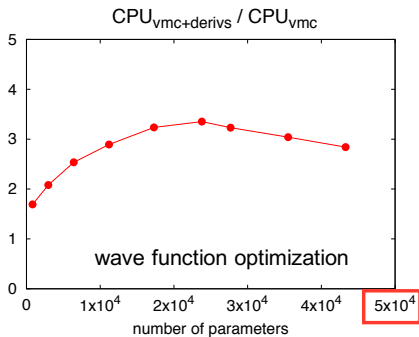
+ Optimization tools → QMC "internally consistent"

Filippi, Assaraf, Moroni, JCP (2016); JCTC (2017)

Beyond VMC?

Removing or reducing wave function bias?

$\Rightarrow$ Projection Monte Carlo methods

Dependence of VMC from wave function Ψ
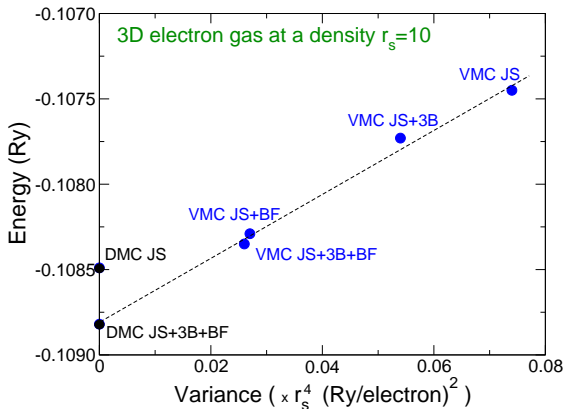
Kwon, Ceperley, Martin, Phys. Rev. B **58**, 6800 (1998)

Why going beyond VMC?

What goes in, comes out! Can we remove wave function bias?

Projector (diffusion) Monte Carlo method

▷ Construct an operator which inverts spectrum of $\mathcal{H}$

Diffusion Monte Carlo $\rightarrow e^{-\tau(\mathcal{H}-E_{\mathrm{ref}})}$

▷ Use it to stochastically project the ground state of $\mathcal{H}$

Consider initial guess $|\Psi^{(0)}\rangle$ (e.g. $\mathcal{J}D$ we used in VMC)

Repeatedly apply projection operator

$$|\Psi^{(n)}\rangle = e^{-\tau(\mathcal{H}-E_{\mathrm{ref}})}|\Psi^{(n-1)}\rangle$$

To understand what one obtains, perform a "thought experiment"

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

$$
\begin{aligned}
|\Psi^{(n)}\rangle &= e^{-n\tau(\mathcal{H}-E_{\mathrm{ref}})}|\Psi^{(0)}\rangle = e^{-n\tau(\mathcal{H}-E_{\mathrm{ref}})}\sum_i |\Psi_i\rangle\langle\Psi_i|\Psi^{(0)}\rangle \\
&= \sum_i e^{-n\tau(\mathcal{H}-E_{\mathrm{ref}})}|\Psi_i\rangle\langle\Psi_i|\Psi^{(0)}\rangle \\
&= \sum_i e^{-n\tau(E_i-E_{\mathrm{ref}})}|\Psi_i\rangle\langle\Psi_i|\Psi^{(0)}\rangle
\end{aligned}
$$

Continuing from previous page

$$
\begin{aligned}
|\Psi^{(n)}\rangle &= e^{-n\tau(E_0 - E_{\mathrm{ref}})} \sum_i e^{-n\tau(E_i - E_0)} |\Psi_i\rangle \langle\Psi_i|\Psi^{(0)}\rangle \\
&= e^{-n\tau(E_0 - E_{\mathrm{ref}})} \left(|\Psi_0\rangle \langle\Psi_0|\Psi^{(0)}\rangle + e^{-n\tau(E_1 - E_0)} |\Psi_1\rangle \langle\Psi_1|\Psi^{(0)}\rangle \ldots \right.
\end{aligned}
$$

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

$$
\lim_{n\to\infty} |\Psi^{(n)}\rangle = |\Psi_0\rangle \langle\Psi_0|\Psi^{(0)}\rangle e^{-n\tau(E_0 - E_{\mathrm{ref}})}
$$

If we choose $E_{\mathrm{ref}} \approx E_0$, we obtain $\boxed{\lim_{n\to\infty} |\Psi^{(n)}\rangle = |\Psi_0\rangle}$

$$|\Psi^{(n)}\rangle = e^{-\tau(\mathcal{H}-E_{\mathrm{ref}})}|\Psi^{(n-1)}\rangle$$

Rewrite projection equation in real-space representation

$$\Rightarrow \langle \mathbf{R}'|\Psi^{(n)}\rangle = \langle \mathbf{R}'|e^{-\tau(\mathcal{H}-E_{\mathrm{ref}})}|\Psi^{(n-1)}\rangle$$

$$\langle \mathbf{R}'|\Psi^{(n)}\rangle = \int \mathrm{d}\mathbf{R}\, \langle \mathbf{R}'|e^{-\tau(\mathcal{H}-E_{\mathrm{ref}})}|\mathbf{R}\rangle\langle \mathbf{R}|\Psi^{(n-1)}\rangle$$

which we rewrite as

$$\Psi(\mathbf{R}', t + \tau) = \int \mathrm{d}\mathbf{R}\, G(\mathbf{R}', \mathbf{R}, \tau)\Psi(\mathbf{R}, t)$$

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

Projection equation in integral form

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

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

▷ Can we sample the wave function?

For the moment, assume we are dealing with $\boxed{\text{bosons}}$, 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 $\boxed{P(\mathbf{R}) = \dfrac{|\Psi(\mathbf{R})|^2}{\int d\mathbf{R}|\Psi(\mathbf{R})|^2}}$

$\boxed{\text{Construct } M}$ which satisfies stationarity condition so that

$$\lim_{n\to\infty} \int d\mathbf{R}_n \cdots d\mathbf{R}_1 M(\mathbf{R}, \mathbf{R}_n) \cdots M(\mathbf{R}_3, \mathbf{R}_2) M(\mathbf{R}_2, \mathbf{R}_1) P_{\text{init}}(\mathbf{R}_1) = P(\mathbf{R})$$

DMC Opposite procedure!

The matrix M is given $\rightarrow$ $\boxed{M \equiv G = \langle \mathbf{R}' | e^{-\tau(\mathcal{H} - E_{\text{ref}})} | \mathbf{R} \rangle}$

We do not know $\boxed{P}$!

$$\lim_{n\to\infty} \int d\mathbf{R}_n \cdots d\mathbf{R}_1 G(\mathbf{R}, \mathbf{R}_n) \cdots G(\mathbf{R}_3, \mathbf{R}_2) G(\mathbf{R}_2, \mathbf{R}_1) P_{\text{init}}(\mathbf{R}_1) = \Psi_0(\mathbf{R})$$

In either case, we want to find the dominant eigenvector of M

$$\Psi(t) = e^{-t(\mathcal{H} - E_{\mathrm{ref}})}\Psi^{(0)}$$

$\Psi(\mathbf{R}, t)$ satisfies the imaginary-time Schrödinger equation

$$(\mathcal{H} - E_{\mathrm{ref}})\Psi(\mathbf{R}, t) = -\frac{\partial \Psi(\mathbf{R}, t)}{\partial t}$$

$$\Psi(\mathbf{R}, t) = \int \mathrm{d}\mathbf{R}_0 \, G(\mathbf{R}, \mathbf{R}_0, t)\Psi^{(0)}(\mathbf{R}_0)$$

$G(\mathbf{R}', \mathbf{R}, t)$ satisfies the imaginary-time Schrödinger equation

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

with $G(\mathbf{R}', \mathbf{R}, t) = \langle \mathbf{R}'|e^{-t(\mathcal{H} - E_{\mathrm{ref}})}|\mathbf{R}\rangle$ and $G(\mathbf{R}', \mathbf{R}, 0) = \delta(\mathbf{R}' - \mathbf{R})$

Can we interpret $G(\mathbf{R}', \mathbf{R}, \tau)$ as a transition probability? (1)

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

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

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

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

The Green's function is given by

$$G(\mathbf{R}', \mathbf{R}, \tau) = \exp\left[-\tau \left(\mathcal{V}(\mathbf{R}) - E_{\text{ref}}\right)\right] \delta(\mathbf{R} - \mathbf{R}'),$$

$\boxed{\text{Positive}}$ 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 $\rightarrow$ $\boxed{e^{(A+B)\tau} = e^{A\tau}e^{B\tau} + \mathcal{O}(\tau^2)}$

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

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

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

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

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

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

$\rightarrow$ Diffusion + reweight factor leading to survival/death/cloning

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

Diffusion and branching in a harmonic potential

$V(x)$

$\Psi_0(x)$
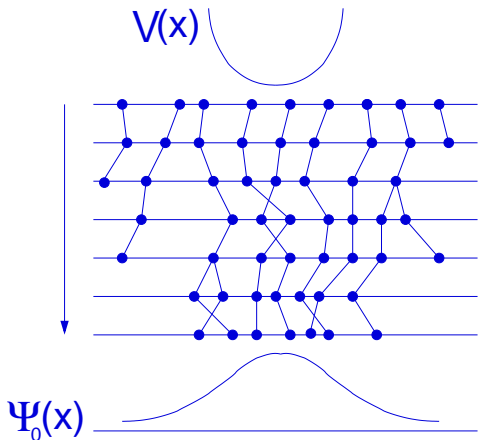
Walkers proliferate/die where potential is lower/higher than E_{ref}

Time-step extrapolation

Example: Energy of Li_2 versus time-step τ
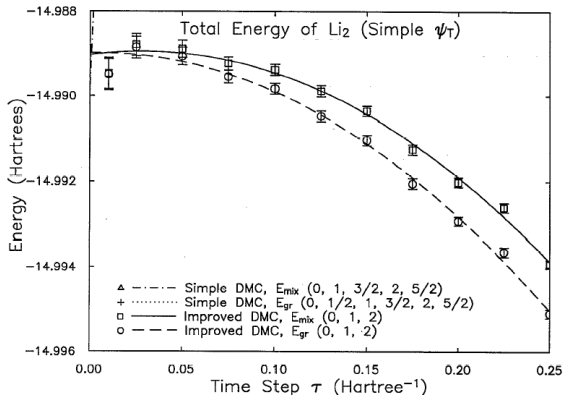
Umrigar, Nightingale, Runge, J. Chem. Phys. **94**, 2865 (1993)

Diffusion Monte Carlo as a branching random walk

The basic DMC algorithm is rather simple:

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

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

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

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

3. For each walker, compute the factor

   $$\boxed{p = \exp\left[-\tau(\mathcal{V}(\mathbf{R}) - E_{\mathrm{ref}})\right]}$$

   p is the probability to survive/proliferate/die

4. Adjust E_{ref} so that population fluctuates around target M_0

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

Problems with simple algorithm

The simple algorithm is inefficient and unstable

▷ Potential can vary a lot and be unbounded

  e.g. electron-nucleus interaction $\rightarrow$ Exploding population

▷ Branching factor grows with system size

Importance sampling

Start from integral equation

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

Multiply each side by trial Ψ_{T} and define $\boxed{\pi(\mathbf{R}, t) = \Psi_{\mathrm{T}}(\mathbf{R})\Psi(\mathbf{R}, t)}$

$$\pi(\mathbf{R}', t + \tau) = \int d\mathbf{R}\, \tilde{G}(\mathbf{R}', \mathbf{R}, \tau)\pi(\mathbf{R}, t)$$

where the importance sampled Green's function is

$$\boxed{\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = \Psi_{\mathrm{T}}(\mathbf{R}')\langle\mathbf{R}'|e^{-\tau(\mathcal{H} - E_{\mathrm{ref}})}|\mathbf{R}\rangle/\Psi_{\mathrm{T}}(\mathbf{R})}$$

We obtain $\boxed{\lim_{n \to \infty} \pi(\mathbf{R}) = \Psi_{\mathrm{T}}(\mathbf{R})\Psi_0(\mathbf{R})}$

Importance sampled Green's function

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

$$-\frac{1}{2}\nabla^2 \tilde{G} + \nabla \cdot [\tilde{G}\, \mathbf{V}(\mathbf{R})] + [E_{\mathrm{L}}(\mathbf{R}) - E_{\mathrm{ref}}]\, \tilde{G} = -\frac{\partial \tilde{G}}{\partial \tau}$$

with quantum velocity $\mathbf{V}(\mathbf{R}) = \dfrac{\nabla \Psi_{\mathrm{T}}(\mathbf{R})}{\Psi_{\mathrm{T}}(\mathbf{R})}$ and $E_{\mathrm{L}}(\mathbf{R}) = \dfrac{\mathcal{H}\Psi_{\mathrm{T}}(\mathbf{R})}{\Psi_{\mathrm{T}}(\mathbf{R})}$

We now have $\boxed{\text{drift}}$ in addition to diffusion and branching terms

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

The drift-diffusion-branching Green's function

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

$$\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R} - \tau\mathbf{V}(\mathbf{R}))^2}{2\tau}\right] \times$$
$$\times \exp\left\{-\tau\left(E_{\mathrm{L}}(\mathbf{R}) - E_{\mathrm{ref}}\right)\right\}$$

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

▷ $\mathbf{V}(\mathbf{R})$ pushes walkers where Ψ is large

▷ $E_{\mathrm{L}}(\mathbf{R})$ is better behaved than the potential $\mathcal{V}(\mathbf{R})$

  Cusp conditions $\Rightarrow$ No divergences when particles approach

  As $\Psi_{\mathrm{T}} \to \Psi_0$, $E_{\mathrm{L}} \to E_0$ and branching factor is smaller

Basic DMC algorithm with importance sampling

1. Sample initial walkers from $|\Psi_{\mathrm{T}}(\mathbf{R})|^2$

2. Drift and diffuse the walkers as $\mathbf{R}' = \mathbf{R} + \tau\mathbf{V}(\mathbf{R}) + \xi$
   where ξ is sampled from $g(\xi) = (2\pi\tau)^{-3N/2}\exp\left(-\xi^2/2\tau\right)$

3. Branching step as in the simple algorithm but with the factor

   $$p = \exp\left\{-\tau[(E_{\mathrm{L}}(\mathbf{R}) + E_{\mathrm{L}}(\mathbf{R}'))/2 - E_{\mathrm{ref}}]\right\}$$

4. Adjust the trial energy to keep the population stable

$\rightarrow$ After many iterations, walkers distributed as $\Psi_{\mathrm{T}}(\mathbf{R})\Psi_0(\mathbf{R})$

Electrons are fermions!

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

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

Fermion Sign Problem

All fermion QMC methods suffer from sign problems

These sign problems look different but have the same "flavour"

Arise when you treat something non-positive as probability density

The DMC Sign Problem

How can we impose antisymmetry in simple DMC method?

Idea Evolve separate positive and negative populations of walkers

Simple 1D example: Antisymmetric wave function $\Psi(x, \tau = 0)$

Rewrite $\Psi(x, \tau = 0)$ as

$$\Psi = \Psi_+ - \Psi_-$$

where

$$\Psi_+ = \frac{1}{2}(|\Psi| + \Psi)$$
$$\Psi_- = \frac{1}{2}(|\Psi| - \Psi)$$

$\Psi(x, \tau = 0)$

$\Psi_+(x, \tau = 0)$ $\Psi_-(x, \tau = 0)$
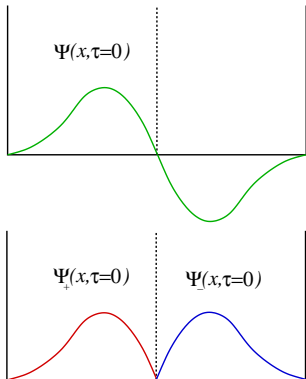
Particle in a box and the fermionic problem

The imaginary-time Schrödinger equation

$$\mathcal{H}\Psi = -\frac{\partial \Psi}{\partial t}$$

is linear, so solving it with the initial condition

$$\Psi(x, t = 0) = \Psi_+(x, t = 0) - \Psi_-(x, t = 0)$$

is equivalent to solving

$$\mathcal{H}\Psi_+ = -\frac{\partial \Psi_+}{\partial t} \qquad \text{and} \qquad \mathcal{H}\Psi_- = -\frac{\partial \Psi_-}{\partial t}$$

separately and subtracting one solution from the other

▷ Since $E_0^s < E_0^a$, both Ψ_+ and Ψ_- evolve to Ψ_0^s

$\Psi_\pm \longrightarrow$
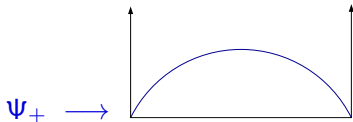
▷ Antisymmetric component exponentially harder to extract

$$\frac{|\Psi_+ - \Psi_-|}{|\Psi_+ + \Psi_-|} \quad \propto \quad \frac{e^{-E_0^a t}}{e^{-E_0^s t}} \quad \text{as} \quad t \to \infty$$

The Fixed-Node Approximation

Problem Small antisymmetric part swamped by random errors

Solution Fix the nodes! (If you don't know them, guess them)

impenetrable barrier

impenetrable barrier
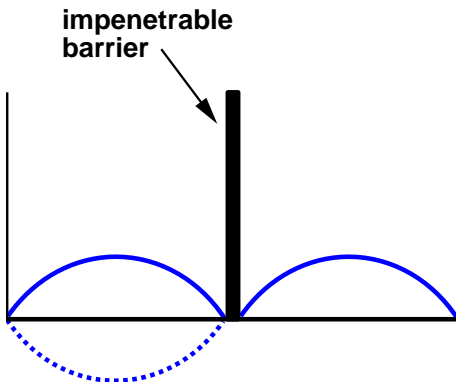
Fixed-node algorithm in simple DMC
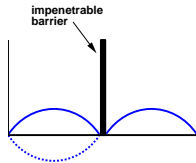
How do we impose this additional boundary condition?

▷ Annihilate walkers that bump into barrier (and into walls)

→ This step enforces $\boxed{\Psi = 0}$ boundary conditions

→ In each nodal pocket, evolution to ground state in pocket

Numerically $\boxed{\text{stable}}$ algorithm (no exponentially growing noise)

→ Solution is exact if nodes are exact

→ Best solution consistent with the assumed nodes

For many electrons, what are the nodes? A complex beast

Many-electron wave function $\Psi(\mathbf{R}) = \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$

Node $\rightarrow$ surface where $\Psi = 0$ and across which Ψ changes sign
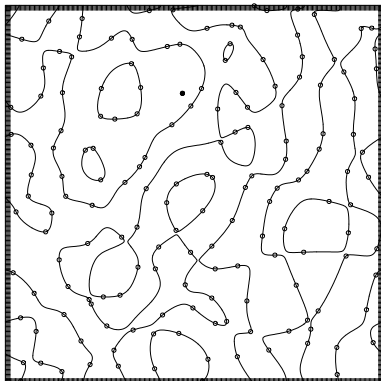
A 2D slice through the 321-dimensional nodal surface
of a gas of 161 spin-up electrons.

Some known properties of the nodes

Physical space has d (=1,2,3) dimensions

▶ Node is $(dN - 1)$-dimensional surface in dN dimensions

$\boxed{\text{One}}$ constraint ($\Psi = 0$) $\Rightarrow$ $\boxed{(dN - 1)}$-dimensional node

▶ Equations as $\mathbf{r}_i = \mathbf{r}_j$ define $(dN - d)$-dimensional <u>coincidence</u>
surfaces and do not define the node completely if $d > 1$

▶ If $d = 1$, coincidence points $x_i = x_j$ define the ground-state
node completely $\rightarrow$ One-dim problems are easy to simulate

Start from $\mathbf{R}_0$ and continously reach all points with $\Psi(\mathbf{R}) \neq 0$

$\Rightarrow$ Nodal pocket accessible from $\mathbf{R}_0$

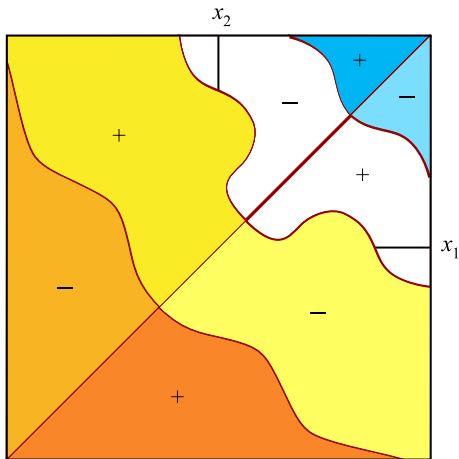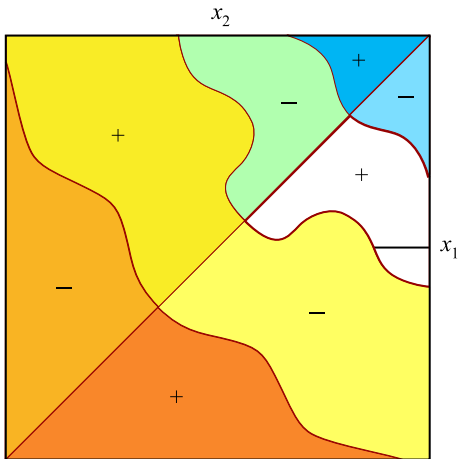Map this subvolume over rest of the space with permutations

Start from $\mathbf{R}_0$ and continuously reach all points with $\Psi(\mathbf{R}) \neq 0$

$\Rightarrow$ Nodal pocket accessible from $\mathbf{R}_0$

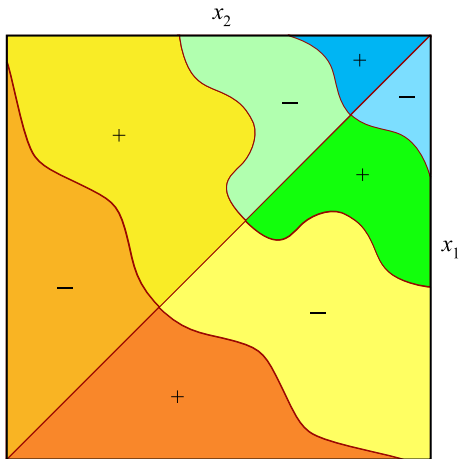Map this subvolume over rest of the space with permutations

Start from $\mathbf{R}_0$ and continously reach all points with $\Psi(\mathbf{R}) \neq 0$

$\Rightarrow$ Nodal pocket accessible from $\mathbf{R}_0$

Map this subvolume over rest of the space with permutations

Start from $\mathbf{R}_0$ and continously reach all points with $\Psi(\mathbf{R}) \neq 0$
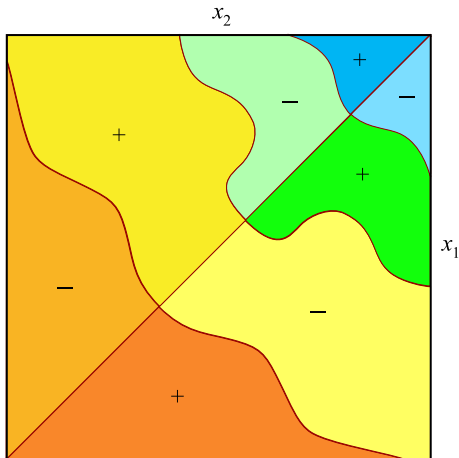
$\Rightarrow$ Nodal pocket accessible from $\mathbf{R}_0$

Map this subvolume over rest of the space with permutations

Start from $\mathbf{R}_0$ and continuously reach all points with $\Psi(\mathbf{R}) \neq 0$

$\Rightarrow$ Nodal pocket accessible from $\mathbf{R}_0$

Map this subvolume over rest of the space with permutations

Start from $\mathbf{R}_0$ and continously reach all points with $\Psi(\mathbf{R}) \neq 0$

$\Rightarrow$ Nodal pocket accessible from $\mathbf{R}_0$

Map this subvolume over rest of the space with permutations

Start from $\mathbf{R}_0$ and continously reach all points with $\Psi(\mathbf{R}) \neq 0$

$\Rightarrow$ Nodal pocket accessible from $\mathbf{R}_0$

Map this subvolume over rest of the space with permutations

Start from $\mathbf{R}_0$ and continously reach all points with $\Psi(\mathbf{R}) \neq 0$

$\Rightarrow$ Nodal pocket accessible from $\mathbf{R}_0$

Map this subvolume over rest of the space with permutations

Start from $\mathbf{R}_0$ and continously reach all points with $\Psi(\mathbf{R}) \neq 0$

$\Rightarrow$ Nodal pocket accessible from $\mathbf{R}_0$

Map this subvolume over rest of the space with permutations
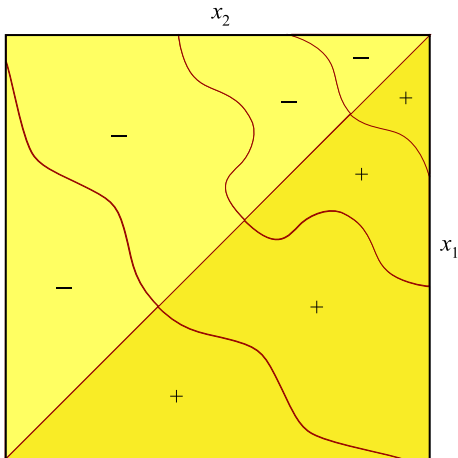
Consider Hamiltonian with a local potential

For ground-state wavefunction, all pockets are in the same class

Consider Hamiltonian with a local potential

For ground-state wavefunction, all pockets are in the same class

Consider Hamiltonian with a local potential

For ground-state wavefunction, all pockets are in the same class

Consider Hamiltonian with a local potential

For ground-state wavefunction, all pockets are in the same class

Consider Hamiltonian with a local potential

For ground-state wavefunction, all pockets are in the same class

Use the nodes of trial $\Psi_T \to$ Fixed-node approximation

Use the nodes of the best available trial Ψ_T wave function

$\Psi(R)=0$

$\Psi(R)>0$

R
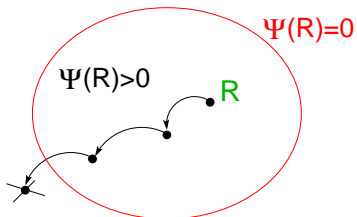
Find best solution with same nodes as trial wave function Ψ_T

Fixed-node solution exact if the nodes of trial Ψ_T are exact

Easy to implement in DMC with importance sampling: $\pi \geq 0$

Fixed-node solution and importance-sampling DMC

Given trial $\Psi_{\mathrm{T}}(\mathbf{R})$, evolve $\boxed{\pi(\mathbf{R}, t) = \Psi_{\mathrm{T}}(\mathbf{R})\Psi(\mathbf{R}, t)}$ as

$$-\frac{1}{2}\nabla^2\pi + \nabla \cdot [\pi\,\mathbf{V}(\mathbf{R})] + [E_{\mathrm{L}}(\mathbf{R}) - E_{\mathrm{ref}}]\,\pi = -\frac{\partial \pi}{\partial \tau}$$

with $\mathbf{V}(\mathbf{R}) = \dfrac{\nabla\Psi_{\mathrm{T}}(\mathbf{R})}{\Psi_{\mathrm{T}}(\mathbf{R})}$ and $E_{\mathrm{L}}(\mathbf{R}) = \dfrac{\mathcal{H}\Psi_{\mathrm{T}}(\mathbf{R})}{\Psi_{\mathrm{T}}(\mathbf{R})}$

Fixed-node approximation $\rightarrow$ $\boxed{\pi(\mathbf{R}, t) \geq 0}$

Have we solved all our problems?

Results depend on the nodes of the trail wave function Ψ

Diffusion Monte Carlo as a black-box approach?

ϵ_{MAD} for atomization energy of the G1 set

	DMC			CCSD(T)/aug-cc-pVQZ	
	HF orb	Optimized orb	CAS		
ϵ_{MAD}	3.1	2.1	1.2	2.8	kcal/mol

Petruzielo, Toulouse, Umrigar, J. Chem. Phys. **136**, 124116 (2012)

With "some" effort on Ψ, we can do rather well

Diffusion Monte Carlo as a black-box approach?

Non-covalent interaction energies for 9 compounds from S22 set

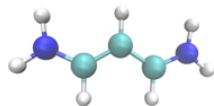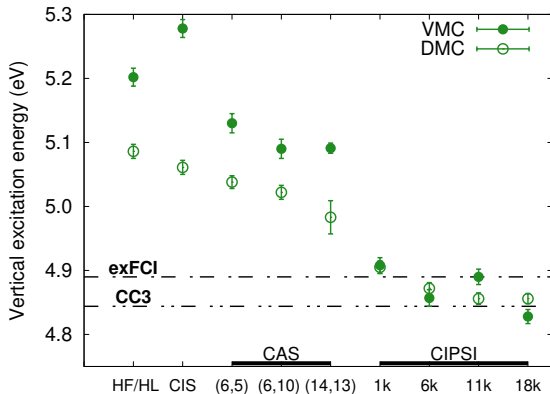DMC with B3LYP/aug-cc-PVTZ orbitals versus CCSD(T)/CBS
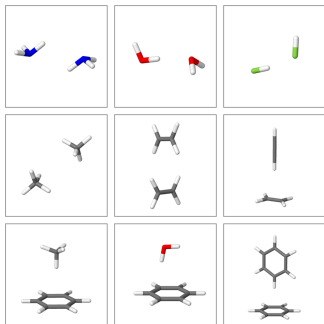
$$\Delta_{\mathrm{MAD}} = 0.058 \text{ kcal/mol}$$

Dubecky *et al.*, JCTC **9**, 4287 (2013)

With "practically no" effort on Ψ, we can do rather well

Diffusion Monte Carlo end excitation energy

Excitation energy for different choices of wave function

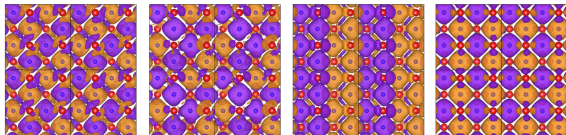Cuzzocrea, Scemama, Briels, Moroni, Filippi, JCTC **16**, 4203 (2020)

DMC is not a panacea but effort on Ψ pays off!

DMC and solid state calculations

Example: Structural/magnetic properties of superconducting FeSe

$\rightarrow$ Accurate lattice constants, bulk moduli, and band dispersion

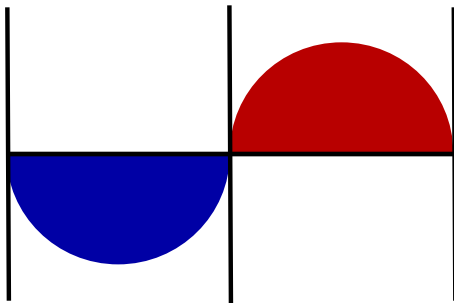$\rightarrow$ Resolving relative energetics of different magnetic ordering

Busemeyer, Dagrada, Sorella, Casula, and Wagner PRB (2016)
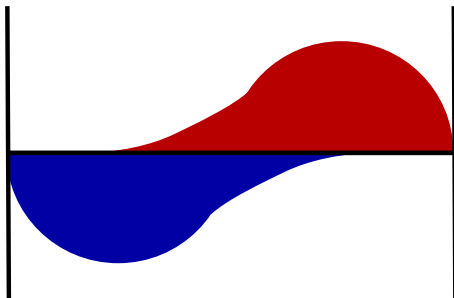
First do a fixed-node DMC simulation

Then release the nodes

Then release the nodes

- ▶ Red and blue solutions collapse to boson ground state, but their difference approaches the fermion ground state
- ▶ Back to the sign problem: exponentially growing noise

Given single-particle basis, perform projection in determinant space

Different way to deal with fermionic problem

– Determinantal QMC by Zhang and Krakauer

  Appears less plagued by fixed phase than DMC by FN

– Full-CI QMC by Alavi

  Start from $\Psi_{\mathrm{CI}} = \sum_i c_i D_i$

$$\mathcal{H}\Psi = -\frac{\partial \Psi}{\partial t} \rightarrow H_{ij} c_j = -\frac{\partial c_i}{\partial t}$$

The fixed-node DMC method is (in general)

▶ Easy to do

▶ Stable

▶ Accurate enough for many applications in quantum chemistry
... especially in large systems

▶ Accurate enough also for subtle correlation physics

Use of fixed-node DMC for computation of excited states

▶ In the general landscape, we are not doing too badly

▶ Sensitivity to wave function but relatively robust

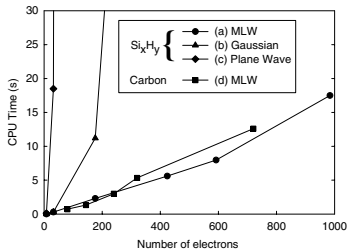$\rightarrow$ basis, choice of CAS (minimal CAS), truncation on CSFs

Beauty of quantum Monte Carlo $\rightarrow$ Highly parallelizable

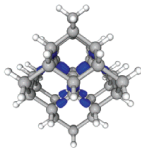$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) \rightarrow$ Ensemble of walkers diffusing in 3N dimensions

VMC $\rightarrow$ Independent walkers $\Rightarrow$ Trival parallelization

DMC $\rightarrow$ Nearly independent walkers $\Rightarrow$ Few communications

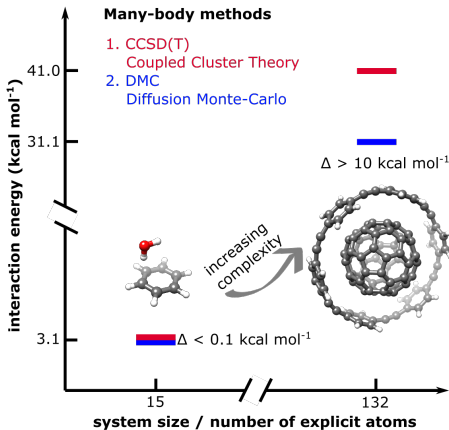Easily take great advantage of parallel supercomputers!

Williamson, Hood, Grossman (2001)

Up to $Si_{123}H_{100}$ and C_{180} !

Going to larger systems pose new problems

Many-body methods

1. CCSD(T)
   Coupled Cluster Theory
2. DMC
   Diffusion Monte-Carlo

$\Delta > 10$ kcal mol^{-1}

interaction energy (kcal mol^{-1})

41.0

31.1

3.1

$\Delta < 0.1$ kcal mol^{-1}

increasing complexity

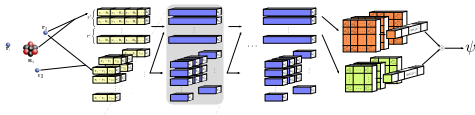15   132

system size / number of explicit atoms

Y.S. Al-Hamdani *et al.* Nature Commun. (2021)

To conclude: ongoing research in QMC

▶ Search for different forms of trial wave function

   Neural network architecture $\rightarrow \Psi$ of multi-electron orbitals

   Pfau, Spencer, Matthews, Foulkes, Phys. Rev. Res. (2020)

▶ Push optimization techniques to larger systems
▶ More work on transition metals
▶ Alternatives to fixed-node diffusion Monte Carlo

Other applications of quantum Monte Carlo methods

► Electronic structure calculations

► Strongly correlated systems (Hubbard, t-J, ...)

► Quantum spin systems (Ising, Heisenberg, XY, ...)

► Liquid-solid helium, liquid-solid interface, droplets

► Atomic clusters

► Nuclear structure

► Lattice gauge theory

Both zero (ground state) and finite temperature

Hands-on tutorial to learn real-space VMC and DMC

https://trex-coe.github.io/qmc-lttc-2023/index.html

Some references on VMC and DMC

- W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Quantum Monte Carlo simulations of solids, Rev. Mod. Phys. 73, 33-83 (2001).

- A. Luchow, Quantum Monte Carlo methods. Wiley Interdiscip. Rev.: Comput. Mol. Sci. 1, 388-402 (2011).

- B. M. Austin, D. Y. Zubarev, W. A. Lester, Quantum Monte Carlo and Related Approaches. Chem. Rev. 112, 263-288 (2012).

- J. Toulouse, R. Assaraf, and C. J. Umrigar, Introduction to the Variational and Diffusion Monte Carlo Methods, Adv. Quantum Chem. 73, 285-314 (2016).

- J. Feldt and C. Filippi, Excited-state calculations with quantum Monte Carlo in "Excited states: Methods for quantum chemistry and dynamics", edited by R. Lindh and L. Gonzalez (Wiley, 2020).

- F. Becca and S. Sorella, Quantum Monte Carlo Approaches for Correlated Systems (Cambridge University Press, 2017).

Some references on determinantal space QMC

- ▶ S. Zhang, Auxiliary-Field Quantum Monte Carlo at Zero- and Finite-Temperature, in "Many-Body Methods for Real Materials Modeling and Simulation Vol. 9" edited by E. Pavarini, E. Koch, and S. Zhang, Forschungszentrum Jülich, (2019).

- ▶ G. H. Booth, A. J. W. Thom, and A. Alavi, Fermion Monte Carlo without fixed nodes: a game of life, death, and annihilation in Slater determinant space, J. Chem. Phys. 131, 054106 (2009).

- ▶ N. S. Blunt, S. D. Smart, J. A. F. Kersten, J. S. Spencer, G. H. Booth, and A. Alavi, Semi-stochastic full configuration interaction quantum Monte Carlo: developments and application, J. Chem. Phys. 142, 184107 (2015).