Introduction to quantum Monte Carlo methods

Part 1

Claudia Filippi

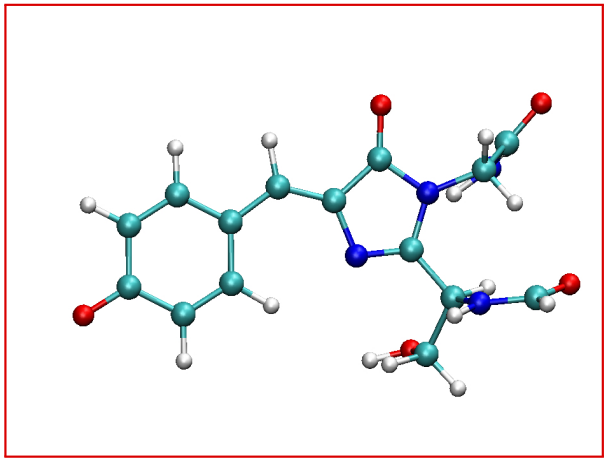
MESA+ Institute for Nanotechnology, Universiteit Twente, The Netherlands

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

Targeting Real Chemical Accuracy at the Exascale project has received funding from the European Union Horizoon 2020 research and innovation programme under Grant Agreement **No. 952165.**

A quick reminder: What is electronic structure theory?

A quantum mechanical and first-principle approach



$\longrightarrow$ Collection of ions + electrons

$\downarrow$

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 $\dfrac{\langle \Psi_n|\mathcal{O}|\Psi_n\rangle}{\langle \Psi_n|\Psi_n\rangle}$

Where is the difficulty?

Electron-electron interaction $\rightarrow$ Non-separable

Is there an optimal electronic structure approach?

- Density functional theory methods

  Large systems but approximate exchange/correlation

  Beyond? Green's function approaches

  Alternatives?

- Quantum chemistry post-Hartree-Fock methods

  Accurate on small-medium systems

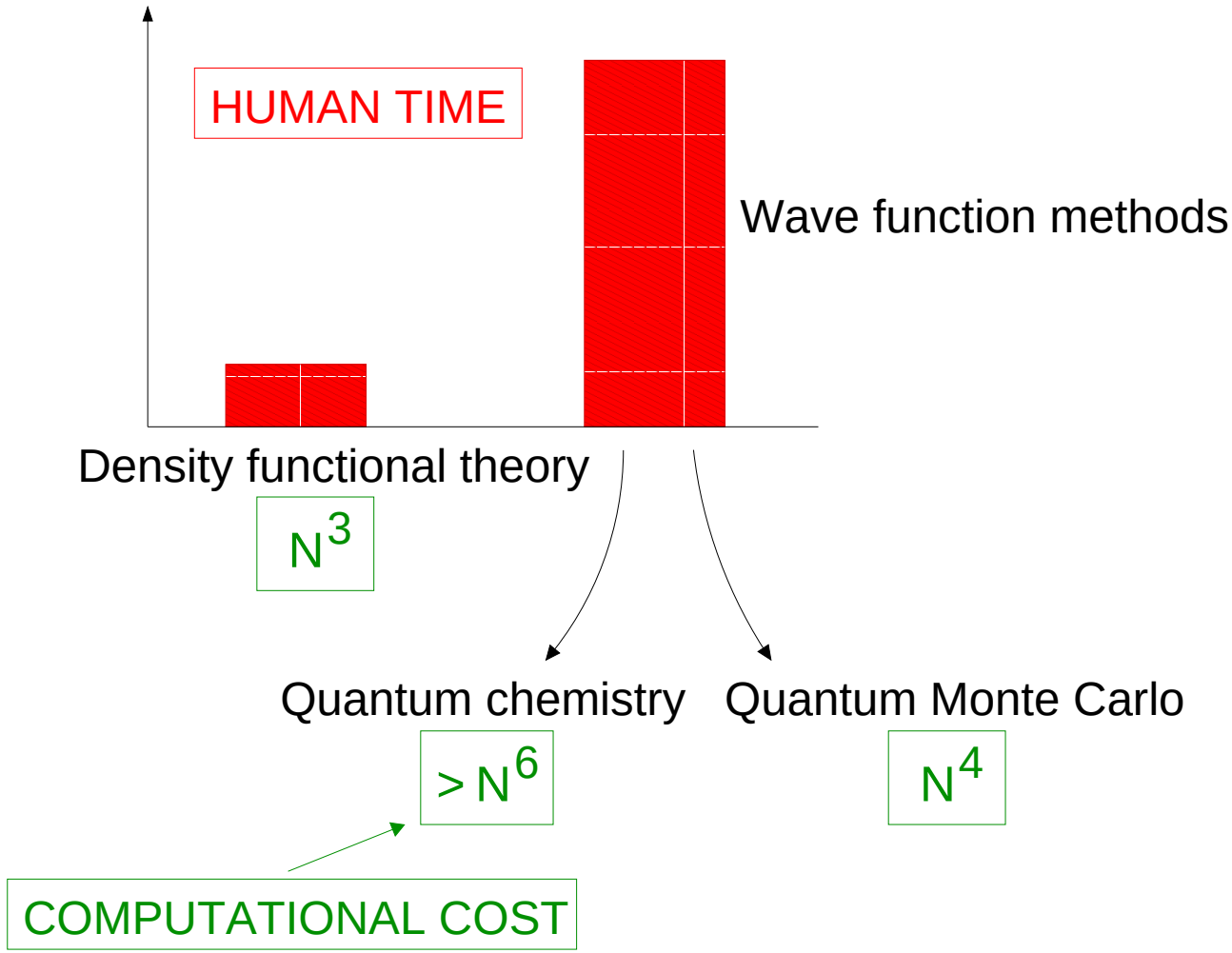
  $\rightarrow$ Jungle of approaches: CI, MCSCF, CC, CASPT2 ...

- Quantum Monte Carlo techniques

  Stochastic solution of the Schrödinger equation

  Accurate correlated calculations for medium-large systems

If you can, use density functional theory!



Density functional theory is cheap(er) and powerful but ...

Many successful stories of DFT + efficient, user-friendly codes

Are we theoreticians out of job? Can anybody do it?

Better posed questions

Is it always a success story?

Do we have a black-box method close to perfection?

In principle $\longrightarrow$ DFT is correct

BUT $\quad E_{xc}[\rho]$ unknown functional of the density

In practice $\longrightarrow$ $E_{xc}[\rho]$ must be approximated

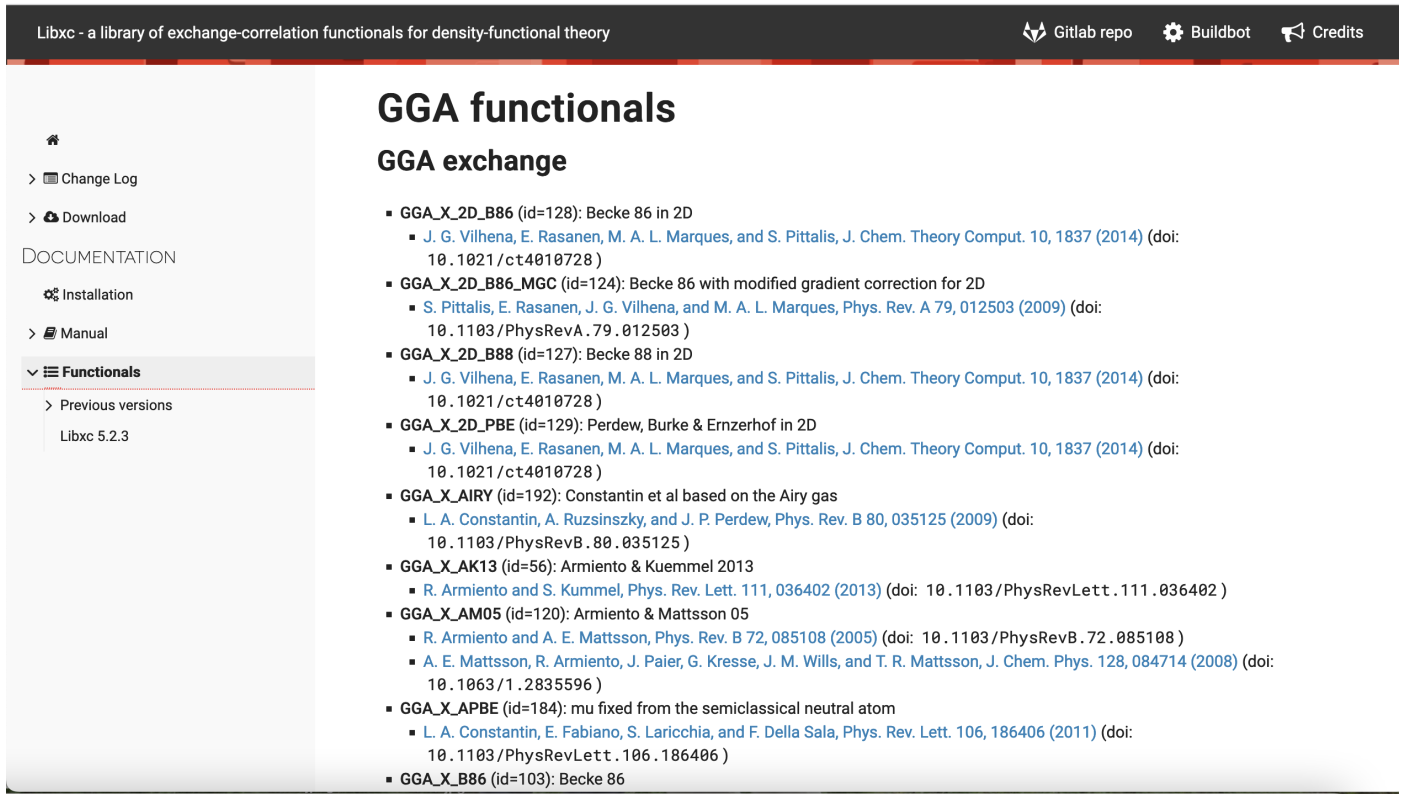
... and sometimes things go wrong

Some open problems

Weakly bound, strongly correlated systems, ...

Excitations with charge-transfer/multi-configurational character ...

→ Effort in development of exchange-correlation functionals



Some open problems

Weakly bound, strongly correlated systems, ...

Excitations with charge-transfer/multi-configurational character ...

→ Effort in development of exchange-correlation functionals

nature > nature communications > articles > article

Article | Open Access | Published: 14 July 2020

Machine learning accurate exchange and correlation functionals of the electronic density

Sebastian Dick & Marivi Fernandez-Serra

Nature Communications **11**, Article number: 3509 (2020) | Cite this article

8536 Accesses | **54** Citations | **28** Altmetric | Metrics

Abstract

Density functional theory (DFT) is the standard formalism to study the electronic structure of matter at the atomic scale. In Kohn–Sham DFT simulations, the balance between accuracy and computational cost depends on the choice of exchange and correlation functional, which only exists in approximate form. Here, we propose a framework to create density functionals using supervised machine learning, termed NeuralXC. These machine-learned functionals are designed to lift the accuracy of baseline functionals towards that provided by more accurate methods while maintaining their efficiency. We show that the functionals

When DFT has problems $\rightarrow$ Wave function based methods

Work in the Born-Oppenheimer approximation

Solve the Schrödinger equation for the electrons in the ionic field

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

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 $\sigma = \pm 1$

When DFT has problems $\rightarrow$ Wave function based methods

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

Optimal wave functions and the variational theorem

$\Psi(\mathbf{X}, a)$ with $\mathbf{X}$ the space-spin variables and $\boxed{a}$ the parameters

$$E_{\mathrm{V}}(a) = \frac{\langle \Psi(a) | \mathcal{H} | \Psi(a) \rangle}{\langle \Psi(a) | \Psi(a) \rangle} \geq E_0$$

$$E_{\mathrm{V}}(a) = E_0 \quad \Leftrightarrow \quad \Psi(\mathbf{X}, a) = \Psi_0(\mathbf{X})$$

The variational method and the linear basis approach

Wave function as a linear combination of basis functions $f_n(\mathbf{X})$

$$\boxed{\Psi(\mathbf{X}, a) = \sum_n a_n f_n(\mathbf{X})} \;\Rightarrow\; E_{\mathrm{V}}(a) = \frac{\sum_{n,m} a_n^* a_m H_{nm}}{\sum_{n,m} a_n^* a_m S_{nm}}$$

where $H_{nm} = \langle f_n | \mathcal{H} | f_m \rangle$ and $S_{nm} = \langle f_n | f_m \rangle$

$$\frac{\mathrm{d}E}{\mathrm{d}a} = 2\left[Ha - E_v Sa\right] = 0 \;\Rightarrow\; \boxed{Ha_\lambda = E_\lambda S a_\lambda}$$

Linear basis approach $\rightarrow$ Generalized eigenvalue problem

Ψ as a linear combination of many-body functions $f_n(\mathbf{X})$

$$\Psi(\mathbf{X}, a) = \sum_n a_n f_n(\mathbf{X}) \;\Rightarrow\; \boxed{Ha_\lambda = E_\lambda S a_\lambda}$$

Important properties

- For a basis of size M, $\exists$ M eigenvalues and eigenfunctions
- McDonald's theorem

$$E_1 \leq E_2 \leq \ldots \leq E_M \quad \text{with} \quad \boxed{E_n^{\text{exact}} \leq E_n}$$

Merits and problems of the variational method

Find approximate solution Ψ to Schrödinger equation

Merits

- Upper bound is guaranteed
- Linear basis $\rightarrow$ Generalized eigenvalue problem
- Linear basis $\rightarrow$ McDonald's theorem for excited states

Problems

- How do we compute the matrix elements H_{nm} and S_{nm}?
- How do we access convergence?
- What goes in, comes out

How do we compute the matrix elements S_{nm} and H_{nm}?

Integrals H_{nm} and S_{nm} too slow to perform unless one-particle basis

$\rightarrow$ Problem which can be solved by Monte Carlo integration

Many-body wave functions in traditional quantum chemistry

Interacting $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) \leftrightarrow$ Non-interacting basis $\psi(\mathbf{x})$

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

Single-particle orbitals expanded in Gaussian basis

$\Rightarrow$ All integrals can be computed analytically

Determinants as a non-interacting basis (1)

Interested in interacting electron system with full Hamiltonian

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

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

Much easier problem $\rightarrow$ $\mathcal{H}^{\text{non-int}} = h^{(1)}(\mathbf{r}_1) + \ldots + h^{(1)}(\mathbf{r}_N)$

e.g. Hartree-Fock or Kohn-Sham equations

Determinants as a non-interacting basis

If $\mathcal{H}^{\mathrm{non-int}} = h^{(1)}(\mathbf{r}_1) + \ldots h^{(1)}(\mathbf{r}_N)$, proceed as follows:

If $\mathcal{H}^{\mathrm{non-int}} = h^{(1)}(\mathbf{r}_1) + \ldots h^{(1)}(\mathbf{r}_N)$, proceed as follows:

- Solve for one-electron and build spin-orbitals

$$h^{(1)}(\mathbf{r}_1)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \quad \rightarrow \quad \psi_i(\mathbf{x}) = \phi_i(\mathbf{r})\chi_{s_i}(\sigma)$$

If $\mathcal{H}^{\rm non-int} = h^{(1)}(\mathbf{r}_1) + \ldots h^{(1)}(\mathbf{r}_N)$, proceed as follows:

- Solve for one-electron and build spin-orbitals

$$h^{(1)}(\mathbf{r}_1)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \quad \rightarrow \quad \psi_i(\mathbf{x}) = \phi_i(\mathbf{r})\chi_{s_i}(\sigma)$$

- Create a product state by occupying N spin-orbitals

$$\Psi^{\rm non-int}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \psi_1(\mathbf{x}_1)\ldots\psi_N(\mathbf{x}_N)$$

$$\rightarrow \quad \mathcal{H}^{\rm non-int}\Psi^{\rm non-int} = \left(\sum_{i=1}^{N}\epsilon_i\right)\Psi^{\rm non-int}$$

If $\mathcal{H}^{\rm non-int} = h^{(1)}(\mathbf{r}_1) + \ldots h^{(1)}(\mathbf{r}_N)$, proceed as follows:

- Solve for one-electron and build spin-orbitals

$$h^{(1)}(\mathbf{r}_1)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \quad \rightarrow \quad \psi_i(\mathbf{x}) = \phi_i(\mathbf{r})\chi_{s_i}(\sigma)$$

- Create a product state by occupying N spin-orbitals

$$\Psi^{\rm non-int}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \psi_1(\mathbf{x}_1)\ldots\psi_N(\mathbf{x}_N)$$

$$\rightarrow \quad \mathcal{H}^{\rm non-int}\Psi^{\rm non-int} = \left(\sum_{i=1}^{N}\epsilon_i\right)\Psi^{\rm non-int}$$

- Anti-symmetrize as a Slater determinant (non-int solution)

$$D(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \mathcal{A}\{\psi_1(\mathbf{x}_1)\ldots\psi_N(\mathbf{x}_N)\} \rightarrow \mathcal{H}^{\rm non-int}D = \left(\sum_{i=1}^{N}\epsilon_i\right)D$$

Determinants as a non-interacting basis (3)

Starting point $\rightarrow$ Non-interacting Hartree-Fock wave function

$$D_{\mathrm{HF}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \begin{vmatrix} \psi_1(\mathbf{x}_1) & \ldots & \psi_1(\mathbf{x}_N) \\ \vdots & & \vdots \\ \psi_N(\mathbf{x}_1) & \ldots & \psi_N(\mathbf{x}_N) \end{vmatrix}$$

Optimal spin-orbitals $\psi_i(\mathbf{x}) = \phi_i(\mathbf{r})\chi_{s_i}(\sigma)$ satisfy HF equations

$$\left[-\frac{1}{2}\nabla^2 + v_{\mathrm{ext}}(\mathbf{r}) + \sum_{j=1}^{N} \int \mathrm{d}\mathbf{r}' \, \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \right] \phi_i(\mathbf{r}) + [\hat{v}_{\mathrm{HF}}\phi_i](\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r})$$

$\Rightarrow$ occupied orbitals $(\psi_1 \ldots \psi_N)$ + virtual orbitals $(\psi_{\mathrm{N+1}} \ldots)$

Many-body wave functions in traditional quantum chemistry (1)

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

Expansion in linear combination of determinants

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

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

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

by constructing single, double, ... up to N-body excitations

Many-body wave functions in traditional quantum chemistry (2)

Pros and cons of CI expansion in Slater determinants

$$\Psi_{\mathrm{CI}} = c_0 D_{\mathrm{HF}} + \sum_{ab} c_{a\to b} D^{a\to b} + \sum_{abcd} c_{ab\to cd} D^{ab\to cd} + \ldots$$

Optimal CI coefficients by solving generalized eigenvalue equation

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

Orbitals on a Gaussian basis $\rightarrow$ **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?

$\rightarrow$ Real-space Monte Carlo methods in quantum mechanics

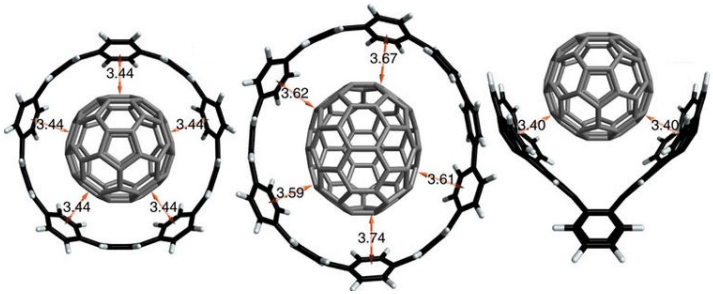
Some general words about quantum Monte Carlo methods

Stochastically solve interacting Schrödinger equation

Why (real-space) quantum Monte Carlo?

- Favorable scaling $\rightarrow$ Energy is $\boxed{O(N^4)}$
- Flexibility in choice of functional form of wave function
- Easy parallelization
- Among most accurate calculations for medium-large systems

  Routinely, molecules of up to 100 (mainly 1st/2nd-row) atoms



upto $C_{136}H_{44}$ (Alfé 2017)

Monte Carlo methods in general

Approaches which make repeated use of random numbers:

- to simulate truly stochastic events
- to solve **deterministic problems** using probabilities

Very important class of methods in statistical mechanics

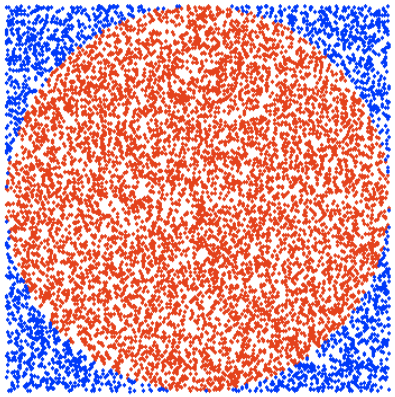
$\rightarrow$ Sampling Boltzmann distribution

Computation of averages (integrals in many dimensions)

For quantum mechanical simulations $\rightarrow$ **Quantum Monte Carlo**

A simple example of a Monte Carlo simulation

Basic idea of Monte Carlo through the "dartboard method"



$\rightarrow$ Throw darts, compute A_{circle}, compute π

Throw darts which land randomly within the square

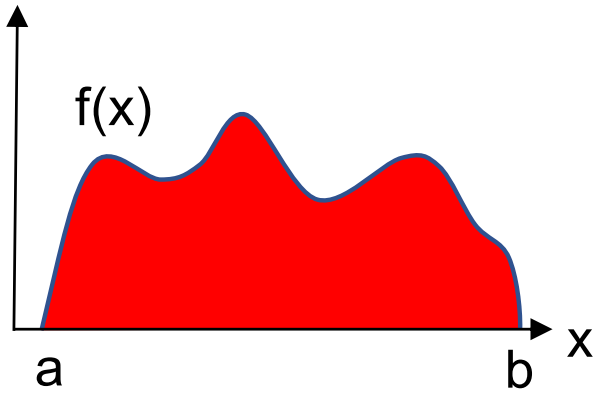
$$\frac{\#\text{ hits inside circle}}{\#\text{ hits inside the square}} = \frac{A_{\text{circle}}}{A_{\text{square}}} = \frac{\pi}{4}$$

$\uparrow$

many, many hits

Monte Carlo integration

We want to compute the integral of f(x) in the interval $[a, b]$



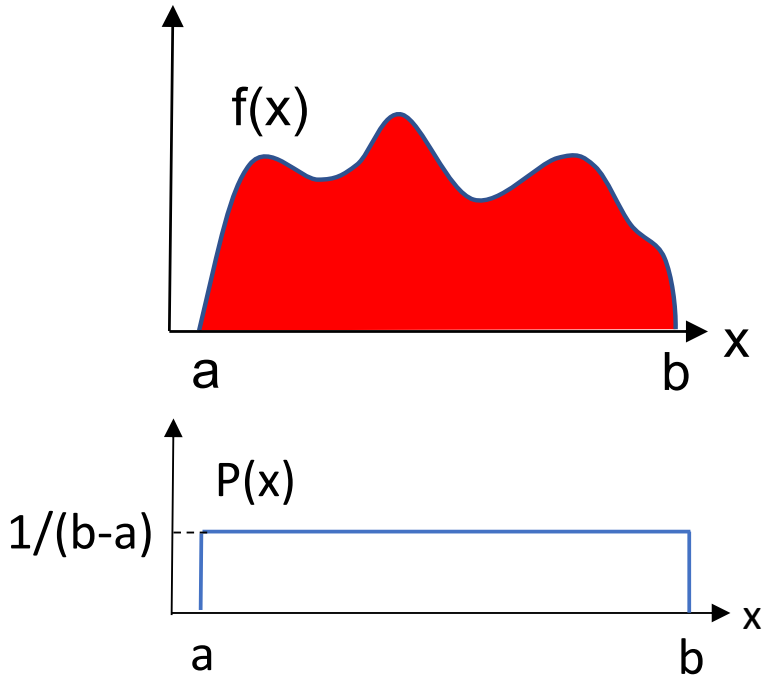
$$
\begin{aligned}
I &= \int_a^b f(x)\,\mathrm{d}x = (b-a)\int_a^b f(x)\frac{1}{b-a}\,\mathrm{d}x \\
&= (b-a)\langle f\rangle_{[a,b]}
\end{aligned}
$$

where $\langle f\rangle_{[a,b]}$ is the average of the function in the range $[a, b]$

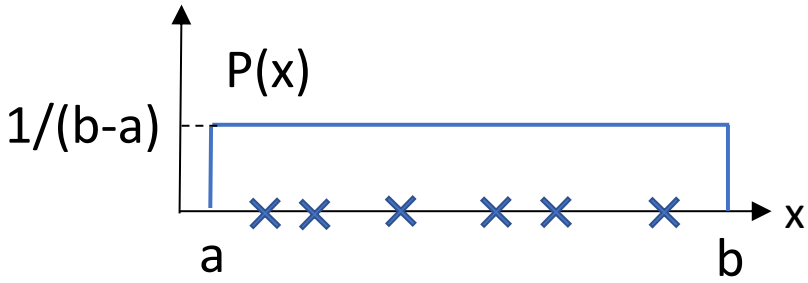
Monte Carlo integration

$$\langle f \rangle_{[a,b]} = \int_a^b f(x) \frac{1}{b-a} \, \mathrm{d}x$$

$$= \int_a^b f(x) P(x) \, \mathrm{d}x$$



Draw M random numbers distributed unformely in $[a, b]$

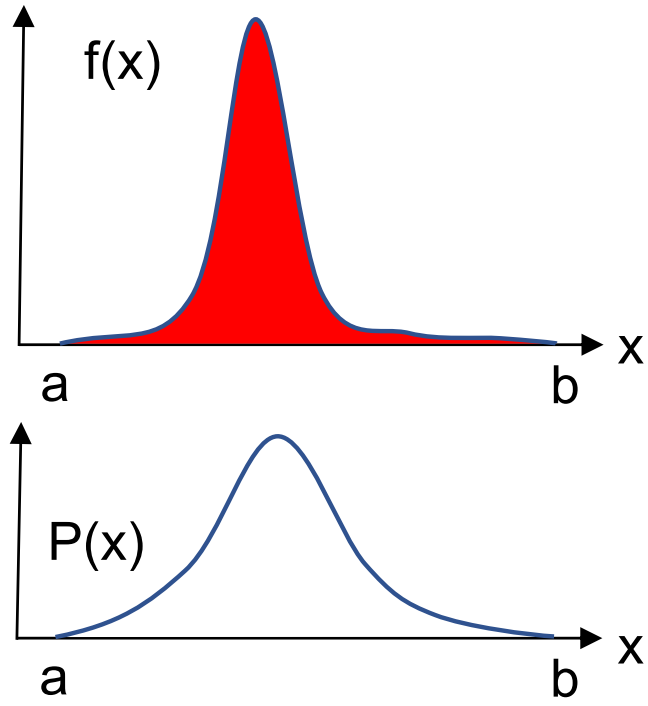


$$\longrightarrow \quad \langle f \rangle_{[a,b]} \approx \frac{1}{M} \sum_{i=1}^{M} f(x_i)$$

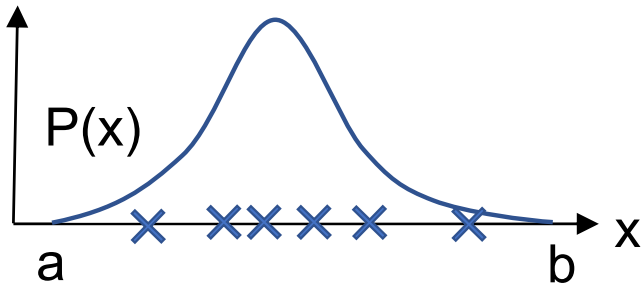
A less uniform function

$$I = \int_a^b f(x)\,\mathrm{d}x$$

$$= \int_a^b \frac{f(x)}{P(x)} P(x)\,\mathrm{d}x$$



Draw M random numbers distributed as $P(x)$



$$\longrightarrow \quad I \approx \frac{1}{M} \sum_{i=1}^{M} \frac{f(x_i)}{P(x_i)}$$

Monte Carlo integration in a nutshell

We want to compute $\boxed{\langle A \rangle = \int_a^b A(x)P(x)}$

with $\boxed{P(x) \geq 0 \text{ and } \int_a^b P(x) = 1}$ $\leftarrow$ a probability density!

Monte Carlo $\rightarrow$ Sample $\{x_1, \ldots, x_M\}$ from $P(x)$

$$\text{Estimate} \quad \langle A \rangle \approx \frac{1}{M} \sum_{i=1}^{M} A(x_i)$$

Statistical physics: $P(x) = \dfrac{e^{-\beta E(x)}}{Z}$, the Boltzman distribution

A different way of writing the expectation values

Back to quantum mechanics!

Consider the expectation value of the Hamiltonian on Ψ

$$E_V = \frac{\langle\Psi|\mathcal{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\int \mathrm{d}\mathbf{R}\, \Psi^*(\mathbf{R})\mathcal{H}\Psi(\mathbf{R})}{\int \mathrm{d}\mathbf{R}\, \Psi^*(\mathbf{R})\Psi(\mathbf{R})} \geq E_0$$

$$= \int \mathrm{d}\mathbf{R}\, \frac{\mathcal{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \boxed{\frac{|\Psi(\mathbf{R})|^2}{\int \mathrm{d}\mathbf{R}|\Psi(\mathbf{R})|^2}}$$

$$= \int \mathrm{d}\mathbf{R}\, E_{\mathrm{L}}(\mathbf{R})\, P(\mathbf{R}) = \langle E_{\mathrm{L}}(\mathbf{R})\rangle_P$$

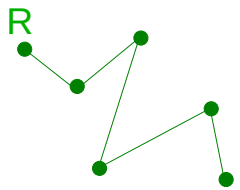
$P(\mathbf{R})$ is a probability density and $E_{\mathrm{L}}(\mathbf{R}) = \dfrac{\mathcal{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$ the local energy

Variational Monte Carlo: a random walk of the electrons

Use Monte Carlo integration to compute expectation values

- ▷ Sample $\mathbf{R}$ from $P(\mathbf{R})$ using Metropolis algorithm
- ▷ Average local energy $E_{\mathrm{L}}(\mathbf{R}) = \dfrac{\mathcal{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$ to obtain E_V as

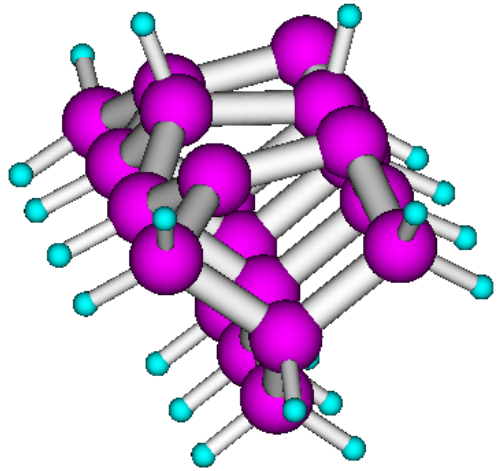
$$E_V = \langle E_{\mathrm{L}}(\mathbf{R}) \rangle_P \approx \frac{1}{M} \sum_{i=1}^{M} E_{\mathrm{L}}(\mathbf{R}_i)$$



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

Just a trick to evaluate integrals in many dimensions

Is it really "just" a trick?



$Si_{21}H_{22}$

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

Number of dimensions $3 \times 106 = \boxed{318}$

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

Monte Carlo integration

We want to compute an integral

$$E_V = \int \mathrm{d}\mathbf{R} E_\mathrm{L}(\mathbf{R}) P(\mathbf{R})$$

We sample $P(\mathbf{R}) \rightarrow$ $\boxed{E_V = \langle E_\mathrm{L}(\mathbf{R}) \rangle_P \approx \frac{1}{M} \sum_{i=1}^{M} E_\mathrm{L}(\mathbf{R}_i)}$

– Does the trick always work?

– How efficient is it?

The Central Limit Theorem

Probability density P and function f with finite mean and variance

$$\boxed{\mu} = \int \mathrm{d}x\, f(x)P(x) \qquad \boxed{\sigma^2} = \int \mathrm{d}x\, (f(x)-\mu)^2 P(x)$$

Sample M independent random variables $x_1, \ldots, x_M$ from $P(x)$

Define $$\boxed{F_M = \frac{1}{M}\sum_{i=1}^{M} f(x_i)}$$

As M increases, F_M is normally distributed as $\frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma_M^2}$

with a mean $\boxed{\mu}$ and variance $\boxed{\sigma_M^2 = \sigma^2/\mathrm{M}}$

$\rightarrow$ $\boxed{\text{Irrespective}}$ of the original probability density function

Compare to deterministic integration

Consider 1-dim integral over (a, b) and $M_{\rm int}$ integration points

If we use for instance **Simpson 1/3 rule** to perform an integral,

$$\int_a^b f(x)\,dx = \frac{1}{3}\left[f(x_a) + 4\sum_{i=1}^{M_{\rm int}/2} f(x_{2i-1}) + 2\sum_{i=1}^{M_{\rm int}/2-1} f(x_{2i}) + f(x_b)\right] + \epsilon(h^4)$$

with ϵ the integration error and

$$h = \frac{(b-a)}{M_{\rm int}} \quad \rightarrow \quad \boxed{\epsilon \propto \frac{1}{M_{\rm int}^4}}$$

What about integration on a grid in higher dimensions?

Monte Carlo versus deterministic integration

Integration error ϵ using $M_{\rm int}$ integration/$M_{\rm MC}$ Monte Carlo points

– Monte Carlo methods

$$\boxed{\epsilon \propto \frac{1}{\sqrt{M_{\rm MC}}}}$$ independent on dimension !

It follows from Central Limit Theorem

$\rightarrow$ width of Gaussian decreases as $\frac{\sigma}{\sqrt{M_{\rm MC}}}$ for finite variance

– Deterministic integration methods

1-dim Simpson rule: $\epsilon \propto \frac{1}{M_{\rm int}^4}$

d-dim Simpson rule: $\boxed{\epsilon \propto \frac{1}{M_{\rm int}^{4/d}}}$ $\leftarrow$ $M_{\rm int}$ grows as $\boxed{M_{{\rm int},1{\rm dim}}^d}$

Scaling with number of electrons

Roughly, Monte Carlo integration advantageous if $\boxed{d > 8}$

... for many-body wave functions $\boxed{d = 3N_{\mathrm{elec}}}$!

Assume that we want to obtain a given error ϵ

– Simpson rule integration ($\mathrm{M}_{\mathrm{int}}$ integration points)

$$\epsilon = \frac{c}{M_{\mathrm{int}}^{4/d}} = \frac{c}{M_{\mathrm{int}}^{4/3N_{\mathrm{elec}}}} \quad \Rightarrow \quad M_{\mathrm{int}} = \left(\frac{c}{\epsilon}\right)^{3N_{\mathrm{elec}}/4}$$

Exponential

– Monte Carlo integration (M_{MC} Monte Carlo samples)

$$\epsilon = \frac{\sigma}{\sqrt{M_{\mathrm{MC}}}} = c\sqrt{\frac{N_{\mathrm{elec}}}{M_{\mathrm{MC}}}} \quad \Rightarrow \quad M_{\mathrm{MC}} = \left(\frac{c}{\epsilon}\right)^2 N_{\mathrm{elec}}$$

Linear

Summary of variational Monte Carlo

Expectation value of the Hamiltonian on Ψ

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

$$\begin{aligned} E_V &= \int \mathrm{d}\mathbf{R}\, E_\mathrm{L}(\mathbf{R})\, P(\mathbf{R}) \\ \sigma^2 &= \int \mathrm{d}\mathbf{R} (E_\mathrm{L}(\mathbf{R}) - E_V)^2 P(\mathbf{R}) \end{aligned}$$

Estimate E_V and σ from M independent samples from $P(\mathbf{R})$ as

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

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

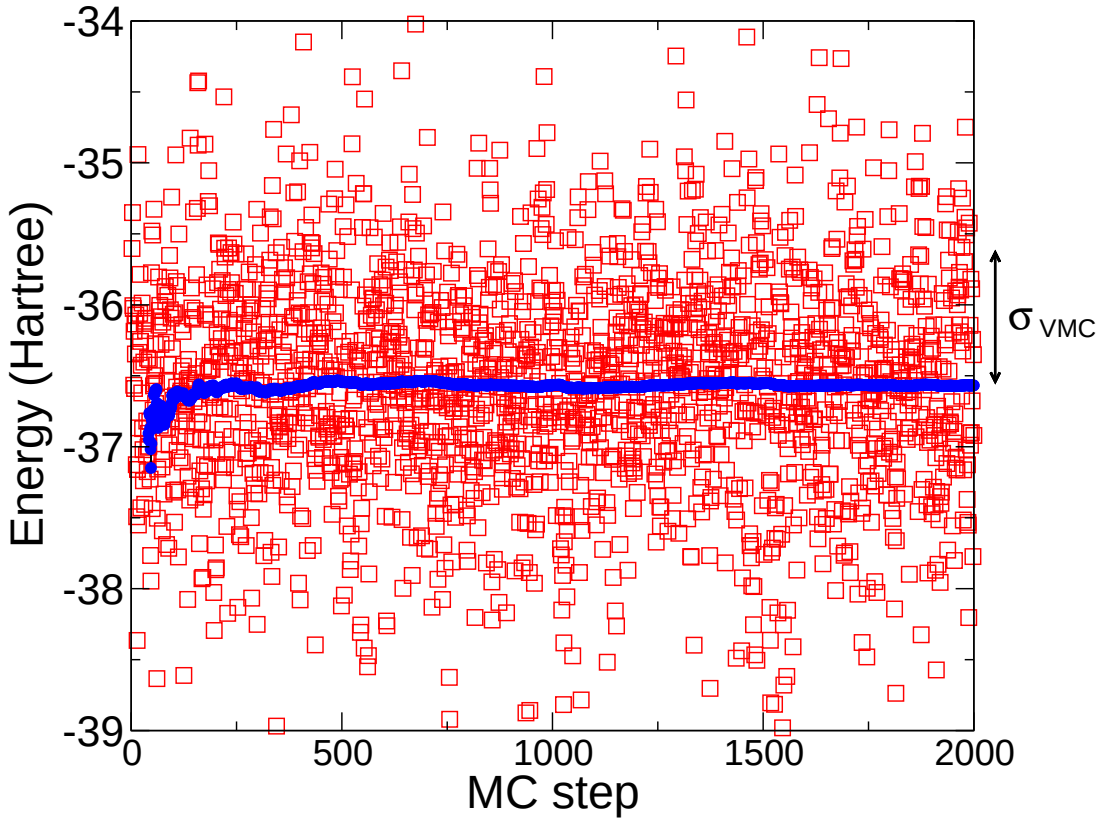
$$\sigma^2 = \frac{\langle \Psi | (\mathcal{H} - E_V)^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle (E_L(\mathbf{R}) - E_V)^2 \rangle_P$$

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

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

Typical VMC run

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: To do list

– Method to **sample** distribution function $P(\mathbf{R}) = \dfrac{|\Psi(\mathbf{R})|^2}{\int \mathrm{d}\mathbf{R}|\Psi(\mathbf{R})|^2}$

$\rightarrow$ Obtain a set of $\{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_M\}$ distributed as $P(\mathbf{R})$

How? As in classical Monte Carlo with Metropolis algorithm!

– Build the wave function $\Psi(\mathbf{R})$. Which **functional form** ?

Here, we spend most of our time, open topic of research

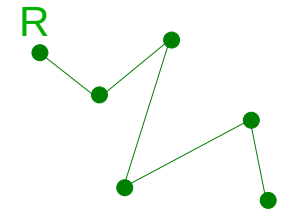
– Compute expectation values $\dfrac{\langle\Psi|\mathcal{O}|\Psi\rangle}{\langle\Psi|\Psi\rangle}$

Reformulate them to reduce fluctuations, open topic of research

How do we sample $P(\mathbf{R})$?

Generate a Markov chain

$$\ldots \xrightarrow{M} \mathbf{R} \xrightarrow{M} \mathbf{R}' \xrightarrow{M} \mathbf{R}'' \xrightarrow{M} \ldots$$



- Start from arbitrary initial state $\mathbf{R}_i$
- Use $M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ as probability for transition $\mathbf{R}_\mathrm{i} \to \mathbf{R}_\mathrm{f}$ so that

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

- Evolve the system by repeated application of M

Starting from an arbitrary distribution P_init, we want to evolve to P

$\to$ Impose stationarity condition

Stationarity condition

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

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

- Stationarity condition

  $\Rightarrow$ If we start with P, we continue to sample P

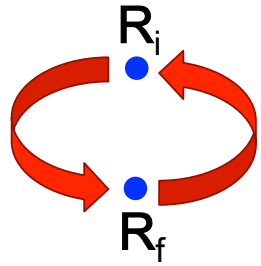
- Stationarity condition + stochastic property of M + ergodicity

  $\Rightarrow$ Any initial distribution will evolve to P

More stringent condition

In practice, we impose **detailed balance** condition

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



Stationarity condition can be obtained by summing over $\mathbf{R}_\mathrm{i}$

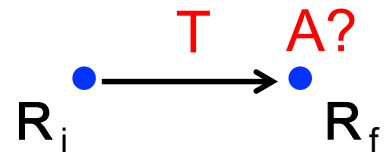
$$\int \mathrm{d}\mathbf{R}_\mathrm{i} M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\, P(\mathbf{R}_\mathrm{i}) = \underbrace{\int \mathrm{d}\mathbf{R}_\mathrm{i} M(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})}_{1}\, P(\mathbf{R}_\mathrm{f}) = P(\mathbf{R}_\mathrm{f})$$

Detailed balance is a sufficient but not necessary condition

How do we construct the transition matrix P in practice?

Metropolis method $\rightarrow$ Write M as proposal T $\times$ acceptance A

$$M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$$



Let us rewrite the detailed balance condition

$$\begin{aligned}
M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ P(\mathbf{R}_\mathrm{i}) &= M(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ P(\mathbf{R}_\mathrm{f}) \\
A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ P(\mathbf{R}_\mathrm{i}) &= A(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ P(\mathbf{R}_\mathrm{f}) \\
\Rightarrow \quad \frac{A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}{A(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})} &= \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ P(\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ P(\mathbf{R}_\mathrm{i})}
\end{aligned}$$

Choice of acceptance matrix A (1)

Detailed balance condition is

$$\frac{A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}{A(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})} = \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ P(\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ P(\mathbf{R}_\mathrm{i})}$$

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

Any function $A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = F\left(\frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ P(\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ P(\mathbf{R}_\mathrm{i})}\right)$ with

$$\frac{F(x)}{F(1/x)} = x \quad \text{and} \quad 0 \leq F(x) \leq 1$$

will do the job!

Choice of acceptance matrix A

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

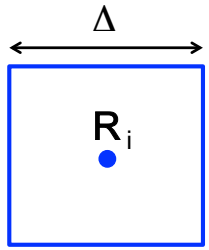
$$A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \min\left\{1, \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ P(\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ P(\mathbf{R}_\mathrm{i})}\right\}$$

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

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

$\rightarrow$ $P(\mathbf{R}) = |\Psi(\mathbf{R})|^2$

Original Metropolis method



Symmetric $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = 1/\Delta^{3N}$ $\Rightarrow$ $A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \min\left\{1, \frac{P(\mathbf{R}_\mathrm{f})}{P(\mathbf{R}_\mathrm{i})}\right\}$

Better choices of proposal matrix T

Sequential correlation $\Rightarrow$ $M_{\text{eff}} < M$ independent observations

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

Aim is to achieve fast evolution and reduce correlation times

Use freedom in choice of T: For example, use available trial Ψ

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

Acceptance and $T_{\rm corr}$ for the total energy E_V

Example: All-electron Be atom with simple wave function

Simple Metropolis

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

Drift-diffusion transition

τ	$T_{\rm corr}$	$\bar{A}$
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 $P(\mathbf{R})$ and proposal matrix $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$
2. Initialize the configuration $\mathbf{R}_i$
3. Advance the configuration from $\mathbf{R}_i$ to $\mathbf{R}'$
   a) Sample $\mathbf{R}'$ from $T(\mathbf{R}'|\mathbf{R}_\mathrm{i})$.
   b) Calculate the ratio $p = \dfrac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}')}{T(\mathbf{R}'|\mathbf{R}_\mathrm{i})} \dfrac{P(\mathbf{R}')}{P(\mathbf{R}_\mathrm{i})}$
   c) Accept or reject with probability p

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

      if $\chi < p$, move accepted $\rightarrow$ set $\mathbf{R}_\mathrm{f} = \mathbf{R}'$

      if $\chi > p$, move rejected $\rightarrow$ set $\mathbf{R}_\mathrm{f} = \mathbf{R}$
4. Throw away first κ configurations of equilibration time
5. Collect the averages

Expectation values in variational Monte Carlo (1)

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

$$
\begin{aligned}
E_V &= \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \\
&= \int \mathrm{d}\mathbf{R} \, \frac{\mathcal{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \, \frac{|\Psi(\mathbf{R})|^2}{\int \mathrm{d}\mathbf{R} |\Psi(\mathbf{R})|^2} \\
&= \int \mathrm{d}\mathbf{R} \, E_{\mathrm{L}}(\mathbf{R}) \, P(\mathbf{R}) \\
&= \langle E_{\mathrm{L}}(\mathbf{R}) \rangle_P \approx \frac{1}{M} \sum_{i=1}^{M} E_{\mathrm{L}}(\mathbf{R}_i)
\end{aligned}
$$

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

$\rightarrow$ 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

$$E_V = \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle E_{\mathrm{L}}(\mathbf{R}) \rangle_P$$

The variance of the local energy is given by

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

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

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