UCSB, Physics 129L, Computational Physics Lecture notes, Week 10B

Zihang Wang (UCSB), zihangwang@ucsb.edu

March 14, 2025

Contents

1 Quantum Monte Carlos

1.1 variational principle and trial quantum state

The **variational principle** states, any expectation value with a trial wavefunction $|\Psi_T\rangle$ has larger or equal energy with respect to the ground state energy,

$$E_T = \frac{\langle\Psi_T|H|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} = \int_V d\mathbf{R}\,\Psi_T^* H\Psi_T(\mathbf{R}) \geq E_0$$

where E_0 is the true ground state energy, and $\mathbf{R}$ is a **specific many-body configuration labeled by the position**. We should also note that the trial wavefunction is the **position basis trial wavefunction** $\Psi_T(\mathbf{R}_i) = \langle\mathbf{R}_i|\Psi_T\rangle$.

This implies that: if E_T is **close to the exact energy**, Ψ_T is a good approximation. On the other hand, if E_T is **far from the exact energy**, we need a better wavefunction.

The **local energy** is defined as:

$$E_L(\mathbf{R}) = \frac{H\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}.$$

If Ψ_T is an exact eigenstate of H, then $E_L(\mathbf{R})$ is a **constant** (zero variance). In reality, the variance is **not zero** due to imperfections in Ψ_T.

$$\sigma^2 = \langle E_L^2 \rangle - \langle E_L \rangle^2.$$

If we have **low variance**, the trial wavefunction Ψ_T is close to an eigenstate.

1.2   Importance sampling on trial wavefunctions

Suppose we wish to compute the expectation value of an operator $O[R]$, (such as differential (∂_R) or function $(f(R))$ operators), with respect to a wavefunction $\Psi(R)$ (which we do not know exactly). The expectation value is

$$\langle O[R] \rangle = \frac{\displaystyle\int dR\, \Psi^*(R)\, O[R]\, \Psi(R)}{\displaystyle\int dR\, |\Psi(R)|^2}.$$

In Monte Carlo integration, we introduce a probability density $P(R)$ and, using the **importance sampling**, we rewrite the integrals as,

$$\langle O[R] \rangle = \frac{\displaystyle\int dR\, \frac{\Psi^*(R)\, O[R]\, \Psi(R)}{P(R)}\, P(R)}{\displaystyle\int dR\, \frac{|\Psi(R)|^2}{P(R)}\, P(R)} \approx \frac{\dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \frac{\Psi^*(R_i)\, O[R]\, \Psi(R_i)}{P(R_i)}}{\dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \frac{|\Psi(R_i)|^2}{P(R_i)}}.$$

In the last equality, we **sample configurations** R_i from $P(R)$, we approximate the integrals by the following sum. This is the **Variational Monte Carlo method**.

A practical choice is to set the probability,

$$P(R) = \frac{|\Psi_T(R)|^2}{\displaystyle\int dR\, |\Psi_T(R)|^2},$$

where $\Psi_T(R)$ is the **trial wave function** used as an approximation to the true (and unknown) wave function $\Psi(R)$.

We define the **local estimator** for $O[R]$ with respect to the trial wavefunction, and define the expectation value as the sum of local estimators,

$$O_L(R_i) = \frac{\Psi_T^*(R_i)\, O[R_i]\, \Psi_T(R_i)}{|\Psi_T(R_i)|^2} = \frac{O[R_i]\, \Psi_T(R_i)}{\Psi_T(R_i)}, \quad \langle O[R] \rangle \approx \frac{1}{N}\sum_{i=1}^{N} O_L(R_i),$$

with the configurations R_i drawn from $P(R)$: When we sample from $|\Psi_T(R)|^2$, we are generating many-body configurations R (i.e., electron positions) that follow the probability distribution $P(R)$ given by the trial wavefunction $\Psi_T(R)$.

This gives a very big problem for fermions: $P(R) \sim \Psi_T(R)$ is **symmetric** under particle exchange, therefore it is equal likely to have one configuration R_i and another configuration R_j under an odd number of particle swaps, such that the trial wavefunction $\Psi_T(R_i) = -\Psi_T(R_j)$ since the **fermion wavefunction is antisymmetric**: For a system of N fermions, suppose you have N single-particle orbitals $\psi_1, \psi_2, \ldots, \psi_N$. The many-fermion wavefunction is written as,

$$\Psi(x_1, x_2, \ldots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(x_1) & \psi_2(x_1) & \cdots & \psi_N(x_1) \\ \psi_1(x_2) & \psi_2(x_2) & \cdots & \psi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(x_N) & \psi_2(x_N) & \cdots & \psi_N(x_N) \end{vmatrix}.$$

Here, x_i represents the coordinates (and possibly spin) of the ith particle. The determinant ensures that swapping any two rows (which corresponds to exchanging two particles) introduces a minus sign in the wavefunction, thereby guaranteeing antisymmetry.

Since we are sampling the symmetric probability, the configurations R_i drawn from this probability are used to calculate the expectation value via the sum of local estimators,

$$O_L(R) = \frac{O[R]\,\Psi_T(R)}{\Psi_T(R)}, \quad \langle O[R] \rangle \approx \frac{1}{N} \sum_{i=1}^{N} O_L(R_i).$$

We can see that **if the numerator $O[R]\,\Psi_T(R)$ does not perfectly agree with the sign change with respect to $\Psi_T(R)$ under odd number particle swaps,**

$$\Psi_T(R_i) = -\Psi_T(R_j) \nRightarrow O[R_i]\,\Psi_T(R_i) = -O[R_j]\,\Psi_T(R_j) \Rightarrow \langle O[R] \rangle = \text{oscillate}, \tag{1}$$

we will experience the **fermionic sign problem: the expectation value estimation has large variance and oscillate (does not stabilize) regardless how many samples we have**. In particular, if the numerator is symmetric under particle swap, the expectation value oscillates around zero regardless of the estimation,

$$\Psi_T(R_i) = -\Psi_T(R_j) \Rightarrow O[R_i]\,\Psi_T(R_i) = O[R_j]\,\Psi_T(R_j) \Rightarrow \langle O[R] \rangle \approx 0. \tag{2}$$

This leads to severe cancellation errors in Monte Carlo simulations.

Let's consider the kinetic energy operator,

$$O = -\frac{\hbar^2}{2m} \sum_{i=1}^{N} \nabla_i^2.$$

We use a simple antisymmetric trial wavefunction for two fermions:

$$\Psi_T(x_1, x_2) = \sin(kx_1) - \sin(kx_2),$$

which is antisymmetric under exchange $x_1 \leftrightarrow x_2$. The kinetic energy operator in one dimension is given by,

$$T = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}.$$

Applying this to the trial wavefunction, we have,

$$\frac{d^2}{dx_1^2}\Psi_T(x_1, x_2) = -k^2\sin(kx_1), \quad \frac{d^2}{dx_2^2}\Psi_T(x_1, x_2) = -k^2\sin(kx_2).$$

The total kinetic energy estimator is,

$$O_L(x_1, x_2) = -\frac{\hbar^2}{2m}\frac{\sum_{i=1}^{2}\nabla_i^2\Psi_T(x_1, x_2)}{\Psi_T(x_1, x_2)} = -\frac{\hbar^2 k^2}{2m}\frac{\sin(kx_1) + \sin(kx_2)}{\sin(kx_1) - \sin(kx_2)}.$$

We can see that the denominator $\sin(kx_1) - \sin(kx_2)$ changes sign when $x_1 \approx x_2$. This causes large fluctuations in $O_L(x_1, x_2)$.

If we sample positions x_1, x_2 using the probability distribution $|\Psi_T(x_1, x_2)|^2$, we get equal probabilities of getting (x_1, x_2) and (x_2, x_1), leading to positive and negative values of the estimator. The sign problem naturally appears in kinetic energy evaluations due to the antisymmetric nature of $\Psi_T(x_1, x_2)$.

2   Density Functional Theory

Density Functional Theory (DFT) is used to determine the electronic structure of many-body systems (e.g., atoms, molecules, and solids). Instead of working with the many-electron wavefunction, which depends on $3N$ spatial coordinates, DFT reformulates the problem in terms of the **electron density** $\rho(\mathbf{r})$, which depends only on three spatial coordinates.

DFT has its mathematical foundations based on two theorems established by Hohenberg and **Kohn** (UCSB) in 1964. The first Hohenberg–Kohn theorem states that: **The ground-state electron density** $\rho(\mathbf{r})$ **uniquely determines the external potential** $V_{\text{ext}}(\mathbf{r})$ (up to an additive constant), and hence all properties of the system. Mathematically, this means there is a one-to-one mapping:

$$\rho(\mathbf{r}) \quad \longleftrightarrow \quad V_{\text{ext}}(\mathbf{r}),$$

which implies that the ground-state many-electron wavefunction Ψ_0 and all observables are functionals of $\rho(\mathbf{r})$. It **does not** imply that $\rho(\mathbf{r})$ is only induced by external potential.

The first Hohenberg–Kohn theorem states that: There exists **a universal energy functional** $E[\rho]$ **such that the ground-state energy** E_0 **is the**

4

minimum of this functional when evaluated at the true ground-state density $\rho_0(\mathbf{r})$:

$$E_0 = \min_{\rho} \{E[\rho]\}, \quad E[\rho] = F[\rho] + \int d\mathbf{r}\, V_{\text{ext}}(\mathbf{r})\, \rho(\mathbf{r}),$$

where $F[\rho]$ is a universal functional of the density containing the kinetic energy and electron–electron interactions. In other words, for any trial density $\rho(\mathbf{r})$ that is N-representable (i.e., it integrates to the number of electrons N),

$$E[\rho] \geq E_0.$$

3 Kohn–Sham Formalism

Since the exact form of $F[\rho]$ is unknown, Kohn and Sham introduced a practical approach. They **map the interacting system onto a fictitious non-interacting system that yields the same ground-state density**. The Kohn–Sham energy functional is written as:

$$E[\rho] = T_s[\rho] + E_{\text{ext}}[\rho] + E_{\text{H}}[\rho] + E_{\text{xc}}[\rho],$$

where:

- $T_s[\rho]$ is the kinetic energy of non-interacting electrons:

$$T_s[\rho] = -\frac{\hbar^2}{2m} \sum_{i=1}^{N} \int d\mathbf{r}\, \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}),$$

with $\{\phi_i(\mathbf{r})\}$ being the Kohn–Sham orbitals.

- $E_{\text{ext}}[\rho]$ is the interaction energy with the external potential:

$$E_{\text{ext}}[\rho] = \int d\mathbf{r}\, V_{\text{ext}}(\mathbf{r})\, \rho(\mathbf{r}).$$

- $E_{\text{H}}[\rho]$ is the classical electrostatic (**Hartree**) energy:

$$E_{\text{H}}[\rho] = \frac{1}{2} \int d\mathbf{r}\, d\mathbf{r}'\, \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}.$$

- $E_{\text{xc}}[\rho]$ is the exchange–correlation energy functional, which includes all the many-body effects (exchange and correlation) not captured by the previous terms:

$$E_{\text{xc}}[\rho] = (T[\rho] - T_s[\rho]) + (V_{\text{ee}}[\rho] - E_{\text{H}}[\rho]).$$

For example, in LDA, we use the **exchange interaction from the free electron gas, rather than the true (non-local) exchange interaction** (Fock contribution). This can be incorporated via nonlocal Fock exchange energy calculated from the **Kohn–Sham orbitals**.

The Kohn–Sham orbitals are obtained by solving the **self-consistent Kohn–Sham equations**:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}),$$

where the effective potential is given by

$$V_{\text{eff}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + \int d\mathbf{r}' \, \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{\text{xc}}(\mathbf{r}),$$

with the exchange–correlation potential defined as the functional derivative:

$$V_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})}.$$

4   Exchange–Correlation Approximations

Since the exact form of $E_{\text{xc}}[\rho]$ is unknown, various approximations are used:

4.1   Local Density Approximation (LDA)

In LDA, **the exchange–correlation energy is approximated by assuming that the system locally behaves like a homogeneous electron gas**:

$$E_{\text{xc}}^{\text{LDA}}[\rho] = \int d\mathbf{r} \, \rho(\mathbf{r}) \, \varepsilon_{\text{xc}}(\rho(\mathbf{r})),$$

where $\varepsilon_{\text{xc}}(\rho)$ is the exchange–correlation energy per particle of a uniform electron gas. To calculate the exchange coorelation potential $V_{\text{xc}}(\mathbf{r})$, we use the standard property of the functional derivative that:

$$\frac{\delta}{\delta \rho(r)} \left(\rho(r') \, \varepsilon_{\text{xc}}(\rho(r')) \right) = \delta(r - r') \, \varepsilon_{\text{xc}}(\rho(r')) + \rho(r') \, \frac{d}{d\rho(r')} \, \varepsilon_{\text{xc}}(\rho(r')).$$

Applying this to the integrand, we have,

$$V_{\text{xc}}(r) = \varepsilon_{\text{xc}}(\rho(r)) + \rho(r) \, \frac{d}{d\rho(r)} \, \varepsilon_{\text{xc}}(\rho(r)).$$

Generalized Gradient Approximation (GGA) improves upon LDA by including the gradient of the density:

$$E_{\text{xc}}^{\text{GGA}}[\rho] = \int d\mathbf{r} \, f\left(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}) \right),$$

where f is a function of both density and its gradient.

4.2 Self-Consistent Field (SCF) Procedure

The Kohn–Sham equations are solved self-consistently:

- **Initialize the Density:** Begin with an initial guess $\rho^{(0)}(\mathbf{r})$.

- **Compute Effective Potential:** Evaluate $V_{\text{eff}}^{(0)}(\mathbf{r})$ using the guessed density.

- **Solve Kohn–Sham Equations:** Solve

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}^{(0)}(\mathbf{r}) \right] \phi_i^{(0)}(\mathbf{r}) = \varepsilon_i^{(0)} \phi_i^{(0)}(\mathbf{r})$$

  to obtain the **Kohn–Sham orbitals**.

- **Update the Density:** Construct a new density from the **occupied orbitals**:

$$\rho^{(1)}(\mathbf{r}) = \sum_i^{\text{occupied}} |\phi_i^{(0)}(\mathbf{r})|^2.$$

  . This will be used to update the V_{eff}.

We repeat steps 2–4 until the density and energy converge within a given tolerance.

5 Approximating $\varepsilon_{xc}(\rho)$ Using Quantum Monte Carlo

In the Local Density Approximation (LDA) of Density Functional Theory (DFT), the exchange–correlation energy of an inhomogeneous electron system is approximated by assuming that locally the system behaves like a homogeneous electron gas. In mathematical terms, the LDA exchange–correlation energy functional is given by

$$E_{xc}^{\text{LDA}}[\rho] = \int d\mathbf{r}\, \rho(\mathbf{r})\, \varepsilon_{xc}\big(\rho(\mathbf{r})\big),$$

where $\varepsilon_{xc}(\rho)$ is the exchange–correlation energy per electron for a homogeneous electron gas of density ρ.

For the homogeneous electron gas, the total ground-state energy per electron can be expressed as the following sum,

$$\varepsilon(\rho) = T_s(\rho) + \varepsilon_x(\rho) + \varepsilon_c(\rho),$$

where,

- $T_s(\rho)$ is the kinetic energy per electron for a **non-interacting** (free) electron gas.

- $\varepsilon_x(\rho)$ is the exchange energy per electron, which can be derived analytically,

$$\varepsilon_x(\rho) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} \rho^{\frac{1}{3}}.$$

- $\varepsilon_c(\rho)$ is the correlation energy per electron (many-body term).

Quantum Monte Carlo methods are employed to compute the total energy $\varepsilon(\rho)$ for a given electron density ρ. The total energy per electron $\varepsilon(\rho)$ obtained from QMC includes all contributions,

$$\varepsilon(\rho) = T_s(\rho) + \varepsilon_x(\rho) + \varepsilon_c(\rho).$$

Since both $T_s(\rho)$ and $\varepsilon_x(\rho)$ are known exactly for the homogeneous electron gas, the correlation energy per electron is isolated by subtracting these known terms from the total QMC energy,

$$\varepsilon_c(\rho) = \varepsilon(\rho) - T_s(\rho) - \varepsilon_x(\rho).$$

For example, using the **variational Monte Carlo method** we discussed previously, the operator becomes the Hamiltonian $O[R] = H$, and using the trial wavefunction and the variational principle, we can estimate the round state energy $E_T \geq E_0$. In Variational Monte Carlo (VMC), the energy is computed using a trial wavefunction $\Psi_T(R)$, but this is only an approximation of the true ground-state wavefunction $\Psi_0(R)$. This method depends on the quality of the trial wavefunction.

5.1 Diffusion Monte Carlo

Diffusion Monte Carlo improves upon VMC by projecting out the true ground-state wavefunction Ψ_0 using an imaginary time evolution approach.

Using **imaginary time** evolution of the Schrodinger equation, DMC gradually "diffuse out" the excited-state contributions from the initial trial wavefunction $\Psi_T(R_i)$ (real valued and time-independent), leaving only the ground-state wavefunction $\Psi_0(G)$ as the **asymptotic state** (G is the ground state configuration),

$$|\Psi(\tau)\rangle = e^{-\tau(H-E_T)} |\Psi(0)\rangle,$$

where the function $|\Psi(R,\tau)\rangle$ represents the quantum state evolving in imaginary time τ. The operator $e^{-\tau(H-E_T)}$ acts on the initial state $|\Psi(0)\rangle$ and evolves it over time. E_T is the trial energy with respect to a specific many-body configuration R_i (collection of many-body spacial coordinates), drawn from the probability distribution, $P(R) = |\Psi_T|^2$,

$$E_T = \int_V d\mathbf{R}\Psi_T^* H\Psi_T(\mathbf{R}) \geq E_0.$$

As $\tau \to \infty$, $\Psi(R,\tau)$ converges to the ground-state wavefunction Ψ_0.

As a side note, imaginary time evolution is very convenient and natural in describing amplitude averages of quantum evolutions in a given time interval: Imagine a plane wave with frequency must larger than the system's inverse time scale. $\omega_h \gg 1/t$. The real time evolution e^{iHt} is fast oscillating, and rapid oscillations cancel out,

$$\int_0^t e^{i\omega_h t} \approx 0. \tag{3}$$

This is also seen in quantum optics as the **rotating wave approximation**. **The imaginary time translates this rapid oscillating behavior to the concept of diffusion**: rapid oscillations directly becomes an effective large "diffusion constant".

For each drawn configuration, the trial wavefunction Ψ_T can be written as a sum of energy eignstates $|\phi_n\rangle$ with energies E_n and weights c_n:

$$|\Psi_T\rangle = \sum_n c_n |\phi_n\rangle.$$

We can then we apply the imaginary time evolution operator:

$$|\Psi(\tau)\rangle = e^{-\tau(H-E_T)} \sum_n c_n |\phi_n\rangle = e^{-\tau(E_n-E_T)} \sum_n c_n |\phi_n\rangle,$$

Since the ground state (lowest energy eigenstate) has the lowest energy E_0, the excited states decay exponentially faster than the ground state. Therefore, as $\tau \to \infty$, the leading order will be the term with E_0,

$$|\Psi(\tau)\rangle \approx c_0 e^{-\tau(E_0-E_T)} |\phi_0\rangle.$$

If the Hilbert space dimension is large, direct imaginary time evolution is impractical. Let's split the Hamiltonian into two the kinetic and potential contributions,

$$H = K + V, \tag{4}$$

and the variation of the updated wavefunction over a small step from $\tau \to \tau + d\tau$ is given by,

$$|\Psi(d\tau)\rangle = \left[e^{-(\tau+d\tau)(H-E_T)} \right] |\Psi(0)\rangle.$$

and we must evaluate the term $e^{-d\tau(K+V)}$. However, directly exponentiating the Hamiltonian is computationally expensive for many-body systems. Let's split the kinetic and potential term into two parts such that it approximates the full term $e^{-d\tau(K+V)}$. The simplest will be the **Trotter expansion**,

$$e^{-d\tau K} e^{-d\tau V} \approx e^{-d\tau(K+V)+(d\tau)^2/2\cdot[K,V]+\mathcal{O}[(d\tau)^3]}. \tag{5}$$

and while it is accurate up to the third order $(d\tau)^3$, we have to evaluate the commutator. Therefore, we need to introduce a commutator that cancels the second order contribution. We can go further by considering the general idea in truncation accuracy we looked at in the previous lectures: **Combing two lower order expressions may result in a higher order accuracy**.

9

Let's consider the following expression,

$$
\begin{aligned}
e^{-d\tau K/2} e^{-d\tau V} e^{-d\tau K/2} &= e^{-d\tau K/2 - d\tau V + (d\tau)^2/4 \cdot [K,V]} e^{-d\tau K/2} \\
&= e^{-d\tau K/2 - d\tau V + (d\tau)^2/4 \cdot [K,V] - d\tau K/2 + (d\tau)^2/4 \cdot [K+V,K] + \mathcal{O}[(d\tau)^3]} \\
&= e^{-d\tau K/2 - d\tau V + (d\tau)^2/4 \cdot [K,V] - d\tau K/2 + (d\tau)^2/4 \cdot [V,K] + \mathcal{O}[(d\tau)^3]} \\
&= e^{-d\tau K/2 - d\tau V + (d\tau)^2/4 \cdot [K,V] - d\tau K/2 - (d\tau)^2/4 \cdot [K,V] + \mathcal{O}[(d\tau)^3]} \\
&= e^{-d\tau K/2 - d\tau V - d\tau K/2 + \mathcal{O}[(d\tau)^3]}.
\end{aligned}
\tag{6}
$$

We in fact increase the accuracy to $\mathcal{O}[(d\tau)^3]$ by using the symmetric property of the expansion. This is the **Trotter-Suzuki splitting**. It usually is compared with the mid point rule: we first propagate the wavefunction by half time step, then scatter via a uniform potential over the full time step, and finally propagate the remaining half-step. This can be nicely illustrated by inserting two resolution of identities,

$$
\begin{aligned}
|\Psi(\tau)\rangle &= e^{-d\tau K/2} e^{-d\tau V} e^{-d\tau K/2} |\Psi(0)\rangle \\
&= \int_V dR \int_{V'} dR' e^{-d\tau K/2} |R\rangle \langle R| e^{-d\tau V} |R'\rangle \langle R'| e^{-d\tau K/2} |\Psi(0)\rangle,
\end{aligned}
\tag{7}
$$

and we define the **propagator (Green's function) for free-particle diffusion (plane wave solution in real time)**,

$$
G(R, R', \frac{d\tau}{2}) = \langle R| e^{-d\tau K/2} |R\rangle,
\tag{8}
$$

It is a Green's function because it satisfies the imaginary-time Schrodinger evolution (or diffusion equation),

$$
\partial_\tau |\Psi^{\text{free}}(\tau)\rangle = -K |\Psi^{\text{free}}(\tau)\rangle,
\tag{9}
$$

and when projecting onto the position basis $\Psi^{\text{free}}(R, \tau) = \langle R|\Psi^{\text{free}}(\tau)\rangle$, the solution is given by,

$$
\Psi^{\text{free}}(R, \tau) = \int_V dR' \langle R| e^{-\tau K} |R'\rangle \Psi^{\text{free}}(R', 0) = \int_V dR' G(R, R', \tau, 0) \Psi^{\text{free}}(R', 0),
\tag{10}
$$

where in the last line, we use the Green's function definition, and we assume a **time independent Hamiltonian** (no time ordering). In other words, the Green's function of the free-particle diffusion follows the following point source equation,

$$
\left(\partial_\tau - \frac{D}{2} \partial_R^2 \right) G(R, R', \tau, 0) = \delta(R - R') \delta(\tau - 0),
\tag{11}
$$

where $D \sim 1/m$ is the diffusion constant or the "inverse mass" of the free particle. The solution is classic Gaussian,

$$
G(R, R', \tau, 0) = \left(\frac{1}{2\pi D d\tau} \right)^{3N/2} \exp\left(-D \frac{|R - R'|^2}{2 D d\tau} \right),
\tag{12}
$$

where N is the number of particles (in 3 dimension), and recall that in stochastic calculus we discuss previously, it is the simplest solution of a purely diffusive solution of the **Fokker–Planck equation**. $G(R, R', \tau, 0)$ sometimes is called the **heat kernel**.

Then, since we know that the Fokker–Planck equation describes the he probability density of a stochastic random variable (in our case $dX \rightarrow dR = R - R'$, we can directly write down the corresponding **stochastic differential equation (or Langevin or Ito process)** under the Ito formulation,

$$dR = \sqrt{2D}dW_\tau \tag{13}$$

Let's say we initially start from the **quantum state** $|\Psi(0)\rangle$ (with special wavefunction $\Psi(R', 0) = \langle R'|\Psi(0)\rangle$), and after time $d\tau$, the probability of the evolved quantum state in a particular position configuration $|R\rangle$ is given by the overlap $\Psi(R, d\tau) = \langle R|\Psi(d\tau)\rangle$, we can rewrite the above evolution equation in the position basis as,

$$\Psi(R, d\tau) = \int_{V''} dR'' \int_{V'''} dR''' G(R, R'', d\tau, \frac{d\tau}{2}) \langle R''|e^{-d\tau V}|R'''\rangle G(R''', R', \frac{d\tau}{2}, 0)\Psi(R', 0). \tag{14}$$

To simplify the above expression, let's assume the **diagonal approximation** where no exchange interaction between different many-body interactions. Therefore, the interaction becomes,

$$\langle R''|e^{-d\tau V}|R'''\rangle = e^{-d\tau V(R'')}\delta(R'' - R'''), \tag{15}$$

and the half-step approximation becomes,

$$\Psi(R, d\tau) = \int_{V''} dR'' e^{-d\tau V(R'')} G(R, R'', d\tau, d\tau/2) G(R'', R', \frac{d\tau}{2}, 0)\Psi(R', 0). \tag{16}$$

Then, we assume a slow-varying potential such that,

$$e^{-d\tau V(R'')} \approx e^{-d\tau V(R')}, \tag{17}$$

We recover the full step free-particle diffusion Green's function, and ingratiate all possible configurations at τ,

$$\Psi(R, d\tau) = \int_V dR' e^{-d\tau V(R'')} G(R, R', d\tau, 0)\Psi(R', 0). \tag{18}$$

Since we want to sample from the trial wavefunction, we multiply both sides by an additional $\Psi_T(R)$, and define a reweighted product $\Phi(R, d\tau) = \Psi_T(R)\Psi(R, d\tau)$,

$$\Phi(R, d\tau) = \Psi_T(R)\Psi(R, d\tau) = \int_V dR' \frac{\Psi_T(R)}{\Psi_T(R')} e^{-d\tau V(R')} G(R, R', d\tau, 0)\Psi_T(R')\Psi(R', 0)$$

$$= \int_V dR' \frac{\Psi_T(R)}{\Psi_T(R')} e^{-d\tau V(R')} G(R, R', d\tau, 0)\Phi(R, 0) \tag{19}$$

11

We expand the following expression around $R = R' + r$

$$\frac{\Psi_T(R)}{\Psi_T(R')} = \exp\left(\log\left[\frac{\Psi_T(R)}{\Psi_T(R')}\right]\right) \approx \exp\left(\log\left[1 + \frac{r \cdot \nabla_R \Psi_T(R')}{\Psi_T(R')}\right]\right) \approx \exp\left(\frac{r \cdot \nabla_R \Psi_T(R')}{\Psi_T(R')}\right).$$
$$(20)$$

Using the previous free particle diffusion Green's function,

$$G(R, R', \tau, 0) = \left(\frac{D}{2\pi\tau}\right)^{d/2} \exp\left(-D\frac{|r|^2}{2\tau}\right),$$
$$(21)$$

we have,

$$\Phi(R, d\tau) = \left(\frac{1}{2\pi D d\tau}\right)^{3N/2} \int_V dR' \exp\left(\frac{r \cdot \nabla_R \Psi_T(R')}{\Psi_T(R')} - \frac{|r|^2}{2D d\tau} - d\tau V(R')\right) \Phi(R', 0)$$
$$(22)$$

We introduce the drift velocity,

$$\mathbf{v}_D(R') = D\frac{\nabla\Psi_T(R')}{\Psi_T(R')}.$$
$$(23)$$

The potential term has the following form,

$$-V d\tau = -d\tau(E_L(R') - E_T),$$
$$(24)$$

where $E_L(R')$ is the local energy,

$$E_L(R') = \frac{V(R')\,\Psi_T(R')}{\Psi_T(R')}.$$
$$(25)$$

The full modified Green's function with diffusion, drift, and potential is:

$$\Phi(R, d\tau) = \left(\frac{1}{2\pi D d\tau}\right)^{3N/2} \int_V dR' \exp\left(\frac{r \cdot \mathbf{v}_D(R')}{D} - \frac{|r|^2}{2D d\tau} - d\tau V(R')\right) \Phi(R', 0)$$
$$(26)$$

Then, we can complete the square:

$$\frac{1}{2D d\tau}\left(2r \cdot \mathbf{v}_D(R')d\tau - |r|^2\right) = \frac{1}{2D d\tau}\left(2r \cdot \mathbf{v}_D(R')d\tau - |r|^2 - |\mathbf{v}_D(R')d\tau|^2 + |\mathbf{v}_D(R')d\tau|^2\right)$$
$$= -\frac{(r - \mathbf{v}_D(R')d\tau)^2}{2D d\tau} + \frac{|\mathbf{v}_D(R')d\tau|^2}{2D d\tau}.$$
$$(27)$$

So the Green's function becomes:

$$\Phi(R, d\tau) = \left(\frac{1}{2\pi D d\tau}\right)^{3N/2} \int_V dR' \exp\left(-\frac{(r - \mathbf{v}_D(R')d\tau)^2}{2D d\tau} - d\tau\left(E_L(R') - E_T + \frac{|\mathbf{v}_D(R')|^2}{2D}\right)\right) \Phi(R', 0).$$
$$(28)$$

12

The above integral for the reweighted product is a general solution the following partial differential equation (PDE),

$$\partial_\tau \Phi(R,\tau) = \underbrace{D\nabla_R^2 \Phi(R)}_{\text{Diffusion}} - \underbrace{\nabla_R \cdot (\mathbf{v}_D \Phi(R))}_{\text{Drift}} - \underbrace{\left(E_L(R) - E_T + \frac{|\mathbf{v}_D|^2}{2D}\right)\Phi(R)}_{\text{Source}},$$

(29)

where in the source free case, above PDE becomes the **Fokker–Planck equation**. In Ito formulation, we obtain the **stochastic differential equation** (Langevin) of the configurational change,

$$dR = \sqrt{2D}dW_\tau + \mathbf{v}_D d\tau = \sqrt{2Dd\tau}\mathcal{N}(0,1) + \mathbf{v}_D d\tau,$$

(30)

where $\mathcal{N}(0,1)$ is the standard normal distribution.

Numerically, similar to the above SDE, Diffusion Monte Carlo separates the evolution of many-body configuration R from the source, and incorporates the source as an decision problem (like the metropolis hasting algorithm we discussed previously).

The DMC simulate SDE by **walkers**: they are stochastic representations of the quantum system's configuration R evolving in imaginary time τ. Initially, they could be randomly distributed according to the trial wavefunction. Each walker samples the configuration space and carries a weight that represents its statistical significance in approximating the wavefunction.

For each stochastic walker (represented by R_i), the displacement evolution updates the walker's configuration (called **diffusion**),

$$R_i(\tau + d\tau) = R_i(\tau) + \sqrt{2D}dW_\tau + \mathbf{v}_D d\tau.$$

Then, the walker is screened by the following weight (called **branching**),

$$w_i(\tau + d\tau) = w_i(\tau)e^{-d\tau(E_L(R_i)-E_T)},$$

where $E_L(\mathbf{R}_i)$ is the local energy and E_T is the trial energy.

Life: The walker is "alive" and continues to exist in the simulation as long as its weight $w_i(\tau)$ is positive. The weight decays over time based on the difference between the local energy $E_L(\mathbf{R}_i)$ and the reference energy E_T. If $E_L(\mathbf{R}_i) > E_T$, the walker decays more quickly, and if $E_L(\mathbf{R}_i) < E_T$, the walker decays more slowly. The equation for weight evolution is:

$$w_i(\tau + d\tau) = w_i(\tau)e^{-d\tau(E_L(\mathbf{R}_i)-E_T)}$$

This exponential factor indicates that the weight decreases more rapidly when the local energy exceeds the reference energy, and more slowly when the local energy is below the reference energy.

Death: If the walker's weight $w_i(\tau)$ falls below a threshold, the walker is considered "dead" and no longer contributes to the simulation. Essentially, the walker dies when the weight decays to a point where it doesn't have a significant effect on the system anymore.

In practice, the trial energy is by using the average local energy of all walkers in the system. The average local energy over time reflects the overall state of the system,

$$E_T(\tau) = \frac{1}{N} \sum_{i=1}^{N} E_L(\mathbf{R}_i(\tau)). \tag{31}$$

The exponential weights constantly keep walkers with lower energies (since negative sign) alive while exponentially increases the chance of death if a walker has an energy larger than the current walker ensemble energy. The reweighted wavefunction can be approximated via the number statistics of walkers,

$$\Phi(R, \tau) \approx \frac{1}{N} \sum_{i=1}^{N} \delta(R - R_i(\tau)). \tag{32}$$

This process will continue until $\Phi(R, \tau)$ reaches stationary.

However, from the definition,

$$\Phi(R, d\tau) = \Psi_T(R)\Psi(R, d\tau). \tag{33}$$

While the above expression works nicely with bosons, for fermions, as we discussed previously, if there is a relative sign mismatch, the expression $\Psi_T(R)\Psi(R, d\tau)$ will not be positively defined, and therefore, we cannot treat them as probability distributions. This is another realization of **the fermionic sign problem**, and various approximations are used (e.g. **fixed node approximation**, walkers will reflect back if it reaches a wave function node). However, this remains as an unsolved issue in modern quantum Monte Carlo research.