

## Quantum Monte Carlo method

Quantum mechanical systems are described by the Schrödinger equation, but in only handful of cases we can solve it analytically. Solving becomes more challenging in many-particle systems like He and higher atomic number elements, molecules, nucleus, electronic structures in solids and so on particularly because such systems consist of large number of interacting electrons, ions or nucleons. Nevertheless, we often wish to determine their ground state properties like energy etc. In that case we have methods like variational method, diffusion monte carlo, biased random walks etc.

**Variational Monte Carlo :** A quantum mechanical Hamiltonian of a system is of the form,

$$\hat{H} = \sum_i^N -\frac{\hbar^2}{2m} \nabla_i^2 + \sum_{i>j}^N V(|r_i - r_j|) \quad (1)$$

where  $N$  can be fairly large. We are interested to know the ground state energy of the above Hamiltonian by solving the differential equation (1). But we are aware that for even  $N > 2$  it requires approximate methods to solve it. The variational principle in QM proposes to make an educated guess of the wave function, the *trial* wave function  $\psi_T(\mathbf{r}, \boldsymbol{\alpha})$ , and calculate the expectation value of  $\hat{H}$  with respect to  $\psi_T$  (usually not an eigenstate of  $\hat{H}$ ), which we always assume to be real,

$$\langle \hat{H} \rangle = \frac{\int d\mathbf{r} \psi_T^*(\mathbf{r}, \boldsymbol{\alpha}) \hat{H} \psi_T(\mathbf{r}, \boldsymbol{\alpha})}{\int d\mathbf{r} \psi_T^*(\mathbf{r}, \boldsymbol{\alpha}) \psi_T(\mathbf{r}, \boldsymbol{\alpha})} \geq E_0 \quad (2)$$

Therefore,  $\langle \hat{H} \rangle$  is the upper bound of  $E_0$  which can be improved by tuning the parameter set  $\boldsymbol{\alpha}$

$$\frac{\partial \langle \hat{H} \rangle}{\partial \alpha_i} = 0 \quad \text{where } i = 1, 2, \dots, N \quad (3)$$

by using some minimization algorithm. In most cases a wavefunction has only small values in large parts of configuration space, therefore scanning the space uniformly will slow down the calculation to a grinding halt. This suggests the need for importance sampling *i.e.* the regions of configuration space where the wavefunction takes appreciable values are sampled more efficiently. The importance sampling can be implemented by defining appropriate probability distribution function,

$$p(\mathbf{r}) = \frac{\psi_T^*(\mathbf{r}, \boldsymbol{\alpha}) \psi_T(\mathbf{r}, \boldsymbol{\alpha})}{\int d\mathbf{r} \psi_T^*(\mathbf{r}, \boldsymbol{\alpha}) \psi_T(\mathbf{r}, \boldsymbol{\alpha})} \quad (4)$$

which when plugged in the expression (2), we get

$$\langle \hat{H} \rangle = \int d\mathbf{r} p(\mathbf{r}) \hat{\mathcal{E}}(\mathbf{r}), \quad \text{where } \hat{\mathcal{E}}(\mathbf{r}) = \frac{1}{\psi_T(\mathbf{r}, \boldsymbol{\alpha})} \hat{H} \psi_T(\mathbf{r}, \boldsymbol{\alpha}) \quad (5)$$

where the new operator  $\hat{\mathcal{E}}$  is called *local energy density*. This equation expresses the variational Monte Carlo approach. It is obvious that  $\langle \hat{H} \rangle = \langle \hat{\mathcal{E}} \rangle$  and therefore,

$$\langle \hat{H} \rangle = \langle \hat{\mathcal{E}}(\mathbf{r}) \rangle = \int d\mathbf{r} p(\mathbf{r}) \hat{\mathcal{E}}(\mathbf{r}) = \frac{1}{N_s} \sum_i \mathcal{E}(\mathbf{r}_i) \quad (6)$$

where  $\mathbf{r}_i$  are to be distributed according to the p.d.f.  $p(\mathbf{r})$  and  $N_s$  are the number of sample points. The points to sum over in  $\mathbf{r}$ -space are selected by importance sampling.

1. The Monte Carlo gets started with an initial choice of  $\{\mathbf{r}\}$  and variational parameters  $\{\boldsymbol{\alpha}\}$ .

2. Determine the  $p(\mathbf{r}) \propto |\psi_T(\mathbf{r}, \boldsymbol{\alpha})|^2$  (how to guess  $\psi_T$  will be discussed below.) The initial energy  $\langle \hat{H} \rangle = \langle \hat{\mathcal{E}} \rangle$  also gets determined.

3. Make small changes in  $\mathbf{r} \rightarrow \mathbf{r}'$  depending on some pre-determined MC step  $\epsilon$ ,

$$\mathbf{r}' = \mathbf{r} + \rho \star \epsilon \quad \text{where} \quad \rho \in [0, 1] \quad (7)$$

4. Determine  $p(\mathbf{r}')$  and perform Metropolis accept/reject test,

$$\text{If} \quad R = \frac{p(\mathbf{r}')}{p(\mathbf{r})} \geq 1 \quad \text{accept } \{\mathbf{r}'\} \quad (8)$$

$$\text{else} \quad R < 1 \quad \text{accept with probability } \rho \in [0, 1] \quad (9)$$

5. Update local energy  $\mathcal{E}$  and go to step 3 and re-do the loop for pre-determined MC trajectory length.

It is to be noted that updating  $\mathbf{r}$  in (7) does not require the knowledge of  $\psi_T$ , the wavefunction only enters the calculation through determining  $p(\mathbf{r})$ , hence important sampling is an important issue here. After a large number of iterations, the local energy  $\mathcal{E}$  should fluctuate around an average which gives the estimate for  $E_0$ . The trial wavefunction has to be chosen with some care to reduce the variance of the local energy  $\mathcal{E}$  as small as possible.

#### *Simple Harmonic Oscillator*

As with any new proposal, it is customary to try it out on harmonic oscillator especially 1-dimensional harmonic oscillator,

$$\hat{H} = -\frac{1}{2} \frac{d^2}{dx^2} + \frac{1}{2} x^2 \quad (10)$$

putting  $m = \hbar = k = 1$  for simplicity. Analytically, the ground state of this quantum mechanical SHO is

$$E_0 = \frac{1}{2} \quad \text{and} \quad \psi_0(x) = \mathcal{N} e^{-x^2/2} \quad (11)$$

For the purpose of variational MC, let the trial wavefunction, decided by its properties that  $\psi_T \rightarrow 0$  as  $x \rightarrow \pm\infty$ , and consequently the local energy be

$$\psi_T(x, \alpha) = \mathcal{N} e^{-\alpha x^2} \quad (12)$$

$$\mathcal{E} = \frac{1}{\psi_T} H \psi_T = e^{\alpha x^2} \left[ -\frac{1}{2} \frac{d^2}{dx^2} + \frac{1}{2} x^2 \right] e^{-\alpha x^2} = \alpha + x^2 \left( \frac{1}{2} - 2\alpha^2 \right) \quad (13)$$

Obviously, for  $\alpha = 1/2$  the energy  $\mathcal{E} = 1/2$  because the trial and the true ground state wavefunctions are identical. However, the question is whether this energy expectation value is minimum as a function of  $\alpha$ . The expectation  $\langle \mathcal{E} \rangle$  is hence calculated,

$$\langle \mathcal{E} \rangle = \int_{-\infty}^{\infty} dx \psi_T^2 \mathcal{E} = \frac{\alpha}{2} + \frac{1}{8\alpha} \quad \text{and} \quad \sigma_{\mathcal{E}}^2 = \frac{(1 - 4\alpha^2)^2}{32\alpha^2} \quad (14)$$

To do the MC integration in (14) for  $\langle \mathcal{E} \rangle$  choose  $x$  as random numbers (say,  $N_s$  of them) with Gaussian p.d.f  $|\psi_T|^2 \sim \exp(-2\alpha x^2)$  and not just any Gaussian. Then use equation (6) to calculate the average. Perform a series of such MC calculations each with slightly different value of  $\alpha$  (hence different Gaussian random distribution) and obtain  $\langle \mathcal{E} \rangle$  as a function of  $\alpha$ . At this point, use a suitable minimization function to determine  $\mathcal{E}_{\min}(\alpha_{\min})$  as the best estimate of  $E_0$ .

### Hydrogen atom

The next most common system we choose for illustrating variational MC is hydrogen atom. From analytical calculation we know that the overall wavefunction of hydrogen atom is complex, but this complex part enters from the  $\theta$  and  $\phi$  dependence. Therefore, we can concentrate solely on the radial part which is real. In fact, this is also a lesson for other atoms, molecules and nucleus for which we can argue that their radial parts are all real. The radial part of hydrogen atom Schrödinger equation is

$$-\frac{\hbar^2 r^2}{2m} \left( \frac{\partial}{\partial r} \left[ r^2 \frac{\partial R(r)}{\partial r} \right] \right) - \frac{ke^2}{r} R(r) + \frac{\hbar^2 l(l+1)}{2mr^2} R(r) = E R(r) \quad (15)$$

This equation is put in a simpler form by the transformation  $u(r) = rR(r)$  which gives,

$$-\frac{\hbar^2}{2m} \frac{\partial^2 u(r)}{\partial r^2} - \left( \frac{ke^2}{r} - \frac{\hbar^2 l(l+1)}{2mr^2} \right) u(r) = E u(r) \quad (16)$$

While solving the above transformed hydrogen atom Schrödinger equation it is imperative that it is written in terms of dimensionless variables because large variation of several constants result in loss of numerical precision. Defining the following

$$\rho = \frac{r}{\beta}, \quad \frac{mke^2\beta}{\hbar^2} = 1, \quad \lambda = \frac{m\beta^2}{\hbar^2} E \quad (17)$$

the equation that we are going to solve numerically is now

$$\left[ -\frac{1}{2} \frac{\partial^2}{\partial \rho^2} - \frac{1}{\rho} + \frac{l(l+1)}{2\rho^2} \right] u(\rho) = \lambda u(\rho) \Rightarrow H = -\frac{1}{2} \frac{\partial^2}{\partial \rho^2} - \frac{1}{\rho} + \frac{l(l+1)}{2\rho^2} \quad (18)$$

The analytical ground state solution is  $\lambda = -1/2$  and  $u(\rho) = \rho \exp(-\rho)$ . Taking clue from the boundary condition that at  $\rho = 0$ ,  $u(\rho) = 0$  and  $\rho \rightarrow \infty$ ,  $u(\rho) = 0$ , the trial wavefunction can be chosen as

$$u_T(\rho, \alpha) = \alpha \rho e^{-\alpha \rho} \quad (19)$$

where  $\alpha$  is a parameter to be tuned to minimize the upper bound of the ground state energy estimate. Please note, this is not the only form of wavefunction that satisfies the given boundary condition but certainly is the simplest. The local energy and probability distribution function, therefore, respectively are

$$\mathcal{E}_L = \frac{1}{u_T} \left[ -\frac{1}{2} \frac{\partial^2}{\partial \rho^2} - \frac{1}{\rho} + \frac{l(l+1)}{2\rho^2} \right] u_T = -\frac{1}{\rho} - \frac{\alpha}{2} \left( \alpha - \frac{2}{\rho} \right) \quad (20)$$

$$p(\rho) = \alpha^2 \rho^2 e^{-2\alpha \rho} \quad \text{where } \rho \in [0, \infty]. \quad (21)$$

### Helium atom

Generally Schrödinger equations for atomic, molecular and solid state systems, where we have interacting electrons and ions, can only be solved analytically through various approximations. Even then, a majority of such cases need helps from computer. As a simplest representative of this class is Helium atom containing many of the complexities of such many particle systems. The most common approximation used to solve Schrödinger equation for systems of interacting electrons is the *Born-Oppenheimer approximation*. It involves separating the nuclear and electronic degrees of freedom by using the fact that nucleus being much heavier little momentum will be transferred between electrons

and nucleus. If we label the distances between the electrons and nucleus as  $r_1$  and  $r_2$  and the electrons separation by a  $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ , then the Hamiltonian is

$$H = -\frac{1}{2} \nabla_1^2 - \frac{1}{2} \nabla_2^2 - \frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}}, \quad \text{where } Z = 2 \quad (22)$$

Even in this form Helium atom cannot be solved exactly and it is the last term above, the interaction between the electrons, that resists the effort. Let us put the repulsion term between the two electrons of Helium atom to zero! Then we have two completely separable Hamiltonians,

$$H_1 = -\frac{1}{2} \nabla_1^2 - \frac{Z}{r_1} \rightarrow H_1 \psi_1 = E_1 \psi_1 \quad (23)$$

$$H_2 = -\frac{1}{2} \nabla_2^2 - \frac{Z}{r_2} \rightarrow H_2 \psi_2 = E_2 \psi_2 \quad (24)$$

$$\psi = \psi_1 \psi_2 \quad \text{and} \quad E = E_1 + E_2 \quad (25)$$

Choosing trial functions for variational MC is not so straight forward since the obvious choice  $\exp(-\alpha r)$  will include situation when the electrons are closed to each other and we can no longer neglect the repulsion term. This problem aside, the trial wave function and its derivative must be well-defined at  $r_{1,2} = 0$ . We can put constraints on the wavefunction when the distance between one electron and the nucleus or two electrons approaches zero. This constraints are called *cusp conditions* and are related to the derivatives of the wavefunctions. Let us consider electron 1 is close to nucleus  $r_1 \rightarrow 0$  but two electrons are far from each other and  $r_2 \neq 0$ . The local energy can be written as (using  $r$  for collective symbol of  $r_{1,2}$ ),

$$\begin{aligned} \mathcal{E}_L(r) &= \frac{1}{\psi_T(r)} H \psi_T(r) = \frac{1}{\psi_T(r)} \left( -\frac{1}{2} \nabla_1^2 - \frac{Z}{r_1} \right) \psi_T(r) + \text{finite terms} \\ &= \frac{1}{\mathcal{R}_T(r_1)} \left( -\frac{1}{2} \frac{d^2}{dr_1^2} - \frac{1}{r_1} \frac{d}{dr_1} - \frac{Z}{r_1} \right) \mathcal{R}_T(r_1) + \text{finite terms} \end{aligned} \quad (26)$$

where  $\mathcal{R}_T(r_1)$  is the radial part of the wavefunction for electron 1 and assumed  $l = 0$ . For small values of  $r_1$ , the dominating term is

$$\lim_{r_1 \rightarrow 0} \mathcal{E}_L(r) = \frac{1}{\mathcal{R}_T(r_1)} \left( -\frac{1}{r_1} \frac{d}{dr_1} - \frac{Z}{r_1} \right) \mathcal{R}_T(r_1) \quad (27)$$

since  $d^2/dr_1^2$  does not diverge because  $\psi(r_1)$  is finite at origin. This indicates that in order to remove the diverging term in the limit  $r_1 \rightarrow 0$  we should have

$$\frac{1}{\mathcal{R}_T(r_1)} \frac{d\mathcal{R}_T(r_1)}{dr_1} = -Z \Rightarrow \mathcal{R}_T(r_1) \propto e^{-Zr_1} \quad (28)$$

A similar cusp condition applies to electron 2. For  $l \neq 0$  the  $Z$  in the above expression (28) is replaced by  $Z/(l+1)$ . The other constraint on the wavefunction corresponds to stopping electrons coming close to each other. When  $r_{12} \rightarrow 0$ , the resulting equation for  $r_{12}$  dependence is similar to electron-nucleus where  $-Z/r$  is replaced by  $1/r_{12}$  and the kinetic term is twice as large because the reduced mass of the two electrons is half the electron mass,

$$\left( -2 \frac{d^2}{dr_{12}^2} - \frac{4}{r_{12}} \frac{d}{dr_{12}} + \frac{2}{r_{12}} + \frac{l(l+1)}{r_{12}^2} \right) \mathcal{R}_T(r_{12}) = 0 \quad (29)$$

Changing variable  $\rho(r_{12}) = r_{12}^{-l} \mathcal{R}_T(r_{12})$  yields the cusp condition as

$$\frac{1}{\mathcal{R}_T(r_{12})} \frac{d\mathcal{R}_T(r_{12})}{dr_{12}} = \frac{1}{2(l+1)} \quad (30)$$

For general system containing more than two electrons, there is a cusp condition for each pair of  $ij$  electron pair. The cusp conditions (28) and (30) suggest a possible trial function (for  $l = 0$  ground state) as,

$$\psi_T(r) = e^{-\alpha(r_1+r_2)+r_{12}/2} \quad (31)$$

The problem is still not over. For helium atom we can place both the electrons in the hydrogen ground state  $1s$ . Since we are not considering spins of the electrons, there is no need to invoke Pauli principle but such approach cannot be taken for atoms beyond helium. Given this trial function the first thing would be to construct the pdf according to (4). The next step is calculation of full local energy  $\mathcal{E}_L(r)$  should involve the full helium Hamiltonian as in (22) and then use the expression (6) to calculate the ground state energy.

**Diffusion Monte Carlo :** In an improvement over variational method, the ground state of a quantum Hamiltonian is found by simulating diffusion process. In 1-dimension, let the probability of finding a given number of particles in an interval  $x + dx$  and  $x$  at a time  $t$  is  $\rho(x, t)$ . This quantity is the probability distribution function, which is equivalent to wavefunction in quantum mechanics. Rewriting Schrödinger equation in the form of diffusion equation is the first step of diffusion monte carlo.

Consider equally spaced positions along a line (1-dimension lattice)  $\Delta x$  distance apart. After a fixed time interval  $\Delta t$ , a walker either jumps to the right with probability  $p$  or to the left with probability  $q = 1 - p$ . What is the probability  $\rho(x, N)$  that the walker is at position  $x$  after  $N$  steps? The probability the walker makes  $n$  jumps to the right and  $N - n$  jumps to the left is

$$\rho(n, N) = \binom{N}{n} p^n q^{N-n} \quad \text{and} \quad x = n\Delta x - (N - n)\Delta x \quad (32)$$

Therefore, the average position of the walker and its variance after  $N$  step and hence the probability, for large  $N$ , are

$$\langle x \rangle = N\Delta x(p - q) \quad \text{and} \quad \sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 = 4N(\Delta x)^2 pq \quad (33)$$

$$\rho(x, N) \approx \frac{1}{\sqrt{2\pi\sigma^2}} e^{(x - \langle x \rangle)/(2\sigma^2)} \quad (34)$$

To connect the above expression to *free diffusion*, we use the equations in (33),

$$\sigma^2 = 4Npq(\Delta x)^2 = 4pq \frac{(\Delta x)^2}{\Delta t} N\Delta t \equiv 2D\Delta t \quad (35)$$

$$\text{where, } D = 2pq \frac{(\Delta x)^2}{\Delta t} \quad (36)$$

$$\langle x \rangle = (p - q)N\Delta x = (p - q) \frac{\Delta x}{\Delta t} N\Delta t \equiv v\Delta t \quad (37)$$

$$\text{where, } v = (p - q) \frac{\Delta x}{\Delta t} \quad (38)$$

where  $D$  is the *diffusion coefficient* and  $v$  is *drift velocity*. Suppose the walker arrives at position  $x$  in  $N + 1$  steps either from the position to right or to the left with respective probabilities  $p$  and  $q$ ,

$$\rho(x, N + 1) = p\rho(x - \Delta x, N) + q\rho(x + \Delta x, N) \quad (39)$$

The above equation (39) is called the *master equation* for a stochastic process which gives a recursive relation for the evolution of probability distribution. From (38) we note that

$$\frac{\Delta x}{\Delta t} = v + 2q \frac{\Delta x}{\Delta t} = 2p \frac{\Delta x}{\Delta t} - v \quad (40)$$

Subtracting  $\rho(x, N)$ , dividing with  $\Delta t$  and using the (38) yields,

$$\begin{aligned}\frac{\rho(x, N+1) - \rho(x, N)}{\Delta t} &= -p \frac{\Delta x}{\Delta t} \frac{\rho(x, N) - \rho(x - \Delta x, N)}{\Delta x} + q \frac{\Delta x}{\Delta t} \frac{\rho(x + \Delta x, N) - \rho(x, N)}{\Delta x} \\ \frac{\rho(x, N+1) - \rho(x, N)}{\Delta t} &= -vp \frac{\rho(x, N) - \rho(x - \Delta x, N)}{\Delta x} - vq \frac{\rho(x + \Delta x, N) - \rho(x, N)}{\Delta x} \\ &\quad + 2pq \frac{(\Delta x)^2}{\Delta t} \frac{\rho(x + \Delta x, N) + \rho(x - \Delta x, N) - 2\rho(x, N)}{(\Delta x)^2}\end{aligned}$$

In the limit  $\Delta x \rightarrow 0$ ,  $\Delta t \rightarrow 0$ ,

$$\frac{\partial \rho(x, t)}{\partial t} = -v \frac{\partial \rho(x, t)}{\partial x} + D \frac{\partial^2 \rho(x, t)}{\partial x^2} \quad (41)$$

The above equation (41) is known as *Fokker-Planck* equation. One of the solutions of Fokker-Planck equation is the p.d.f. (34) where  $\sigma^2$  and  $\langle x \rangle$  is given in the equations (35) and (37).

$$\rho(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{(x-vt)/4Dt} \quad (42)$$

The Fokker-Planck equation is also a continuity equation, in this case for probability  $\rho(x, t)$ ,

$$\frac{\partial \rho(x, t)}{\partial t} = \frac{\partial j}{\partial x} \quad \text{where, } j = -v\rho(x, t) + D \frac{\partial \rho(x, t)}{\partial x} \quad (43)$$

which when generalized to 3-dimension we get

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} = \nabla \cdot \mathbf{j} \quad \text{where, } \mathbf{j} = -\mathbf{v}\rho(\mathbf{r}, t) - D \nabla \rho(\mathbf{r}, t) \quad (44)$$

Consider the following function  $G(x, y; t)$ , as a function of  $y$  and  $t$ , keeping  $x$  fixed. It satisfies the diffusion equation (41) at  $t > 0$  and in the limit  $t \rightarrow 0$ ,

$$G(x, y; t) = \frac{1}{\sqrt{4\pi Dt}} e^{(x-y)^2/4Dt} \quad (45)$$

$$G(x, y; t) \rightarrow \delta(x - y) \quad \text{for } t \rightarrow 0 \quad (46)$$

$$\int dy G(x, y; t) = 1 \quad (47)$$

where in the last expression we have taken  $x$  to be some arbitrary initial state. The  $G(x, y; t)$  is called the Green's function of the diffusion equation (41) with  $v = 0$ . This function is used to determine the time evolution of  $\rho(x, 0) \rightarrow \rho(x, t)$ ,

$$\rho(y, t) = \int dx G(x, y; t) \rho(x, 0) \quad (48)$$

Let us solve the diffusion equation (41) by Fourier transform,

$$\begin{aligned}u(x, t) &= \int dk u(k, t) e^{ikx} \Rightarrow \frac{1}{D} \frac{\partial u(k, t)}{\partial t} = -k^2 u(k, t) \\ \Rightarrow u(k, t) &= u(k, t = t') e^{-Dk^2(t-t')}\end{aligned} \quad (49)$$

Assuming a point source at  $t = t'$ ,  $u(x, t = t') = \delta(x - x')$ , then

$$\begin{aligned} u(k, t = t') &= \frac{1}{2\pi} \int dx \delta(x - x') e^{-ikx} = \frac{1}{2\pi} e^{-ikx'} \\ \text{for } t > t' \Rightarrow u(x, t) &= \frac{1}{2\pi} \int dk e^{-ik(x-x')} e^{-Dk^2(t-t')} \end{aligned} \quad (50)$$

The integral can be done by completing the square and the result is the point source solution  $u(x, t)$ , which is actually the Green's function,

$$G(x, x'; t, t') = \frac{1}{\sqrt{4\pi D(t-t')}} e^{-(x-x')^2/4D(t-t')} \quad (51)$$

Compare this form with (45) for  $t' = 0$ .

Every new positions of a walker are elements of Markov chain and the Green function gives the transition probability of Markov process  $T(x \rightarrow y) \equiv G(x, y; \Delta t)$ .

The general form of diffusion equation (41) is

$$\frac{\partial \rho(x, t)}{\partial t} = \mathcal{L} \rho(x, t) \Rightarrow \rho(x, t) = e^{\mathcal{L}t} \rho(x, 0) \text{ hence, } G(x, y; t) = \langle x | e^{\mathcal{L}t} | y \rangle \quad (52)$$

Consider Schrödinger equation in imaginary time  $\tau = it$ ,

$$i \frac{\partial |\psi\rangle}{\partial t} = H |\psi\rangle \Rightarrow \frac{\partial |\psi\rangle}{\partial \tau} = -H |\psi\rangle \quad (53)$$

The imaginary time Schrödinger equation is a particular form of diffusion equation

$$\frac{\partial \rho}{\partial \tau} = \frac{1}{2} \frac{\partial^2 \rho(x, \tau)}{\partial x^2} - V(x) \rho(x, \tau) \quad (54)$$

The formal solution to this imaginary time Schrödinger equation is

$$|\psi(\tau + \Delta\tau)\rangle = e^{-H\Delta\tau} |\psi(\tau)\rangle \quad (55)$$

If we express  $|\psi\rangle$  in terms of *energy ordered* eigenstates of the Hamiltonian  $|\phi_i\rangle$ , then

$$\begin{aligned} |\psi(\tau)\rangle &= \sum_i c_i |\phi_i(\tau)\rangle \Rightarrow |\psi(\tau + \Delta\tau)\rangle = \sum_i c_i e^{-\epsilon_i \Delta\tau} |\phi_i(\tau)\rangle \\ \lim_{\Delta\tau \rightarrow \infty} |\psi(\tau)\rangle &\approx c_0 e^{-\epsilon_0 \Delta\tau} |\phi_0(\tau)\rangle \end{aligned} \quad (56)$$

This is an interesting observation – imaginary time evolution results in excited states decaying exponentially fast. This contamination of excited states to the ground state in variational MC makes the energy estimate larger. The diffusion MC is a realization of the above equation (56) in position space,

$$\lim_{\tau \rightarrow \infty} \psi(\mathbf{r}, \tau) = c_0 e^{-\epsilon_0 \tau} \phi_0(\mathbf{r}) \quad (57)$$

In this form it looks mighty similar to the diffusion equation (52).

Consider again the equation (55) and use  $H = K + V$  where  $K = p^2/2$  is the kinetic energy operator and  $V$  potential energy. For  $\Delta\tau$  small, we can neglect Baker-Campbell-Hausdorff relation and write

$$e^{-\Delta\tau(K+V)} = e^{-K\Delta\tau} e^{-V\Delta\tau} + \mathcal{O}(\Delta\tau^2) \quad (58)$$

However, it might look profitable to split the exponential as,

$$e^{-\Delta\tau(K+V)} = e^{-V\Delta\tau/2} e^{-K\Delta\tau} e^{-V\Delta\tau/2} + \mathcal{O}(\Delta\tau^3) \quad (59)$$

but, because diffusion steps are carried out successively, the first term and the last term are calculated at different time  $\tau$ . To find the Green's function, we need to find the matrix element of this exponential operator on the right hand side, only for  $K$  since  $V$  is just a function of  $x$  and yields a delta-function,  $\langle x|V|y\rangle = \delta(x-y)$ ,

$$\begin{aligned} G_{\text{kin}}(x, y; \Delta\tau) &= \langle x|e^{-\Delta\tau \hat{p}^2/2}|y\rangle = \int dp \langle x|p\rangle \langle p|e^{-\Delta\tau \hat{p}^2/2}|y\rangle \\ &= \int dp dp' \langle x|p\rangle \langle p|e^{-\Delta\tau \hat{p}^2/2}|p'\rangle \langle p'|y\rangle \\ &= \int dp dp' \langle x|p\rangle e^{-\Delta\tau p^2/2} \langle p|p'\rangle \langle p'|y\rangle \\ &= \int dp dp' \langle x|p\rangle e^{-\Delta\tau p^2/2} \delta(p-p') \langle p'|y\rangle \\ &= \frac{1}{2\pi} \int dp e^{-\Delta\tau p^2/2} e^{ipx} e^{-ipy} = \frac{1}{2\pi} \int dp e^{-\Delta\tau p^2/2} e^{ip(x-y)} \end{aligned} \quad (60)$$

where we have used  $\langle x|p\rangle = \exp(ipx)/\sqrt{2\pi}$ . Completing the square, the expression in the exponential in the above equation (60) and doing the Gaussian integral, we get

$$G_{\text{kin}}(x, y; \Delta\tau) = \frac{1}{\sqrt{2\pi\Delta\tau}} e^{(x-y)^2/(2\Delta\tau)} \quad (61)$$

Comparing this with the expression for Green's function in (45), the kinetic part of Schrödinger Green's function is equal to the Green's function of the simple diffusion equation. If the potential term is included, the full Green's function reads

$$G(x, y; \Delta\tau) = G_{\text{kin}}(x, y; \Delta\tau) e^{-V\Delta\tau} + \mathcal{O}(\Delta\tau^2) \quad (62)$$

The problem with this definition of Green's function is, because of the presence of  $V$ , the normalization (47) is lost. This is addressed by multiplying Green's function in (62) by a presently unknown prefactor  $\exp(E_T\tau)$  to normalize it back again. In that case, the original Hamiltonian modifies to,

$$H = K + V \rightarrow H' = K + V - E_T = -\frac{1}{2} \frac{\partial^2}{\partial x^2} + [V(x) - E_T] \quad (63)$$

Hence, the imaginary time Schrödinger equation in (53) reads,

$$\frac{\partial|\psi\rangle}{\partial\tau} = \frac{1}{2} \frac{\partial^2|\psi\rangle}{\partial x^2} - [V(x) - E_T]|\psi\rangle \quad (64)$$

$$\frac{\partial\psi(x, \tau)}{\partial\tau} = \frac{1}{2} \frac{\partial^2\psi(x, \tau)}{\partial x^2} - [V(x) - E_T]\psi(x, \tau) \quad (65)$$

where in the last expression we have written the imaginary time Schrödinger equation in terms of the wavefunctions. The potential has been shifted by  $E_T$  and, if  $E_T$  is so chosen that Green's function is normalized, then the equation above (64) describes a Markov process and have an invariant distribution *i.e.*  $G(x, y; \tau)$ . For stationary distribution, it reduces to stationary Schrödinger equation,

$$-\frac{1}{2} \frac{\partial^2|\psi\rangle}{\partial x^2} - V(x)|\psi\rangle = E_T|\psi\rangle \quad (66)$$



This equation can be used for Monte Carlo methods but it is rather inefficient without importance sampling. The condition is similar to variational MC but going to be far more complicated than just inserting  $p(r)$  of (4) in the expectation values (5). Define a *trial* or a *guiding* wavefunction  $|\psi_g\rangle \rightarrow \psi_T(x)$ , such that the new wavefunction is,

$$f(x, \tau) = \psi_T(x) \psi(x, \tau) \quad (67)$$

and plugging it in (64), we get

$$\begin{aligned} \frac{\partial \psi(x, \tau)}{\partial \tau} &= \frac{1}{\psi_T(x)} \frac{\partial f(x, \tau)}{\partial \tau} \\ \frac{\partial \psi(x, \tau)}{\partial x} &= \frac{1}{\psi_T(x)} \frac{\partial f(x, \tau)}{\partial x} - \frac{1}{\psi_T^2(x)} \frac{\partial \psi_T(x)}{\partial x} f(x, \tau) \\ \frac{\partial^2 \psi(x, \tau)}{\partial x^2} &= -\frac{1}{\psi_T^2(x)} \frac{\partial \psi_T(x)}{\partial x} \frac{\partial f(x, \tau)}{\partial x} + \frac{1}{\psi_T(x)} \frac{\partial^2 f(x, \tau)}{\partial x^2} \\ &\quad - \frac{1}{\psi_T^2(x)} \frac{\partial^2 \psi_T(x)}{\partial x^2} f(x, \tau) - \frac{1}{\psi_T^2(x)} \frac{\partial \psi_T(x)}{\partial x} \frac{\partial f(x, \tau)}{\partial x} + \frac{2}{\psi_T^3(x)} \left( \frac{\partial \psi_T(x)}{\partial x} \right)^2 f(x, \tau) \\ &= \frac{1}{\psi_T(x)} \frac{\partial^2 f(x, \tau)}{\partial x^2} - \frac{1}{\psi_T^2(x)} \frac{\partial^2 \psi_T(x)}{\partial x^2} f(x, \tau) - \frac{2}{\psi_T^2(x)} \frac{\partial \psi_T(x)}{\partial x} \frac{\partial f(x, \tau)}{\partial x} \\ &\quad + \frac{2}{\psi_T^3(x)} \left( \frac{\partial \psi_T(x)}{\partial x} \right)^2 f(x, \tau) \\ &= \frac{1}{\psi_T(x)} \frac{\partial^2 f(x, \tau)}{\partial x^2} - \frac{2}{\psi_T(x)} \frac{\partial}{\partial x} \left( \frac{1}{\psi_T(x)} \frac{\partial \psi_T(x)}{\partial x} f(x, \tau) \right) + \frac{1}{\psi_T^2(x)} \frac{\partial^2 \psi_T(x)}{\partial x^2} f(x, \tau) \\ &= \frac{1}{\psi_T(x)} \frac{\partial^2 f(x, t)}{\partial x^2} - \frac{2}{\psi_T(x)} \frac{\partial}{\partial x} [F(x) f(x, \tau)] + \frac{1}{\psi_T^2(x)} \frac{\partial^2 \psi_T(x)}{\partial x^2} f(x, \tau) \end{aligned} \quad (68)$$

where  $F(x) = \partial(\log \psi_T(x))/\partial x$  is called the *drift* or *quantum force*. Putting everything back in the imaginary time Schrödinger equation (65), we get

$$\frac{\partial f(x, t)}{\partial \tau} = \frac{1}{2} \frac{\partial^2 f(x, \tau)}{\partial x^2} - \frac{\partial}{\partial x} [F(x) f(x, \tau)] - [E_L(x) - E_T(x)] f(x, t) \quad (69)$$

$$\begin{aligned} E_L(x) &= V(x) - \frac{1}{2\psi_T(x)} \frac{\partial^2 \psi_T(x)}{\partial x^2} \\ &= \frac{1}{\psi_T(x)} \left[ -\frac{1}{2} \frac{\partial^2 \psi_T(x)}{\partial x^2} + V(x) \psi_T(x) \right] \equiv \frac{1}{\psi_T} H \psi_T \end{aligned} \quad (70)$$

where  $E_L(x)$  is the *local energy* that includes the  $V(x)$  and the last term of (68), and obviously of the form (5). Often, in the definition of force  $F(x)$  a factor of 2 is included so that the kinetic term can be written as

$$F(x) = \frac{2}{\psi_T(x)} \frac{\partial \psi_T(x)}{\partial x} \rightarrow K = \frac{1}{2} \frac{\partial}{\partial x} \left[ \frac{\partial}{\partial x} - F(x) \right] \rightarrow \frac{1}{2} \hat{p} [\hat{p} - F(\hat{x})] \quad (71)$$

Hence, the kinetic part of the Schrödinger Green's function in (61) modified to

$$G_{\text{kin}}(x, y; \Delta\tau) = \frac{1}{\sqrt{2\pi\Delta\tau}} e^{(x-y-F(x)\Delta\tau/2)^2/(2\Delta\tau)} \quad (72)$$

This Green's function is normalized and can, therefore, be used for constructing a Markov chain. Adding the potential part does not destroy the normalization because of  $E_T(x)$  term,

$$\begin{aligned} G(x, y; \Delta\tau) &= G_{\text{kin}}(x, y; \Delta\tau) e^{-\Delta\tau[E_L(x) - E_T(x)]} + \mathcal{O}(\Delta\tau^2) \\ &= \frac{1}{\sqrt{2\pi\Delta\tau}} e^{(x-y-F(x)\Delta\tau/2)^2/(2\Delta\tau)} e^{-\Delta\tau[E_L(x) - E_T(x)]} + \mathcal{O}(\Delta\tau^2) \end{aligned} \quad (73)$$

To understand the role of  $E_T$ , apart from helping Green's function to normalize, consider the time-dependent solution of the Schrödinger equation (65) using the *energy ordered* eigenstates (56) we get

$$\psi(x, \tau) = \sum_i c_i \phi_i(x) e^{-(\epsilon_i - E_T)\tau} \quad (74)$$

Since we demanded  $\epsilon_0 < \epsilon_1 < \dots < \epsilon_i < \dots$ , we infer that,

- if  $E_T > \epsilon_0$ , the wavefunction  $\psi(x, \tau)$  diverges.
- if  $E_T < \epsilon_0$ , the wavefunction  $\psi(x, \tau)$  vanishes.
- if  $E_T \approx \epsilon_0$ , the wavefunction  $\psi(x, \tau) \approx c_0 \phi_0$ .

Hence, we should perform an imaginary time evolution, by clever choice of  $E_T$ , until we find the ground state. In light of this, we write the Schrödinger equation in (69) in shorthand as

$$-\frac{\partial f}{\partial \tau} = [K + D + B] f \quad (75)$$

where  $K$  is the kinetic or free diffusion term,  $D$  is the drift term involving the the drift force  $F(x)$  and  $B$  is the *branching* term. The Green's functions of each individual term are,

$$\begin{aligned} G_K(x, y, \tau) &\sim \exp\left[-\frac{(x-y)^2}{2\tau}\right] \\ G_D(x, y, \tau) &\sim \delta(x - \mathcal{X}(x)), \quad \text{where } \frac{\mathcal{X}(x)}{d\tau} = F(x) \\ G_B(x, y, \tau) &\sim \exp[E_L(x) - E_T] \delta(x - y) \end{aligned} \quad (76)$$

Finally, the diffusion monte carlo steps are,

1. Put  $N_w$  *walkers* at random points in the configuration space. A *walker* is defined by the positions of all the particles of the system  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  i.e.  $f(\mathbf{x}, \tau)$  at some initial  $\tau$ ,

$$f(\mathbf{x}, \tau) = \sum_i^{N_w} \delta(\mathbf{x} - \mathbf{x}_i(\tau))$$

2. The first step is free diffusion governed by  $G_K$ . Select a walker and perform the shift,

$$\mathbf{x}' = \mathbf{x} + \boldsymbol{\eta} \sqrt{\Delta\tau}$$

where  $\boldsymbol{\eta}$  is a normalized gaussian random vector drawn from  $G_K$ .

3. In the second step, update  $\mathbf{x}$  further using the quantum or drift force  $F(x)$  given in (68),

$$\mathbf{x}'' = \mathbf{x}' + \frac{\Delta\tau}{2} F(\mathbf{x}')$$

4. In the third *i.e.* the branching step, calculate the local energy  $E_L(x)$  as in (70) and evaluate

$$q = \exp \left[ - \Delta \tau (E_L(\mathbf{x}') + E_L(\mathbf{x})) / 2 - E_T \right]$$

The branching process duplicates the walkers that best mimics the ground state distribution or kill walkers that are far from the desired solution. If  $q < 1$ , the walker survives with probability  $q$  and dies with probability  $1 - q$ . If  $q > 1$ , the walker gives birth to either  $[q - 1]$  or  $[q]$  new ones at  $\mathbf{x}$ , where  $[q]$  indicates the integer part of  $q$ . This is implemented in the next step.

5. The final step is the Metropolis accept / reject step by choosing a uniform random number  $r \in [0, 1]$  and defining  $s = q + r$ . Eliminate the walker (or the configuration) when  $[s] = 0$  or create  $[s]$  new ones at  $\mathbf{x}'$  otherwise.
6. Update  $E_T$  and go to step 2.

The last point is how to choose  $E_T$ . The reference energy  $E_T$  is chosen to reduce the fluctuations in the number of walkers and keep the population size in the desired or target range, say  $M$ . When the number of walkers remains statistically constant,  $E_T$  is an estimator of the ground state. Suppose, after last branching, the number of walkers is  $\tilde{M}$ . Then adjust  $E_T$  as

$$E_T = \alpha \ln \left( \frac{M}{\tilde{M}} \right) + \tilde{E}_0 \tag{77}$$

where  $\alpha$  is an adjustable parameter use for modifying  $E_T$ .