Quantum Monte Carlo

Hannah Baek, Tobias Tian
University of British Columbia
December 8th, 2025

Abstract

This paper explores the practicality and applicability of the various Quantum Monte Carlo methods. This report starts by introducing the key properties and ideas of Monte Carlo methods. It then progresses on how Variational Monte Carlo methods approximate ground-state energies, with a focus on how trial wavefunctions are selected and optimized, while highlighting the sampling algorithm that captures the essence of Monte Carlo. Furthermore, it discusses two more QMC methods, Diffusion Monte Carlo, and Path Integral Monte Carlo, and how they can be leveraged to approximate physical equilibria as well. Finally, this paper briefly highlights how the latter two methods suffer from a common fermion sign problem, and how the discussed QMC methods compare and contrast.

1 Introduction

Quantum Monte Carlo (QMC) method is a class of classical computational algorithm that leverages the Monte Carlo method to simulate and approximate the ground-state and thermodynamic equilibria of many-body physical systems. Its essence lies within the fact that one does not need to precisely understand how high-dimensional wavefunctions behave as long as it is possible to sample such wavefunctions so that an approximation to the density distribution is accessible.
In Section 2, we introduce the Monte Carlo method through a classic example of estimating π and summarizing its core idea. In Section 3, we highlight how Variational Monte Carlo can be used to estimate such ground-state energies, by introducing relevant quantum physical concepts, and explaining the complete scope of steps in detail, while providing a code sample that calculates the electronic ground-state energy of a fixed-nucleus helium atom using VMC. In Section 4 and 5, We explore the other two most-used QMC methods, and highlight key differences among them. Finally, in Section 6, we discuss briefly about how the fermion sign problem affects QMC methods, and some conceptual ways to circumvent the problem.

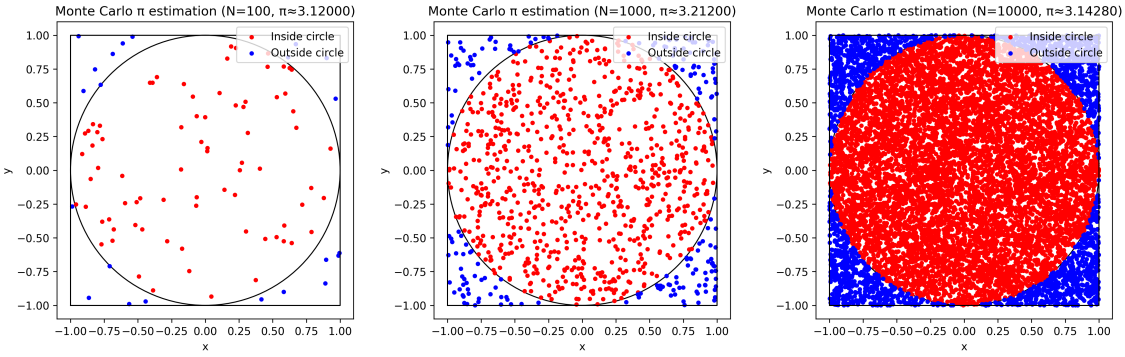
2 Monte Carlo Method

Monte Carlo method is a class of classical computational algorithm which aims to find a numerical value of interest through iterative random sampling according to a probability distribution. Such method is often applied when analytical solutions are too complex to solve. A simple case in which Monte Carlo method can be used is to approximate the value of π with only the knowledge of how to calculate the area of a circle. The process is outlined as such [KW08]:

1. Consider a square with edge length of 2, and centered at $(0, 0)$.
2. Then, draw the inscribed circle of such a square which has radius 1, also centered at $(0, 0)$.

3. **Uniformly at random** select a point within the square.
4. Calculate the ratio of the number of points within the circle and the total number of points sampled. This should approximate $\frac{\pi}{4}$.
5. Repeat Step 3 and 4 until a certain threshold. Calculate the approximated value of π by multiplying the ratio stated in Step 4 by 4.

By applying this process, we can obtain the following result (link to code, see A.1):



The key idea from this demonstration is that random sampling can be used to evaluate a definite integral [KW08]:

$$I = \int_0^1 \int_0^{\sqrt{1-x^2}} \mathrm{d}y\mathrm{d}x$$

From the above demonstration, it is observed that the main goal of Monte Carlo method is to solve a certain mathematical problem with the core process to be iterated random sampling, where each additional iteration gives more accurate results.

3 Variational Monte Carlo

As the number of interacting particles in a given system scale, analytically solving the Schrödinger's equation of the entire system while incorporating the interaction of particles within becomes intractable.In this case, the Variational Monte Carlo (abbreviated as VMC) method provides a practical approximation to the ground state energy.

3.1 Main Process of VMC

To begin with, consider the following relevant physical definitions.

Definition 3.1. *Consider a general quantum system, denoted mathematically as a Hilbert space $\mathcal{H}$, with N particles. Each particle i has 3-dimensional position vector $\mathbf{r}_i$. Define a $3N$-dimensional vector $\mathbf{R} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$. A general quantum state $|\Psi\rangle$ in this system can be denoted as:*

$$|\Psi\rangle = \int \Psi(\mathbf{R}) \, |\mathbf{R}\rangle \, \mathrm{d}\mathbf{R} \tag{1}$$

where $\Psi(\mathbf{R})$ is the position-space wavefunction, which satisfies the following properties:

1. $|\Psi(\mathbf{R})|^2 \geq 0$
2. $\int \Psi^*(\mathbf{R})\Psi(\mathbf{R})\mathrm{d}\mathbf{R} = 1$

Definition 3.2. *An **observable** $\mathcal{O}$ is a Hermitian operator, where measuring $\mathcal{O}$ mathematically amounts to doing a projective measurement with respect to the orthonormal basis of eigenvectors of $\mathcal{O}$.*

Definition 3.3. *Given an observable $\mathcal{O}$ and a quantum state $|\Psi\rangle$, the **expected value** of this observable with respect to this state is*

$$\langle\mathcal{O}\rangle = \langle\Psi|\mathcal{O}|\Psi\rangle \tag{2}$$

*and the **variance** of the above observable and the above state is:*

$$(\Delta\mathcal{O})^2 = \left\langle\mathcal{O}^2\right\rangle - (\langle\mathcal{O}\rangle)^2 \tag{3}$$

To characterize the time evolution of the quantum state, the observable $\hat{H}$ is introduced to give the following.

Definition 3.4. *The **time-dependent Schrödinger's equation** of a given quantum state $|\Psi\rangle$ that is dependent on time t is*

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle \tag{4}$$

where i *is the imaginary unit, $\hbar$ is Planck's constant, $\hat{H}$ is the Hamiltonian operator for the quantum state $|\Psi\rangle$.*

In the context of VMC, the quantity of interest is the ground state energy. After the system reaches this equilibrium, the Hamiltonian can be considered to be time-independent. Thus, from 3.4

Definition 3.5. *If the Hamiltonian $\hat{H}$ is time-independent, it induces the **time-independent Schrödinger's equation** for $|\Psi\rangle$ to be:*

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \tag{5}$$

From Eq. (2) and Eq. (5), it follows that

$$\left\langle\hat{H}\right\rangle = E \tag{6}$$

which indicates that, for an eigenstate $|\Psi\rangle$ of the Hamiltonian $\hat{H}$, the expected value of $\hat{H}$ is the eigenvalue of $|\Psi\rangle$, which physically represents the energy of the system when it is in the state of $|\Psi\rangle$.
If the ground state wavefunction is denoted to be $\Psi_0(\mathbf{R})$, it follows from Eq. (6), the ground state energy E_0 is:

$$\begin{aligned} E_0 &= \langle\Psi_0|\hat{H}|\Psi_0\rangle \\ &= \int \Psi_0^*(\mathbf{R})\hat{H}\Psi_0(\mathbf{R})\mathrm{d}\mathbf{R} \end{aligned}$$

To approximate this energy, a trial wavefunction Ψ_T is first selected, satisfying certain conditions [FMNR01]:

1. Ψ_T and $\boldsymbol{\nabla}\Psi_T$ must be continuous wherever the potential is finite.
2. The integrals $\int \Psi_T^*\Psi_T$, $\int \Psi_T^*\hat{H}\Psi_T$, and $\int \Psi_T^*\hat{H}^2\Psi_T$ all must exist

With this trial wavefunction, the expected value of the Hamiltonian is calculated to be [FMNR01] [Aci97] [KW08]:

$$E_V = \frac{\int \Psi_T^*(\mathbf{R})\hat{H}\Psi_T(\mathbf{R})\mathrm{d}\mathbf{R}}{\int |\Psi_T(\mathbf{R})|^2 \mathrm{d}\mathbf{R}} \geq E_0 \tag{7}$$

To use VMC, sampling needs to be conducted according to a certain probability distribution. The trial energy E_V can be considered as an expected value of all local energies E_{loc}. From Eq. (5), energy of the system in state $|\Psi\rangle$ can also be calculated as:

$$E = \frac{\hat{H}\,|\Psi\rangle}{|\Psi\rangle}$$

Therefore, each local energy sampled from $\mathbf{R}$ is evaluated to be [KW08]:

$$E_{\mathrm{loc}}(\mathbf{R}) = \frac{\hat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \tag{8}$$

This gives the probability density for sampling to naturally be [KW08]:

$$f(\mathbf{R}) = \frac{|\Psi_T(\mathbf{R})|^2}{\int |\Psi_T(\mathbf{R})|^2 \mathrm{d}\mathbf{R}} \tag{9}$$

From Eq. (8) and Eq. (9), Eq. (7) becomes:

$$E_V = \int f(\mathbf{R}) E_{\mathrm{loc}}(\mathbf{R})\mathrm{d}\mathbf{R} \tag{10}$$

which can be approximated by [FMNR01][Aci97]:

$$E_V \approx \frac{1}{M}\sum_{m=1}^{M} E_{\mathrm{loc}}(\mathbf{R}_m) \tag{11}$$

The essential steps of VMC have been established, showing how the ground-state energy can be estimated through Monte Carlo sampling. However, in this process, two key challenges remain: the choice and optimization of the trial wavefunction and the sampling of the configuration $\mathbf{R}$. These two aspects will be discussed in the following sections.

3.2 Selection and Optimization of the Trial Wavefunction

In this section, we focus on zero-temperature VMC applied to electronic systems, where the wavefunction describes N interacting electrons. Choosing an accurate trial wavefunction Ψ_T is crucial: it determines not only the variational energy E_V but also the statistical efficiency of the Monte Carlo sampling. The purpose of Ψ_T is to provide the best possible approximation to the true ground state while remaining mathematically well-behaved and computationally tractable.

3.2.1 Motivation: Why Good Trial Wavefunctions Matter

Although VMC is a powerful many-body method, its accuracy depends heavily on the quality of Ψ_T. Most ground-state calculations in electronic structure theory instead use less expensive methods such as Hartree–Fock (HF) or Kohn–Sham density functional theory (DFT). These methods

scale favorably and are efficient to solve, but they have well-known failures, HF neglects correlation entirely, resulting in limited predictive power. These limitations are further discussed in Ref. [FMNR01], Sec. II C.
Quantum Monte Carlo techniques successfully overcome many of these failures. Continuum QMC calculations can already achieve "chemical accuracy" (~ 1 kcal/mol ≈ 0.004 eV per molecule) for small systems and show no fundamental loss of accuracy for larger systems (Ref. [FMNR01], Sec. I). The computational cost of fermionic VMC and DMC scales approximately as N^3 with the number of electrons N, making simulations of moderately large systems feasible. Furthermore, QMC algorithms are naturally parallelizable, enabling efficient use of modern high-performance computing resources.

3.2.2 Structure of a Useful Trial Wavefunction

A common and highly effective ansatz for electronic systems is the *Slater–Jastrow* form:

$$\Psi_T(\mathbf{R};\alpha) = \Phi(\mathbf{R})\, J(\mathbf{R};\alpha),$$

where $\mathbf{R} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$ denotes the electron coordinates. This is the standard starting point for VMC of electronic systems because it combines an antisymmetric Slater determinant with a Jastrow factor that captures the electron-electron correlation.

Definition 3.6. *The Slater determinant that captures the antisymmetric part of an electron system takes the form of:*

$$\Phi(\mathbf{R}) = \det\left[\phi_i(\mathbf{r}_j)\right]$$

which enforces the required antisymmetry of a fermionic wavefunction. Each orbital ϕ_i is the wavefunction of a single electron independently moving in a self-consistent field generated by the other electrons. To give extra clarification, the determinant is taken over a matrix in the form:

$$\det(\mathcal{D}(\mathbf{R})) = \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_1(\mathbf{r}_2) & \cdots & \phi_1(\mathbf{r}_N) \\ \phi_2(\mathbf{r}_1) & \phi_2(\mathbf{r}_2) & \cdots & \phi_2(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(\mathbf{r}_1) & \phi_2(\mathbf{r}_2) & \cdots & \phi_N(\mathbf{r}_N) \end{vmatrix} \tag{12}$$

where i represents the orbital, and j represents the electron.
A high-quality Slater determinant within VMC must be built from orbitals that respect the symmetries of the Hamiltonian, remain orthonormal for numerical stability, and generate reasonable nodal surfaces, since the behavior of Ψ_T near its nodes strongly influences the stability and accuracy of the variational estimate. It is also desirable for the determinant to remain compact so that computing $\Phi(\mathbf{R})$ and its derivatives during the Monte Carlo sampling procedure is computationally efficient. In practice, Hartree–Fock orbitals naturally satisfy these conditions and therefore form a reliable foundation for constructing the Slater determinant [FMNR01].

Definition 3.7. *The Jastrow factor explicitly introduces electron-electron correlation through*

$$J(\mathbf{R};\alpha) = \exp\left[\sum_{i<j} u(\mathbf{r}_{ij};\alpha)\right],$$

where $u(\mathbf{r}_{ij};\alpha)$ is a function of the distance between two electrons, with a set of tuning parameters α that control how strongly the electrons correlate with one another.

The Jastrow term enforces the correct electron–electron cusp behaviour, accounts for short-range correlation holes, and captures long-range effects such as dispersion that Hartree–Fock fails to describe. By including these Coulomb-driven correlation effects, the Jastrow factor lowers the variational energy and reduces fluctuations in the local energy while leaving the antisymmetry of the Slater determinant unchanged.
The choice of $u(\mathbf{r}_{ij};\alpha)$ depends on the physical system and on the types of particle-particle correlations that must be captured. Simple analytic forms that satisfy the electron–electron cusp condition are often sufficient for small atoms, while more flexible parametrizations are used for larger or more strongly correlated systems.
In our VMC simulation of helium (see Appendix B.1), we select a two-body Jastrow factor whose initial parameters follow directly from the cusp condition and from basic physical intuition about the typical separation between the two electrons. In Section 3.2.4, we describe how these parameters are optimized by minimizing the variance of the local energy.

3.2.3 Selecting the Slater–Jastrow wavefunction.

Constructing a good trial wavefunction therefore involves: (1) performing a Hartree–Fock (or DFT) calculation to generate orbitals with the appropriate symmetry and nodal structure; (2) building the Slater determinant from these orbitals; and (3) choosing a Jastrow factor with the correct cusp conditions and flexible correlation terms. The resulting Slater–Jastrow wavefunction provides a balanced and computationally efficient starting point for subsequent parameter optimization within VMC.

3.2.4 Optimization of Variational Parameters

The trial wavefunction $\Psi_T(\mathbf{R};\alpha)$ contains adjustable parameters α, primarily in the Jastrow factor, that determine the strength and range of the correlation terms. To obtain the most accurate and statistically efficient wavefunction, these parameters must be optimized. In this work we focus on *variance minimization*, which adjusts α so that the fluctuations of the local energy,

$$E_{\mathrm{loc}}(\mathbf{R};\alpha) = \frac{\hat{H}\Psi_T(\mathbf{R};\alpha)}{\Psi_T(\mathbf{R};\alpha)},$$

are minimized over configurations sampled from $|\Psi_T|^2$. For the exact ground state the local energy is constant, so a lower variance directly indicates a trial wavefunction that is closer to the true eigenstate. Variance minimization is therefore a stable and widely used approach.
In practice, the optimization proceeds by estimating the derivatives of the variance with respect to each parameter α_k using configurations drawnfrom an initial Monte Carlo run. These derivatives provide a stochastic gradient that indicates how the parameters should be updated to reduce the variance. Repeating this process iteratively yields progressively improved values of α. The specific update rule used in our implementation is described in Appendix B.2, where we apply this procedure to optimize the Jastrow parameters for the helium VMC calculation.

3.3 Sampling Method

In the previous section, it is discussed how a trial wavefunction can be selected and optimized. Yet, it is still unclear how a configuration $\mathbf{R}$ is sampled such that it satisfies the probability density in Eq. (9).
Naïvely, one should sample according to this probability density function. However, it should be

noted that the integral present in the denominator of Eq. (9) becomes computationally implausible to calculate or normalize when the dimension of $\mathbf{R}$ scales, because as stated before, configuration $\mathbf{R}$ is a $3N$-dimensional vector. Furthermore, uniformly sampling in a $3N$-dimensional space suffers from the "curse of dimensionality" where if only a given number of data points are being sampled, the samples would be so sparse that they cannot represent the population well [MRR+53]. Furthermore, depending on the form of the selected trial wavefunction, the wavefunction can take different shapes and correlations which further complicates a naïve sampling method.
Hence, Metropolis et al. developed an algorithm that accepts or rejects a configuration starting from an initial configuration using only probability ratios [MRR+53]. This algorithm ensures that the final list of accepted configurations follows the same distribution as the probability density distribution [MRR+53].
Analogous to the paper in which Metropolis et al. demonstrated how they sampled according to the Boltzmann distribution, the following steps is taken to sample the selected trial wavefunction [MRR+53]:

1. Place the N particles in any configuration, and denote it as $\mathbf{R}_1$. Calculate $|\Psi_T(\mathbf{R}_1)|^2$.
2. Since the initial configuration is a vector that contains the Cartesian positions of N particles, it is possible to select one particle, say particle i, who has the initial position of (x_i, y_i, z_i), and displace particle within a small cube, say with an edge length of ε. Mathematically, this displacement is denoted as:

$$x_i \mapsto x_i + \varepsilon\delta_x, y_i \mapsto y_i + \varepsilon\delta_y, z_i \mapsto z_i + \varepsilon\delta_z$$

where $\delta_x, \delta_y, \delta_z \in [-1, 1]$. Denote the new configuration as $\mathbf{R}_2$. Calculate $|\Psi_T(\mathbf{R}_2)|^2$.
3. Calculate the ratio between the amplitude probabilities of the two configurations:

$$p := \frac{|\Psi_T(\mathbf{R}_2)|^2}{|\Psi_T(\mathbf{R}_1)|^2}$$

4. If $p \geq 1$, confirm moving the particle to the new position. If $p < 1$, move the particle to the new position with probability p. Mathematically, move the particle to the new position with probability

$$P = \min(1, p)$$

5. If the new configuration is accepted, repeat step 2 to 4 which generates another configuration and determine whether to accept or reject it based on the latest accepted configuration. If the new configuration is rejected, start from the latest accepted configuration, and repeat step 2 to 4.
6. This above process returns a list of M configurations $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_M$.

This algorithm developed by Metropolis et al. prevents one from needing to sample configurations that grows exponentially large as the dimension scales, and efficiently generates a set of configurations who satisfy the probability distribution from the trial wavefunction.
From the list of M configurations, it is then possible to calculate the local energy for each configuration following Eq. (8), and compute the variational approximation of the ground-state energy using Eq. (11).

4 Diffusion Monte Carlo

Similar to VMC, the Diffusion Monte Carlo (abbreviated as DMC) method provides an alternative way of approximating the ground-state properties of a many-body system.

There exists a key difference between VMC and DMC. VMC method selects a trial wavefunction with respect to certain parameters and updates these parameters from results of an iteration to optimize the wavefunction, which leads to the convergence to ground-state properties. DMC method, however, starts with a trial wavefunction, instead of optimizing this wavefunction through updating parameters, it samples certain "walkers", which will be explained in detail later, and allows them to diffuse over time with respect to an appropriate function, such that they converge to the ground state.

4.1 Mathematical Background

Additional to Definition 3.4, DMC is based on the solution to a "modified" version of the time-dependent Schrödinger's equation, where it is written in imaginary time [Tro11] [FMNR01] [KW08] [Aci97]:

Definition 4.1. *A time-dependent Schrödinger's equation written in imaginary time is*

$$-\frac{\partial}{\partial\tau}\Psi(\mathbf{R},\tau) = (\hat{H} - E_T)\Psi(\mathbf{R},\tau)$$

where $\tau = \frac{it}{\hbar}$*, and* E_T *is an energy offset.*

Additionally, in [Tro11], it is shown that as $\tau \to \infty$, the solution wavefunction to the above equation will have the ground state dominating other states. Furthermore, the introduction of the energy offset allows the wavefunction to be normalized [Tro11].
For the wavefunction to "diffuse" in time, Green's function is needed to describe the process, where [Tro11] [FMNR01] [KW08] [Aci97]:

Definition 4.2. *A Green's function that describes the diffusion process for the wavefunction described above,* $\Psi(\mathbf{R},\tau)$*, is given to be:*

$$G(\mathbf{R},\mathbf{R}',\delta\tau) = \left\langle \mathbf{R}\middle|\exp\Big(-\delta\tau(\hat{H} - E_T)\Big)\middle|\mathbf{R}'\right\rangle \tag{13}$$

where $\delta\tau$ *is the change in time.*

This permits us to write the general wavefunction to be the result of a diffusion process [Tro11] [FMNR01] [KW08] [Aci97]:

$$\Psi(\mathbf{R},\tau) = \int G(\mathbf{R},\mathbf{R}',\tau)\Psi(\mathbf{R}',0)\mathrm{d}\mathbf{R}'$$

where the Green's function acts like the probability density that the starting configuration $\mathbf{R}'$ evolves to $\mathbf{R}$ after time τ.
As $\delta\tau \to 0$, while assuming that $\hat{H} = \hat{T} + \hat{V}$, where $\hat{T}$ is the kinetic-energy operator and $\hat{V}$ is the potential-energy operator, by applying the Trotter-Suzuki formula and approximation for small $\delta\tau$, Eq. (13) can be rewritten as [Tro11] [FMNR01] [Aci97]:

$$G(\mathbf{R},\mathbf{R}',\delta\tau) = (2\pi\delta\tau)^{-\frac{3N}{2}}\exp\left(-\frac{(\mathbf{R}-\mathbf{R}')^2}{2\delta\tau}\right)\exp\left(-\delta\tau\frac{V(\mathbf{R})+V(\mathbf{R}')-2E_T}{2}\right) \tag{14}$$

where the rate term [KW08]

$$p := \exp\left(-\delta\tau\frac{V(\mathbf{R})+V(\mathbf{R}')-2E_T}{2}\right) \tag{15}$$

determines the number of "walkers" that survive to the next step [FMNR01]. For each original "walker", the number of descendants that "branches" out is:

$$n_d := \text{int}(p + \eta) \tag{16}$$

where int takes the integer part of the value, and η is a random number drawn uniformly at random on the interval $[0, 1]$ [Tro11] [FMNR01].
Additionally, the rest of Eq. (14) excluding Eq. (15) is a time propagator, G_d, which generates a random walk, and is a Gaussian distribution [Tro11][KW08].

4.2 Main Process

With the mathematical setup, it is possible to describe DMC in the following steps [KW08]:
Step 1: Sample the first generation of points $\mathbf{R}'$ according to a trial wavefunction $\Psi_T(\mathbf{R}')$.
Step 2: For each point in $\mathbf{R}'$, obtain a new set of points, which form a new configuration $\mathbf{R}$, by sampling according to the Gaussian distribution described by G_d, which gives us $\mathbf{R} - \mathbf{R}'$. This is where the idea of Monte Carlo is embedded.
Step 3: For each point in $\mathbf{R}$ and its corresponding point in $\mathbf{R}'$, generate n_d according to Eq. (16). If $n_d = 0$, remove the point from future random walks. If $n_d > 1$, replace the point with n_d clones.
Step 4: Adjust E_T accordingly to control the size of "walkers". After a sufficient enough time of repeating Step 2 and Step 3, for each point in the final configuration $\mathbf{R}_f$, find the ground-state energy by Eq. (11).

4.3 Importance Sampling

The process described above is generally extremely inefficient for a large number of particles due to possible complicated inter-particle interactions [Tro11], which can lead to Eq. (15) fluctuate wildly [FMNR01]. The technique of *importance sampling* is introduced to overcome these difficulties. Without extending into the mathematical details embedded in this technique, the key idea of the technique can be described as the following:

1. Design a *trial wavefunction*, $\Phi(\mathbf{R})$, that approximately describes the exact ground state Φ_0 [Tro11].
2. Introduce a new function $f(\mathbf{R}, \tau)$ such that [Tro11] [FMNR01] [Aci97]

$$f(\mathbf{R}, \tau) := \Phi(\mathbf{R})\Psi_T(\mathbf{R}, \tau) \tag{17}$$

where $f(\mathbf{R}, \tau)$ is interpreted as the probability density distribution of the population of "walkers" instead of $\Psi_T(\mathbf{R}, \tau)$.
3. Follow a similar diffusion process as described in the main process, except that, before deciding to "branch" or not, one need to introduce a *drift*, or *pseudo force*, $\mathbf{F}$, to the new configuration $\mathbf{R}$ relative to $\mathbf{R}'$, where [Tro11]

$$\mathbf{R} = \mathbf{R}' + \frac{\delta\tau}{2}\mathbf{F}(\mathbf{R}') \tag{18}$$

and

$$\mathbf{F} = \frac{2\nabla\Phi(\mathbf{R})}{\Phi(\mathbf{R})} \tag{19}$$

The introduction of the "drift" guides the walkers in regions with high probability, hence, reducing fluctuations in Eq. (15) and local energies [Tro11].
Another difficulty that DMC runs into is that it requires the wavefunction to be positive definite [FMNR01]. This means DMC can simulate Bosonic systems at zero temperature, but for fermionic systems like electronic systems that have positive and negative values due to antisymmetry, DMC runs into the *fermion sign problem*, which will be discussed in Section 6.1. Additionally, DMC also needs tuning if the system of interest consists of bosons in excited states [Tro11].

5 Path Integral Monte Carlo

Though the name may be deceptive, Path Integral Monte Carlo (abbreviated as PIMC) shares the same kernel as DMC and borrows the idea of a "path" from VMC. The key difference between PIMC and the other two QMC methods is that PIMC places its focus on reaching the thermodynamic equilibrium of a physical system, while VMC and DMC focuses on approximating the ground-state energies, with DMC having additional restrictions on the systems it can be used. PIMC explores the properties of physical systems with finite temperatures, which renders it to be more useful than DMC considering that DMC works best for zero-temperature Bosonic systems.

5.1 Mathematical Background

All QMC methods start by investigating which Hamiltonian is of interest, in most cases, similar to what was mentioned in DMC, a general Hamiltonian that describes a many-body system takes the form of:

$$\hat{H} = \hat{T} + \hat{V}$$

where $\hat{T}$ represents the kinetic term, and $\hat{V}$ represents the potential term, which may arise from different sources (one of such sources is the interaction between particles).
The key element that PIMC utilizes is the density matrix for the canonical ensemble which can be described as the following [Tro11] [Bar79] [KW08]:

Definition 5.1. *The thermal density matrix* $\exp\left(-\beta\hat{H}\right)$ *from which one can obtain thermal equilibrium takes the form where*

$$\rho(\mathbf{R}, \mathbf{R}', \beta) = \left\langle \mathbf{R} \middle| \exp\left(-\beta\hat{H}\right) \middle| \mathbf{R}' \right\rangle \tag{20}$$

denotes the matrix elements, and β *is the inverse of the product of the Boltzmann constant and the temperature:* $\beta = \frac{1}{k_B T}$

Definition 5.2. *The partition function, a function of the thermodynamic state variables, in concern is described to be [Bar79]*

$$Z = \frac{1}{N!} \int_{\mathbf{R}} \rho(\mathbf{R}, \mathbf{R}, \beta) \mathrm{d}\mathbf{R} \tag{21}$$

By leveraging Eq. (20) and the product property of the thermal density matrix, the following stands [Tro11] [Bar79] [KW08]

$$\rho(\mathbf{R}_1, \mathbf{R}_3, \beta_1 + \beta_2) = \int \rho(\mathbf{R}_1, \mathbf{R}_2, \beta_1)\rho(\mathbf{R}_2, \mathbf{R}_3, \beta_2)\mathrm{d}\mathbf{R}_2 \tag{22}$$

By taking M small time steps where $\tau = \frac{\beta}{M}$ as $M \to \infty$, it is possible to extend Eq. (22) that [Tro11] [Bar79]:

$$\rho(\mathbf{R}_1, \mathbf{R}_{M+1}, \beta) = \int \int \cdots \int \rho(\mathbf{R}_1, \mathbf{R}_2, \tau)\rho(\mathbf{R}_2, \mathbf{R}_3, \tau) \times \cdots \times \rho(\mathbf{R}_M, \mathbf{R}_{M+1}, \tau)\mathrm{d}\mathbf{R}_2\mathbf{R}_3 \cdots \mathrm{d}\mathbf{R}_M \tag{23}$$

When M is sufficiently large, or temperature T is sufficiently high, it follows that τ must be small, which allows replacing unknown $\rho(\mathbf{R}_j, \mathbf{R}_{j+1}, \tau)$ with known functions through approximation [Tro11]. One such approximation echoes Eq. (14) in DMC [Bar79],

$$\rho(\mathbf{R}, \mathbf{R}', \tau) \approx (2\pi\tau)^{-\frac{\nu N}{2}} \exp\left(-\frac{(\mathbf{R} - \mathbf{R}')^2}{2\tau} - \tau(\frac{V(\mathbf{R}) + V(\mathbf{R}')}{2})\right) \tag{24}$$

in atomic units, where ν is the dimensionality of space (can take $\nu = 3$ for 3D space), and

$$(\mathbf{R} - \mathbf{R}')^2 = \sum_i (\mathbf{r}_i - \mathbf{r}_i')^2$$

The Green's function in this case specifically refers to the part:

$$G(\mathbf{R}, \mathbf{R}', \tau) = \exp\left(-\frac{(\mathbf{R} - \mathbf{R}')^2}{2\tau} - \tau(\frac{V(\mathbf{R}) + V(\mathbf{R}')}{2})\right) \tag{25}$$

from Eq. (24).

5.2 Main Process

With the mathematical setup, the main process follows several steps as described below:
Step 1: Write the partition function in Definition 5.2 in terms of a path integral like Eq. (23).
Step 2: For each particle of all N particles, represent it in a "world line" such that $(\mathbf{r}_1^i, \mathbf{r}_2^i, \ldots, \mathbf{r}_M^i)$ describes the path of particle i over M steps. On aggregate, this should give us a path of configurations as well $(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_M)$.
Step 3: To obtain one path of configurations, each move is accepted or rejected based on the Metropolis algorithm similar to that of in Section 3.3 [Tro11], and the displacement of configurations is being sampled from a Gaussian distribution similar to that in Section 4.1, which, in this case, is according to Eq. (24).
Step 4: Obtain say n paths. From these n sample paths, it is possible to estimate a thermodynamic observable following a similar logic as estimating the ground-state energy:

$$\langle \mathcal{O} \rangle = \frac{1}{n} \sum_{i=1}^{n} \mathcal{O}(\{\mathbf{R}_i\})$$

where $\mathcal{O}(\{\mathbf{R}_i\})$ represents calculating the observable on the ith path sampled. The details of how to obtain the properties for a single path is not discussed here, as the mathematical details, once again, overshadow the key ideas of PIMC.
It is important, however, to highlight that PIMC suffers from the fermion sign problem as well, if not worse, compared to the DMC method. This is because DMC only evolves "walkers" (single points in a distribution), while PIMC evolves entire set of particles from one configuration to another.

6 Fermion Sign Problem

To generalize the fermion sign problem, it arises because of antisymmetric fermionic wavefunctions which produce weights of both signs in Monte Carlo [TW05]. This will lead to an exponential growth in statistical noise that prevents a Monte Carlo sampling from being possible and practical [TW05]. Computationally, the sign problem is NP-hard, which reflects a generic solution of the sign problem is non-existent unless NP = P [TW05].
Both DMC and PIMC suffer from this problem, and the following sections discuss briefly how each route overcomes this difficulty.

6.1 DMC - Fixed-node Approximation

The principle that DMC uses to "circumvent" the sign problem is to force the ground state of a fermionic system, described by a wavefunction Ψ_F, to have the same nodal structure as the trial wavefunction, Ψ_T [Tro11]. Specifically, given a trial wavefunction Ψ_T that describes a N-particle system, find a trial nodal surface to be a $(3N-1)$-dimensional surface on which the function Ψ_T evaluates to 0, and crossing such a surface would change the sign of Ψ_T [FMNR01]. For a DMC simulation to proceed, this nodal constraint forces the "walkers" not to cross the nodal surface described above [Tro11]. This further indicates that the accuracy would be limited, and the result from such a modification only provides an upper bound to the actual ground-state energy [Tro11].

6.2 PIMC - Restricted PIMC

The idea of "circumventing" the sign problem in PIMC is similar to that in DMC. The key difference between the two methods lies within from where do we restrict the nodes. In DMC, as described above, the nodal surface is found in a spatial situation; whereas in PIMC, the restricting condition can be described in terms of Eq. (22), where we need to enforce

$$\forall \tau \in [0, \beta], \rho(\mathbf{R}(\tau), \mathbf{R}(0), \tau) > 0 \tag{26}$$

This prevents any sampled path from ever entering the negative sign area, which leads to cancellation interference.

7 Discussion

To concisely summarize, this paper discussed three types of QMC methods: VMC, DMC, and PIMC. VMC and DMC generally share the same goal – approximating the ground-state energy of a many-body system, while PIMC offers a pathway to estimating thermodynamic properties in equilibrium. Both VMC and PIMC relies on the Metropolis algorithm to sample the next configuration, where PIMC's objective of using such an algorithm is to generate a configuration path. Both DMC and PIMC suffer from the fermion sign problem due to the nature of antisymmetry in fermionic systems. This issue can be "circumvented" but not solved classically, as it is NP-hard, through imposing certain restricting conditions.

Acknowledgments

Tobias would like to thank his past self for making the decision to go on an exchange program to ETH Zürich, at which he attended the course of Quantum Physics for Non-physicists taught by Prof. Philipp Kammerlander, and was able to have a foundation in quantum physics, allowing him to understand the various physical concepts present in this paper. Tobias would also like to express his gratefulness to Elisabetta Schneider, whom he met in the aforementioned quantum physics course, because she provided her lecture notes on Computational Quantum Physics for him, giving him a place to start researching.

Hannah would like to thank Tobias for taking the time to explain the background physics concepts, which greatly supported her in overcoming gaps in prior knowledge during the development of this paper.

Appendix

A Monte Carlo Method

A.1 Using Monte Carlo Method to Find π

The link to the Jupyter Notebook that demonstrates this process can be accessed here on GitHub.

B Variational Monte Carlo

B.1 Using VMC to Find the Ground-state Energy of the Two-electron System in a Fixed-Nucleus Helium Atom

The accompanying Jupyter Notebook (access here) contains a full implementation of a Variational Monte Carlo (VMC) simulation for the ground-state electronic energy of the helium atom in the Born–Oppenheimer approximation. We treat the nucleus as fixed at the origin and consider only the two interacting electrons. All calculations are performed in atomic units.

B.2 Variance-Based Optimization of the Jastrow Parameter

We optimize the Jastrow parameter β by minimizing the variance of the local energy. Since helium has only one variational parameter, a simple scan over a discrete grid,

$$\beta \in \{\beta_1, \beta_2, \ldots, \beta_M\}$$

For each β in this grid, the function `optimize_beta` runs a short VMC simulation, computes the corresponding variance of E_{loc}, and identifies the value of β that yields the smallest variance. This optimized parameter is then used in a longer VMC run to produce the final energy estimate.

References

[Aci97] Paulo H. Acioli. Review of quantum monte carlo methods and their applications. *Journal of Molecular Structure: THEOCHEM*, 394(2):75–85, 1997. doi:10.1016/S0166-1280(96)04821-X. [pp. 4, 8, 9]

[Bar79] J.A. Barker. A quantum-statistical monte carlo method; path integrals with boundary conditions. *The Journal of Chemical Physics*, 70(6):2914–2918, 1979. doi:10.1063/1.437829. [pp. 10, 11]

[FMNR01] W. M. C. Foulkes, L. Mitas, R.J. Needs, and G. Rajagopal. Quantum monte carlo simulations of solids. *Reviews of Modern Physics*, 73(1):33–83, 2001. doi:10.1103/RevModPhys.73.33. [pp. 3, 4, 5, 8, 9, 10, 12]

[KW08] Malvin H. Kalos and Paula A. Whitlock. *Monte Carlo Methods*. Wiley, 2008. [pp. 1, 2, 4, 8, 9, 10]

[MRR+53] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953. doi:10.1063/1.1699114. [p. 7]

[Tro11] Matthias Troyer. Computational quantum physics. Lecture notes, at ETH Zürich, 2011. URL: https://edu.itp.phys.ethz.ch/fs11/cqp/cqp.pdf. [pp. 8, 9, 10, 11, 12]

[TW05] Matthias Troyer and Uwe-Jens Wiese. Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations. *Physical Review Letters*, 94(17):170201, 2005. doi:10.1103/PhysRevLett.94.170201. [p. 12]