

Quantum Monte Carlo

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

Abstract

(To be completed)

1 Introduction

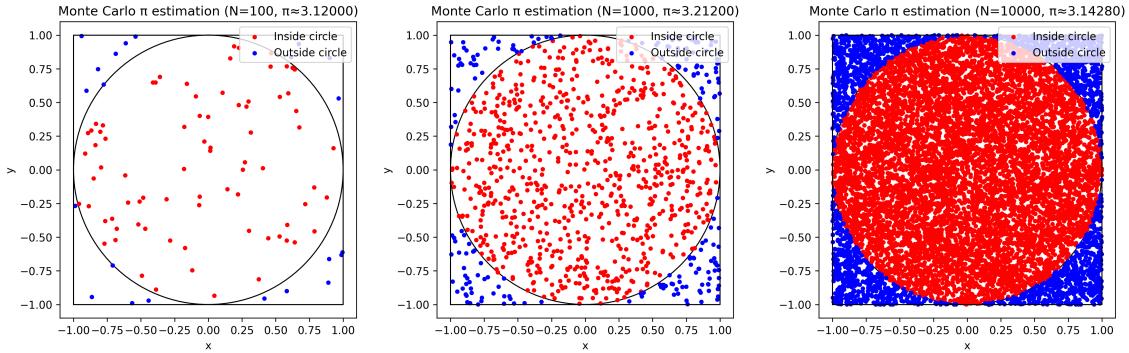
(To be completed)

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:



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

$$I = \int_0^1 \int_0^{\sqrt{1-x^2}} dy dx$$

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

In quantum many-body problems, investigating the ground-state energy of the many-body system is essential, as the system's ground state captures its equilibrium properties [FMNR01]. However, 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 $\mathbf{R} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$. A general quantum state $|\Psi\rangle$ in this system can be denoted as:

$$|\Psi\rangle = \int \Psi(\mathbf{R}) |\mathbf{R}\rangle d\mathbf{R} \quad (1)$$

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

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

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

$$\langle \mathcal{O} \rangle = \langle \Psi | \mathcal{O} | \Psi \rangle \quad (2)$$

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

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

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

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

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:

$$\hat{H} |\Psi\rangle = E |\Psi\rangle \quad (5)$$

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

$$\langle \hat{H} \rangle = E \quad (6)$$

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:

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

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

1. Ψ_T and $\nabla \Psi_T$ 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]:

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

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]:

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

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

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

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

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

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

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

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 (beyond exchange), while common DFT functionals struggle with strongly interacting systems, transition-metal atoms, van der Waals interactions, and dissociation regions. These limitations are further discussed in Ref. [FMNR01], Sec. II C. As a result, their predictive power is often limited.

Quantum Monte Carlo techniques successfully overcome many of these failures. Continuum QMC calculations can already achieve “chemical accuracy” ($\sim 1 \text{ kcal/mol} \approx 0.004 \text{ eV}$ 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:

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

where $R = (r_1, \dots, r_N)$ is the configuration of all electrons.

Slater determinant (antisymmetric part). The Slater determinant

$$\Phi(R) = \det [\phi_i(r_j)]$$

ensures antisymmetry under particle exchange, as required for fermions. Each orbital ϕ_i is a single-particle wavefunction describing where one electron “prefers” to be in space.

A high-quality Slater determinant should satisfy the following properties:

- **Correct symmetry.** The orbitals must transform appropriately under rotations, translations, reflections, and crystal periodicity (in solids).
- **Orthonormality.** Using orthonormal orbitals keeps the determinant numerically stable and produces well-behaved nodal surfaces.
- **Good nodal surfaces.** The nodal surface is the set of configurations where $\Psi_T(R) = 0$. Electrons can never be found at these nodes. HF orbitals often give qualitatively correct nodal surfaces, typically better than DFT–LDA orbitals for atoms and molecules. Since DMC’s fixed-node approximation depends entirely on the nodes of Ψ_T , poor nodal structure leads directly to higher errors.
- **Physically reasonable shape.** HF orbitals reflect chemical features such as shell structure, electron localization, hybridization, and bonding patterns.
- **Compactness.** HF already provides a single-determinant approximation. Without it, one would require very large configuration-interaction (CI) expansions, slowing down QMC convergence significantly.

These properties and the role of the Slater determinant as the antisymmetric component of the trial wavefunction are discussed in Ref. [FMNR01], Secs. II C and IV B.

Jastrow factor (correlation part). The Jastrow factor introduces explicit electron–electron correlations:

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

where $u(r_{ij})$ is a correlation function depending on the inter-electron distance.

This factor corrects several deficiencies of Hartree–Fock:

- **Cusp conditions.** The exact electronic wavefunction has cusps when two electrons approach each other. HF lacks these entirely.
- **Correlation holes.** The Jastrow factor accounts for electrons avoiding each other at short range.
- **Short- and long-range dynamical correlation.** This includes van der Waals interactions, which HF cannot describe.

Together, the Slater and Jastrow parts produce a wavefunction with both accurate symmetry structure (Slater) and realistic correlation behavior (Jastrow).

3.2.3 Choosing a Good Trial Wavefunction

Before optimizing the parameters within Ψ_T , one must first decide on a suitable functional form for the trial wavefunction itself. For electronic systems, the construction almost always begins with a set of single-particle orbitals obtained from a mean-field calculation such as Hartree–Fock (HF) or, in some cases, density functional theory (DFT). Although HF neglects dynamical electron correlation, it provides a high-quality antisymmetric foundation for VMC because:

- HF orbitals obey the correct spatial and spin symmetries of the problem,
- they are orthonormal and numerically stable,
- they produce nodal surfaces that are qualitatively correct for many atoms and molecules,
- and HF already captures the dominant exchange physics arising from antisymmetry.

As discussed in Sec. VII of [FMNR01], these properties make HF orbitals reliable building blocks for constructing the Slater determinant in VMC. In practice, one first performs an HF (or DFT–LDA) calculation for the system of interest and extracts the resulting orbitals $\{\phi_i(r)\}$. These orbitals form the Slater determinant $\Phi(R)$, which enforces the fermionic symmetry and defines the nodal surface that strongly influences the accuracy of both VMC and DMC.

However, the HF determinant lacks several key features of the exact ground state. In particular, it does not incorporate:

- the electron–electron cusp conditions,
- short-range correlation holes,
- long-range dynamical correlation,
- or van der Waals interactions.

To address these deficiencies, VMC introduces a Jastrow correlation factor $J(R; \alpha)$, whose functional form is chosen to satisfy the known cusp conditions and to capture both short- and long-range electron correlation. Typical Jastrow forms include one-body (electron–ion) and two-body (electron–electron) terms, with additional three-body terms used in high-accuracy calculations.

The overall process for selecting a good trial wavefunction therefore consists of:

1. performing an HF or DFT calculation to obtain physically meaningful orbitals,
2. constructing the Slater determinant using these orbitals,
3. choosing a Jastrow factor with the correct cusp behavior and flexible correlation terms,
4. and verifying that the resulting Slater–Jastrow form is compact, well-behaved, and computationally efficient to evaluate.

This Slater–Jastrow wavefunction serves as the starting point for subsequent parameter optimization, which refines the correlation terms while preserving the symmetry and nodal structure inherited from the HF determinant.

3.2.4 Optimization of Variational Parameters

The trial wavefunction contains adjustable parameters α , typically in the Jastrow factor and occasionally in the orbitals. These parameters are optimized by minimizing either

$$E_V(\alpha) \quad \text{or} \quad \text{Var}[E_{\text{loc}}],$$

where E_{loc} is the local energy.

Energy minimization. The variational principle ensures $E_V(\alpha) \geq E_0$, so minimizing E_V is a direct way to improve the wavefunction. However, Monte Carlo estimates of E_V can be noisy, making optimization challenging when Ψ_T is far from optimal.

Variance minimization. Since the local energy of the exact ground state is constant, variance minimization provides a stable and physically motivated strategy. It is widely used in practice, especially in the early stages of optimization.

TODO: METHOD OF OPTIMIZATION OF PARAMETERS

3.3 Sampling Method

4 Diffusion Monte Carlo

5 Path Integral Monte Carlo

Acknowledgments

If you have discussions with your classmates about your chosen topic, you can thank them here.

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](https://doi.org/10.1016/S0166-1280(96)04821-X). [pp. 3, 4]
- [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](https://doi.org/10.1103/RevModPhys.73.33). [pp. 2, 3, 4, 5, 6]
- [KW08] Malvin H. Kalos and Paula A. Whitlock. *Monte Carlo Methods*. Wiley, 2008. [pp. 1, 2, 3, 4]