Variational Monte Carlo Study of Hydrogen Atom and Hydrogen Molecule

Gavin Xie 02410400

Abstract—**A concise summary of aims, methods, results, and conclusions. Explain that the project investigates ground state wavefunctions of the harmonic oscillator, hydrogen atom, and hydrogen molecule using Variational Monte Carlo (VMC), along with optimisation and fitting to Morse potentials.**

I. INTRODUCTION

THIS project investigates the ground state energy for hydrogen molecules using Variational Monte Carlo (VMC). The project starts from choosing a suitable numerical differentiation method and sampling method. Those methods are then used to investigate 1D simple harmonic oscillator by calculating the local energy. A parametrised ansatz is then used to find the ground state wavefunction by minimising the Hamiltonian expectation with respect to the parameter θ.

Those methods are then generalized in to multi-dimension forms to determine the ground state energy of the Hydrogen atom, in which methods of differentiation, random sampling and minimization are verified by comparing the minimun local energy of the 1D simple harmonic system to the value provided from the note. The final stage is to determine the bond length and the dissociation energy of the Hydrogen molecules using the optimized ground state energy from different nuclie separations. A chi-square minimization method is used in the curve fitting for the Morse potential.

II. THEORY

A. Calculating the Expectation Value $\langle H \rangle$

The expectation value of the Hamiltonian

$$\langle H \rangle = \frac{\int d\mathbf{R}\, \Psi_T^*(\mathbf{R})\, \hat{H}\, \Psi_T(\mathbf{R})}{\int d\mathbf{R}\, |\Psi_T(\mathbf{R})|^2} \quad (1)$$

is investigated using Monte Carlo integration. This is reasonable choice because the integration is performed over a $3N$-dimensional configuration space where N denotes the number of electrons. Deterministic quadrature schemes such as Newton-Cotes rules suffer from the curse of dimensionality as the number of grid points required grows exponentially with dimension, introducing computationally infeasibility for quantum many-body systems. Furthermore, Newton–Cotes methods require a structured grid and smooth integrands, whereas the quantum-mechanical wavefunction may contain sharp features, cusps near nuclei, that would require extremely fine grids to resolve.

Monte Carlo integration offers several important advantages. First, with N_s sampling points, Monte Carlo integration converges as $\mathcal{O}(N_s^{-1/2})$, a rate that is independent of the dimensionality of the integral. This dimensional independence makes Monte Carlo methods especially effective for high-dimensional quantum systems. Moreover, sampling from the probability density

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

naturally concentrates computational effort in the physically relevant regions of configuration space where the wavefunction has significant weight. Further more, Monte Carlo methods do not require a uniform grid and therefore handle cusps and other non-smooth features of Ψ_T without difficulty, unlike grid-based quadrature schemes that would require prohibitively fine resolution.

For these reasons, Monte Carlo integration is the standard and practical method for evaluating expectation values in variational quantum calculations.

To make further progress, (1) can be rewritten as

$$\langle H \rangle = \int d\mathbf{R}\, \rho(\mathbf{R})\, \frac{\hat{H}\, \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} = \mathbb{E}_\rho[\, E_L(\mathbf{R})\,] \quad (3)$$

where the local energy is defined by

$$E_L(\mathbf{R}) = \frac{1}{\Psi_T(\mathbf{R})} \hat{H}\Psi_T(\mathbf{R}). \qquad (4)$$

Using N_s independent samples $\mathbf{R}^{(i)} \sim \rho(\mathbf{R})$, the Monte Carlo estimator of the expectation value becomes

$$\langle H \rangle \approx \frac{1}{N_s} \sum_{i=1}^{N_s} E_L(\mathbf{R}^{(i)}), \qquad (5)$$

which converges to the exact variational energy as $N_s \to \infty$ by the central limit theorem.

B. Local Energy

The local energy is defined as:

$$E_L(R) = \frac{1}{\psi(R)} H\psi(R).$$

For each system I explicitly compute or numerically approximate derivatives, justifying:

- step sizes used,
- order of finite-difference schemes (first, second, fourth order),
- numerical stability considerations.

III. 1D HARMONIC OSCILLATOR

A. The Dimensionless Schrödinger Equation

Derive the natural units and obtain the dimensionless Hamiltonian,

$$H = -\frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{2}x^2.$$

B. Analytical Eigenstates

Write the Hermite-polynomial eigenfunctions and eigenvalues:

$$\psi_n(x) = H_n(x)e^{-x^2/2}, \qquad E_n = n + \tfrac{1}{2}.$$

C. Finite Difference Method

Investigate the impact of step size and error order on the numerical error of the finite difference scheme to the second order derivative.

D. Metropolis sampling

In contrast to the transformation and rejection approaches, the Metropolis–Hastings algorithm constructs a Markov chain whose stationary distribution is the target density $\rho(x) = |\psi_n(x)|^2$. One begins from an initial point $x^{(0)}$ and generates proposals $x' = x^{(i)} + \eta$ with a symmetric proposal distribution $q(\eta)$, typically a Gaussian of adjustable width δ. Each proposal is accepted with probability

$$A(x \to x') = \min\left\{ 1, \frac{|\psi_n(x')|^2}{|\psi_n(x)|^2} \right\},$$

and otherwise the chain remains at $x^{(i)}$. After discarding an initial burn-in period, the successive positions $x^{(i)}$ are distributed according to $\rho(x)$ despite being correlated.

The local energy at each sampled point is evaluated using the same finite-difference approximation to the kinetic term,

$$E_L(x^{(i)}) \approx -\frac{1}{2}\frac{\psi_n(x^{(i)} + \Delta x) - 2\psi_n(x^{(i)}) + \psi_n(x^{(i)} - \Delta x)}{\Delta x^2} +$$

and the expectation value is estimated by a running average over the Markov chain. Since Metropolis samples are correlated, it is customary to monitor the autocorrelation time of E_L or to use blocking/binning methods to obtain a reliable estimate of the statistical uncertainty.

The key distinction is that transformation and rejection sampling produce independent samples. Transformation sampling is only available when the cumulative distribution function of $\rho(x)$ can be inverted analytically, which is rarely the case for excited states or more complicated systems. Rejection sampling requires a known envelope function that bounds $\rho(x)$ efficiently; for highly structured or nodal densities this envelope may be very inefficient, leading to a high rejection rate.

Metropolis sampling avoids both limitations because it needs only the ability to evaluate $|\psi_n(x)|^2$ up to a multiplicative constant. This makes it applicable to wavefunctions with nodes, multiple peaks, or high dimensionality. However, the resulting samples are correlated and require careful tuning of the proposal width δ to ensure both a reasonable acceptance rate and efficient exploration of the domain.

Overall, while transformation and rejection methods are preferred when they are analytically feasible

and efficient, the Metropolis algorithm offers a robust and widely applicable alternative for situations in which the structure of ψ_n makes direct sampling difficult or impractical.

E. Monte Carlo Verification

Steps required:
1) Generate samples $x^{(i)} \sim \rho(x) = |\psi_n(x)|^2$ using rejection sampling / transformation method.
2) Compute $E_L(x^{(i)})$ using a finite-difference second derivative.
3) Plot average energy vs. step size; verify convergence to E_n.

F. Error Analysis

Discuss statistical vs numerical error and justify the optimal step size used later in 3D problems.

IV. THE HYDROGEN ATOM

A. Hamiltonian and Ansatz

The dimensionless Hamiltonian is

$$H = -\frac{1}{2}\nabla^2 - \frac{1}{r}.$$

We use the trial wavefunction

$$\psi(r; \theta) = e^{-\theta r}.$$

Justification of ansatz:

It preserves spherical symmetry and resembles the exact ground state e^{-r}, allowing a single-parameter optimisation.

B. Sampling in 3D

Explain method used for sampling:
- Metropolis algorithm in Cartesian space,
- Why step-size tuning is required for 40–60% acceptance rate,
- Validation: radial histogram vs analytical form.

C. Local Energy in 3D

Compute the second derivative numerically using the 3D Laplacian. Discuss:
- interplay between step size and numerical noise;
- why small steps cause roundoff and large steps cause truncation error;
- tested different derivative orders (2nd and 4th).

D. Step Size of 3D Laplacian

For a function $f(\mathbf{r})$, the standard central finite-difference approximation of the Laplacian is

$$\nabla^2 f(\mathbf{r}) \approx \frac{f(x+h, y, z) + f(x-h, y, z) + f(x, y+h, z) +}{}$$

This formula contains two competing sources of numerical error:

(1) Truncation error for large h. The finite-difference expression comes from a Taylor expansion,

$$f(x \pm h) = f(x) \pm h f'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{6} f'''(x) + \cdots.$$

When forming the Laplacian stencil, all odd powers cancel and the first neglected term is $O(h^2)$. Thus the truncation error behaves as

$$\varepsilon_{\text{trunc}} \sim h^2.$$

A large stepsize therefore omits important curvature information and makes the approximation too coarse.

(2) Roundoff error for small h. The stencil involves subtracting nearly equal floating-point numbers, e.g.

$$f(x+h, y, z) - 2f(\mathbf{r}) + f(x-h, y, z),$$

which becomes extremely small as $h \to 0$. In finite precision arithmetic these subtractions lose significant digits (catastrophic cancellation), so the computed numerator has a relative roundoff error of size

$$\varepsilon_{\text{round}} \sim \frac{\epsilon_{\text{mach}}}{h^2},$$

where ϵ_{mach} is machine precision. Because we then divide by h^2, the roundoff error grows like $1/h^2$ and becomes dominant when h is too small.

Balance of errors. The total error is therefore

$$\varepsilon(h) \sim h^2 + \frac{\epsilon_{\text{mach}}}{h^2},$$

which has a minimum at an intermediate optimal stepsize. This explains why:
- *large h* gives large *truncation* error,
- *small h* gives large *roundoff* error from subtracting nearly equal numbers.

E. Gradient of Energy and Minimisation

Use estimator:

$$\partial_\theta \langle H \rangle \approx \frac{2}{N_s} \sum_i (E_L(R^{(i)}) - \langle H \rangle) \frac{\partial_\theta \psi(R^{(i)}; \theta)}{\psi(R^{(i)}; \theta)}.$$

Justify choice of optimiser:

- stochastic gradient descent (robust to MC noise),
- learning rate schedule,
- convergence criteria using energy uncertainty.

F. Results

- Optimised θ value.
- Computed ground-state energy (should approach -0.5).
- 2D histogram of (x, y) projection.
- Comparison to exact hydrogenic density.

V. THE HYDROGEN MOLECULE H_2

A. Hamiltonian

Write the full dimensionless Hamiltonian

$$H = -\frac{1}{2} \sum_{i=1}^{2} \left(\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right) - \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{1}{|\mathbf{r}_i - \mathbf{q}_j|} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} + \frac{1}{|\mathbf{q}_1 - \mathbf{q}_2|}.$$

B. Two-Electron Ansatz

$$\psi(r_1, r_2; \theta) = \left[e^{-\theta_1(|r_1 - q_1| + |r_2 - q_2|)} + e^{-\theta_1(|r_1 - q_2| + |r_2 - q_1|)} \right] e^{-\frac{r_{12}}{1 + \theta_3 r_{12}}}.$$

Justify each parameter:

- θ_1 controls electron–nucleus localisation,
- θ_2 controls electron–electron repulsion (Jastrow factor),
- θ_3 controls correlation length.

C. Sampling Strategy

- Metropolis sampling in 6D,
- Combined electron moves for efficiency,
- Burn-in stage justification,
- Step-size adjustment,
- Correlation analysis via autocorrelation time.

How those choices influence the error rate from MS sampling?

D. Energy Minimisation

Describe:

- gradient estimator,
- regularisation needed to maintain $\theta_1, \theta_3 > 0$,
- convergence diagnostics (energy trace, gradient norm, noise level).

Why gradient descent is no longer robust to MC noise?

Cannot converge even for more than 400 steps but only wandering around the minimun. Cannot achieve a small enough (only 1e-2 level) gradient or abs change in the norm of theta.

Use Quist-Newton and set the tolerance to grad tol 1e-3 (cannot set for even smaller for MC noise?) step size to 1e-4.

Also use mc noise from local energy to set for the converge condition. sigma delta (sq average of the std) comparing to the difference of the Energy. Meaning less if it not changing a lot.

How should I set for the tolerance? According to what information?

E. Projected Electron Density

Plot 2D histograms of (x_1, y_1) and (x_2, y_2) and justify bin size and sample count.

VI. FITTING TO THE MORSE POTENTIAL

Given energies $E(r)$ at various nuclear separations r, fit

$$V_{\text{Morse}}(r) = D(1 - e^{-a(r - r_0)})^2 - D + 2E_{\text{single}}.$$

A. Justification of Fit Procedure

- nonlinear least squares using `curve_fit`;
- weighting by uncertainty in E;
- justification for choice of r-range based on convergence.

B. Results

Report:

$$r_0 = \text{(value)} \pm \text{(std)},$$
$$D = \text{(value)} \pm \text{(std)},$$
$$a = \text{(value)} \pm \text{(std)}.$$

Compare to expected bond length ≈ 1.4 and dissociation energy ≈ 0.17.

VII. Discussion

- Physical interpretation of parameters,
- Comparison to known exact solutions,
- Sources of error: statistical, numerical (finite differences), optimisation noise,
- Robustness of results across sample sizes.

VIII. Conclusion

Summarise:

- successes (correct hydrogen and H_2 energies),
- challenges (correlated sampling, multi-parameter optimisation),
- future improvements (importance sampling, correlated sampling, automatic differentiation).

A. Subsection Heading Here

Subsection text here.

1) Subsubsection Heading Here: Subsubsection text here.

IX. Conclusion

The conclusion goes here.

Appendix A
Proof of Einstein's Famous Equation

The famous equation

$$E = mc^2$$

can be derived.

Appendix B

[Appendix two text goes here.]

Acknowledgment

[acknowledgments go here]