3 Introduction to Variational and Projector Monte Carlo

Cyrus J. Umrigar
Laboratory of Atomic and Solid State Physics
Physics Department, Cornell University
Ithaca, NY 14853, USA

Contents

E. Pavarini, E. Koch, and S. Zhang (eds.)
Many-Body Methods for Real Materials
Modeling and Simulation Vol. 9
Forschungszentrum Jülich, 2019, ISBN 978-3-95806-400-3
http://www.cond-mat.de/events/correl19

1 Introduction

Quantum Monte Carlo (QMC) methods are a broad and versatile class of stochastic algorithms for solving the many-body Schrödinger equation. They have been applied to fermions and bosons at zero and finite temperatures. Zero temperature calculations include both ground and excited states. Application areas range from quantum chemistry to lattice models of correlated materials to nuclear physics. Although the various methods differ greatly in their details, there is a small core of underlying ideas that are common to all the methods, which we will emphasize in this lecture. This lecture will be limited to zero temperature methods for fermionic systems. The emphasis will be on providing a unified description of both variational and projector Monte Carlo methods, both in discrete and in continuous space. The wide range of QMC applications will not be discussed. The finite-temperature path-integral Monte Carlo method has been reviewed in Ref. [1]. Details of zero-temperature methods, which we do not have time for here, can be found in the original papers, in books and review articles [2–8], and in chapters by Becca, Lüchow, Prokof'ev, Sandvik, and Zhang in this volume.

The many-body Schrödinger equation can be solved straightforwardly by expanding the wavefunction in a linear combination of determinants of single-particle orbitals, a method that is known as full configuration interaction. The limitation is that the number of states scales combinatorially in the number of orbitals $N_{\rm orb}$, and the number of electrons $N = N_\uparrow + N_\downarrow$, as ${}^{N_{\rm orb}}C_{N_\uparrow} \times {}^{N_{\rm orb}}C_{N_\downarrow}$, where $N_\uparrow$, $N_\downarrow$ are the number of up- and down-spin electrons respectively, so this brute-force method can be applied only to tiny systems. In contrast, some of the QMC methods scale as a low-order polynomial in N, provided that an approximate solution whose accuracy depends on the quality of a trial wavefunction is acceptable. Frequently high quality trial wavefunctions can be constructed, making QMC one of the few methods that provide accurate solutions for difficult problems.

2 QMC in a nutshell

I will distinguish between *QMC simulations* and *QMC calculations*, although almost all other researchers use these terms interchangeably. To my mind the stochasticity in QMC *simulations* mimics the stochasticity of the experimental system, e.g., the diffusion of neutrons in a nuclear reactor. On the other hand, QMC *calculations*, which is what we will be discussing here, introduce stochasticity into problems that are in fact deterministic. The solution of the many-body Schrödinger equation is perfectly deterministic (not to be confused with the probabilistic interpretation of the wavefunction amplitude) so it is in fact rather remarkable that introducing stochasticity makes the problem more tractable.

QMC methods are most useful when the dimension of the Hilbert space is very large and other many-body methods become impractical. They can be used both when the state of the system is described by discrete degrees of freedom, e.g., spin states or expansion coefficients of the wavefunction in a finite basis, as well as when the state is described by continuous degrees of freedom, e.g., the wavefunction amplitudes as a function of $3N-$dimensional electron coordi-

nates, in which case the Hilbert space is in fact infinite! The basic ideas of QMC are the same in discrete and continuous systems, so I will provide a unified treatment of both cases and will use discrete and continuous notation (e.g., sums and integrals) interchangeably.

2.1 Variational Monte Carlo

Many QMC methods rely on having good approximate wavefunctions to improve their accuracy and efficiency. So, we consider three wavefunctions: the exact wavefunction $|\Psi_0\rangle$ and two approximate wavefunctions which we call the trial wavefunction $|\Psi_T\rangle$ and the guiding wavefunction $|\Psi_G\rangle$: Their expansions in a (complete or incomplete) basis of $N_{\rm st}$ states are

$$\text{Exact} \qquad |\Psi_0\rangle \;=\; \sum_i^{N_{\rm st}} e_i\,|\phi_i\rangle, \quad \text{where} \quad e_i = \langle\phi_i|\Psi_0\rangle \tag{1}$$

$$\text{Trial} \qquad |\Psi_T\rangle \;=\; \sum_i^{N_{\rm st}} t_i\,|\phi_i\rangle, \quad \text{where} \quad t_i = \langle\phi_i|\Psi_T\rangle \tag{2}$$

$$\text{Guiding} \qquad |\Psi_G\rangle \;=\; \sum_i^{N_{\rm st}} g_i\,|\phi_i\rangle, \quad \text{where} \quad g_i = \langle\phi_i|\Psi_G\rangle \tag{3}$$

If the basis is incomplete then "exact" should be construed to mean "exact in that basis." Ψ_T and Ψ_G are frequently chosen to be the same function, but they serve two different purposes and at times there are good reasons for choosing them to be different, as will become apparent shortly. The basis state indices may be discrete (e.g., determinants of single-particle orbitals) or continuous (e.g., the $3N$ coordinates of the N electrons). In either case one can use $|\Psi_T\rangle$ to define the "local energy" of that state

$$E_{\rm L}(i) \;=\; \frac{\sum_j^{N_{\rm st}} H_{ij}t_j}{t_i}\,. \tag{4}$$

The variational energy of $|\Psi_T\rangle$ can be evaluated as follows:

$$\begin{aligned}
E_V \;&=\; \frac{\langle\Psi_{\rm T}|\hat{\boldsymbol{H}}|\Psi_{\rm T}\rangle}{\langle\Psi_{\rm T}|\Psi_{\rm T}\rangle} \;=\; \frac{\sum_{ij}^{N_{\rm st}}\langle\Psi_{\rm T}|\phi_i\rangle\;\langle\phi_i|\hat{\boldsymbol{H}}|\phi_j\rangle\;\langle\phi_j|\Psi_{\rm T}\rangle}{\sum_i^{N_{\rm st}}\langle\Psi_{\rm T}|\phi_k\rangle\;\langle\phi_k|\Psi_{\rm T}\rangle} \\
&=\; \frac{\sum_{ij}^{N_{\rm st}} t_i H_{ij} t_j}{\sum_k^{N_{\rm st}} t_k^2} \;=\; \sum_i^{N_{\rm st}} \frac{t_i^2}{\sum_k^{N_{\rm st}} t_k^2}\,\frac{\sum_j^{N_{\rm st}} H_{ij}t_j}{t_i} \;=\; \sum_i^{N_{\rm st}} \frac{t_i^2}{\sum_k^{N_{\rm st}} t_k^2}\,E_{\rm L}(i) \\
&\approx\; \frac{\left[\sum_i^{N_{\rm MC}} E_{\rm L}(i)\right]_{\Psi_{\rm T}^2}}{N_{\rm MC}} \;\to_{\Psi_G\neq\Psi_T}\; \frac{\left[\sum_i^{N_{\rm MC}}\left(\frac{t_i}{g_i}\right)^2 E_{\rm L}(i)\right]_{\Psi_{\rm G}^2}}{\left[\sum_k^{N_{\rm MC}}\left(\frac{t_k}{g_k}\right)^2\right]_{\Psi_{\rm G}^2}}
\end{aligned} \tag{5}$$

In the last line, we switch from the sum over all states to a sum over states sampled with probability $g_i^2/\sum_k^{N_{\rm st}} g_k^2$, first for $\Psi_{\rm G} = \Psi_{\rm T}$ (the usual case) and then for $\Psi_{\rm G} \neq \Psi_{\rm T}$. In the limit

that $\Psi_{\rm T} \rightarrow \Psi_0$, $E_{{\rm L},i} \rightarrow E_0$, the exact energy, independent of i. It is now apparent that the requirements on $|\Psi_T\rangle$ and $|\Psi_G\rangle$ are rather different. We wish to employ a $|\Psi_T\rangle$ which is not only a good approximation to $|\Psi_0\rangle$ (i.e., one that gives a low E_V and has small rms fluctuations in $E_{\rm L}$), but also one for which $E_{\rm L}$ can be evaluated efficiently. For a discrete basis, this means that $\hat{\boldsymbol{H}}$ should be very sparse in that basis, whereas for the continuous real space basis this requires that $\langle\phi_k|\Psi_{\rm T}\rangle$ and its Laplacian (needed for calculating the kinetic energy) can be evaluated efficiently. Instead the requirements on $|\Psi_G\rangle$ are that $\langle\phi_i|\Psi_{\rm G}\rangle$ can be evaluated efficiently and that it is nonzero whenever $\langle\phi_i|\Psi_0\rangle$ is nonzero, since otherwise the expectation values would be biased relative to the actual variational value.
Hence the minimal ingredients for an accurate and efficient VMC algorithm are:

1. A method (Metropolis-Hastings) for sampling $\langle\phi_k|\Psi_{\rm T}\rangle^2$.

2. A $|\Psi_T\rangle$ with variational parameters that is flexible enough to be a good approximation to $|\Psi_0\rangle$ for optimized parameters, and for which $\langle\phi_k|\Psi_{\rm T}\rangle$ and $E_{\rm L}$ can be evaluated efficiently.

3. Robust and efficient methods for optimizing the variational parameters.

These ingredients will be discussed in Secs. 3, 5, and 6 respectively.

2.2 Projector Monte Carlo

Projector Monte Carlo (PMC) methods evaluate the true energy E_0 (in the absence of a "sign problem") rather than the variational energy $E_{\rm V}$ using a "mixed estimator," $\langle\Psi_0|\hat{\boldsymbol{H}}|\Psi_{\rm T}\rangle/\langle\Psi_0|\Psi_{\rm T}\rangle$, for the energy. Following almost the same steps as in VMC

$$\begin{aligned}
E_0 &= \frac{\langle\Psi_0|\hat{\boldsymbol{H}}|\Psi_0\rangle}{\langle\Psi_0|\Psi_0\rangle} = \frac{\langle\Psi_0|\hat{\boldsymbol{H}}|\Psi_{\rm T}\rangle}{\langle\Psi_0|\Psi_{\rm T}\rangle} = \frac{\sum_{ij}^{N_{\rm st}} \langle\Psi_0|\phi_i\rangle\ \langle\phi_i|\hat{\boldsymbol{H}}|\phi_j\rangle\ \langle\phi_j|\Psi_{\rm T}\rangle}{\sum_k^{N_{\rm st}} \langle\Psi_0|\phi_k\rangle\ \langle\phi_k|\Psi_{\rm T}\rangle} \\
&= \frac{\sum_{ij}^{N_{\rm st}} e_i H_{ij} t_j}{\sum_k^{N_{\rm st}} e_k t_k} = \sum_i^{N_{\rm st}} \frac{e_i t_i}{\sum_k^{N_{\rm st}} e_k t_k} \frac{\sum_j^{N_{\rm st}} H_{ij} t_j}{t_i} = \sum_i^{N_{\rm st}} \frac{e_i t_i}{\sum_k^{N_{\rm st}} e_k t_k} E_{\rm L}(i) \\
&= \frac{\left[\sum_i^{N_{\rm MC}} E_{\rm L}(i)\right]_{\Psi_{\rm T}\Psi_0}}{N_{\rm MC}} \rightarrow_{\Psi_G \neq \Psi_T} \frac{\left[\sum_i^{N_{\rm MC}} \left(\frac{t_i}{g_i}\right) E_{\rm L}(i)\right]_{\Psi_{\rm G}\Psi_0}}{\left[\sum_k^{N_{\rm MC}} \left(\frac{t_k}{g_k}\right)\right]_{\Psi_{\rm G}\Psi_0}}
\end{aligned} \tag{6}$$

The only difference between Eqs. (6) and (5) is that the sampled distribution is $e_i g_i / \sum_k^{N_{\rm st}} e_k g_k$ rather than $g_i^2 / \sum_k^{N_{\rm st}} g_k^2$. At first sight this seems like an insurmountable obstacle since $|\Psi_{\rm G}\rangle$ is known, but $|\Psi_0\rangle$ is not. In Sec. 4 we describe methods for sampling the mixed distribution needed for PMC. The properties required of $|\Psi_{\rm T}\rangle$ and $|\Psi_{\rm G}\rangle$ are exactly the same as for VMC. Note that although an unbiased energy is obtained in the absence of a sign problem regardless of $|\Psi_{\rm T}\rangle$ and $|\Psi_{\rm G}\rangle$, the statistical error depends on $|\Psi_{\rm T}\rangle$ and $|\Psi_{\rm G}\rangle$. When the sign problem is present, approximations will be needed and their accuracy and efficiency depend on $|\Psi_{\rm T}\rangle$ and $|\Psi_{\rm G}\rangle$.

3 Variational Monte Carlo

3.1 Metropolis-Hastings method

As discussed in Sec. 2.1 we need to sample with probability $g_i^2/\sum_k^{N_{\rm st}} g_k^2$. To familiarize the reader with both discrete and continuous notation, we now assume we are working in continuous real space, in which case we need to sample from the probability density function $\rho(\mathbf{R}) = \psi_{\rm T}^2(\mathbf{R})/\int d\mathbf{R}\ \psi_{\rm T}^2(\mathbf{R})$, where we assume that the choice $|\Psi_{\rm G}\rangle = |\Psi_{\rm T}\rangle$ has been made since that is the usual practice. It is possible to sample a large number of probability density functions [9] directly (i.e., with no serial correlations) using the transformation or the rejection methods, but not such complicated probability density functions, so instead the Metropolis-Hastings method [10, 11] is used. It was originally developed to sample the thermal Boltzmann distribution, but is in fact a very powerful method for sampling any *known* discrete or continuous distribution. (In Sec. 4 we will see how to sample the unknown distribution $\psi_{\rm T}\psi_0/\int d\mathbf{R}\psi_{\rm T}\,\psi_0$!)

3.1.1 Markov chains

The Metropolis-Hastings method employs a *Markov chain*. A Markov chain is specified by two ingredients:

1) an initial state

2) a transition matrix $M(\mathbf{R}_{\rm f}|\mathbf{R}_{\rm i})$ (probability of transition from $\mathbf{R}_{\rm i} \to \mathbf{R}_{\rm f}$.) with the properties

$$M(\mathbf{R}_{\rm f}|\mathbf{R}_{\rm i}) \geq 0, \quad \text{and} \quad \int d\mathbf{R}_{\rm f}\, M(\mathbf{R}_{\rm f}|\mathbf{R}_{\rm i}) = 1. \qquad \text{(Column-stochastic matrix)} \tag{7}$$

The first property expresses the fact that probabilities must be non negative. The second property expresses the fact that a point at $\mathbf{R}_{\rm i}$ must go somewhere at the next step. The eigenvalues of a column-stochastic matrix are between 0 and 1, and there is at least one eigenvalue equal to 1 since the vector with all components equal to one is a left eigenvector with eigenvalue 1, and the left and right eigenvalues of any matrix are the same. If in addition M is a "primitive matrix," i.e., there exists an integer n for which all the elements of M^n are strictly positive, then there is a *unique* eigenvector with eigenvalue 1 and the Markov chain is said to be "ergodic."[1]
We wish to choose an M such that repeated application of M results in sampling $\rho(\mathbf{R})$, so we choose an M that satisfies

$$\int d\mathbf{R}_{\rm f}\, M(\mathbf{R}_{\rm i}|\mathbf{R}_{\rm f})\ \rho(\mathbf{R}_{\rm f}) = \rho(\mathbf{R}_{\rm i}) = \int d\mathbf{R}_{\rm f}\, M(\mathbf{R}_{\rm f}|\mathbf{R}_{\rm i})\ \rho(\mathbf{R}_{\rm i}) \quad \forall\ \mathbf{R}_{\rm i} \tag{8}$$

The first equality expresses a *stationarity condition*, namely that the net flux in and out $\mathbf{R}_{\rm i}$ is zero. Hence if one starts with the correct distribution $\rho(\mathbf{R})$, repeated application of M will continue to sample from $\rho(\mathbf{R})$. The second equality follows from the definition of a stochastic matrix. Eq. (8) shows that $\rho(\mathbf{R})$ is a right eigenvector with eigenvalue 1. Since all the other

[1] Here we ignore some subtleties that arise when the space is infinite.

eigenvalues are between 0 and 1, but not equal to 1, $\rho(\mathbf{R})$ is the dominant right eigenvector of M and repeated application of M results eventually in sampling $\rho(\mathbf{R})$.
In practice, the length of Monte Carlo (MC) run should be long enough that there is a significant probability of the system making several transitions between the neighborhoods of any pair of representative states that make a significant contribution to the average. This ensures that states are visited with the correct probability with only small statistical fluctuations. For example in a double-well system many transitions between the two wells should occur, but we can choose our Markov matrix to achieve this even if the barrier between wells is high.
A drawback of the Metropolis-Hastings method is that the sampled states are serially correlated. The rate at which the initial density evolves to the desired density ρ and the autocorrelation time of estimates of various observables is governed by the subdominant eigenvalues. In the ideal situation all the other eigenvalues are zero and every sample is independent.

Construction of M: Detailed balance condition We have as yet not provided a prescription to construct M, such that ρ is its stationary state. To do this we impose the *detailed balance* condition

$$M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ \rho(\mathbf{R}_\mathrm{i}) = M(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ \rho(\mathbf{R}_\mathrm{f}) \tag{9}$$

The detailed balance condition is more stringent than the stationarity condition. Instead of requiring that the net flux in and out of each state is zero, it requires that the net flux between *every pair* of states is zero. It is a sufficient, not a necessary condition, but it provides a practical way to construct M.
To go further, it is convenient employ a 2-step process. Moves from $\mathbf{R}_\mathrm{i}$ to a provisional point $\mathbf{R}'_\mathrm{f}$ are proposed with probability $T(\mathbf{R}'_\mathrm{f}|\mathbf{R}_\mathrm{i})$ and then accepted with probability $A(\mathbf{R}'_\mathrm{f}|\mathbf{R}_\mathrm{i})$. The corresponding Markov matrix is

$$M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) \;=\; \begin{cases} A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) & \text{if } \mathbf{R}_\mathrm{f} \neq \mathbf{R}_\mathrm{i} \\ 1 - \int d\mathbf{R}'_\mathrm{f}\ A(\mathbf{R}'_\mathrm{f}|\mathbf{R}_\mathrm{i})\ T(\mathbf{R}'_\mathrm{f}|\mathbf{R}_\mathrm{i}) & \text{if } \mathbf{R}_\mathrm{f} = \mathbf{R}_\mathrm{i} \end{cases} \tag{10}$$

$M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ and $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ are stochastic matrices, but $A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ is not. The detailed balance condition now becomes

$$\begin{aligned} A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ \rho(\mathbf{R}_\mathrm{i}) &= A(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ \rho(\mathbf{R}_\mathrm{f}) \\ \text{i.e.}\quad \frac{A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}{A(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})} &= \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}\,\frac{\rho(\mathbf{R}_\mathrm{f})}{\rho(\mathbf{R}_\mathrm{i})}\,. \end{aligned} \tag{11}$$

3.1.2 Choice of acceptance matrix

To satisfy Eq. (11) we can choose

$$A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) \;=\; F\left(\frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}\,\frac{\rho(\mathbf{R}_\mathrm{f})}{\rho(\mathbf{R}_\mathrm{i})}\right) \tag{12}$$

where F is any function for which $F(x)/F(1/x) = x$ and $0 \leq F(x) \leq 1$. Two possible choices are $F(x) = x/(1+x)$ and $F(x) = \min\{1, x\}$, corresponding to

$$A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ \rho(\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ \rho(\mathbf{R}_\mathrm{f}) + T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ \rho(\mathbf{R}_\mathrm{i})}, \tag{13}$$

$$\text{and} \quad A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \min\left\{1, \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}\,\frac{\rho(\mathbf{R}_\mathrm{f})}{\rho(\mathbf{R}_\mathrm{i})}\right\}. \tag{14}$$

The latter choice is the optimal choice since it maximizes the acceptance for given ρ and T, and is the choice made in the Metropolis et al. and the Hastings papers. Actually, Metropolis et al. assumed that $T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f}) = T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ in which case the factors of T drop out of Eq. (14), and Hastings made the generalization to $T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f}) \neq T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$, which enables a more efficient algorithm.

3.1.3 Choice of proposal matrix *T*

The optimal choice for the acceptance matrix $A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ is straightforward, Eq. (14), but there is considerable scope for using one's ingenuity to come up with good proposal matrices, $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$, that give small serial correlations between samples. As mentioned before, the ideal choice of $M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ has one eigenvalue equal to 1 and the rest zero. However, in practice the way to find efficient choices for $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ is not to think about eigenvalues, but instead to think about choosing a $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ that has large proposed moves and at the same time has high acceptance probabilities, i.e.,

$$\frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}\,\frac{\rho(\mathbf{R}_\mathrm{f})}{\rho(\mathbf{R}_\mathrm{i})} \approx 1. \tag{15}$$

There is a great deal of freedom in the choice of $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$, the only constraints being that it is a stochastic matrix leading to an ergodic Markov chain, and that it must be possible to efficiently sample $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ with a direct sampling method.

It may appear from Eq. (15) that our goal should be to make $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) \propto \rho(\mathbf{R}_\mathrm{f})$ since in that case the various factors cancel and the product becomes 1. This is in fact the case if it is possible to achieve that condition over all space, but it is not—if it were possible, we would not be using Metropolis-Hastings in the first place. So, we will discuss an alternative goal in a moment.

In order to prevent the acceptance from getting too small, it is common practice to restrict the moves to be in the neighborhood of $\mathbf{R}_\mathrm{i}$ by choosing $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ to be non-zero, or at least not negligible, only within a domain $D(\mathbf{R}_\mathrm{i})$ of volume $\Omega(\mathbf{R}_\mathrm{i})$ around $\mathbf{R}_\mathrm{i}$. For a given functional form of $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ the acceptance decreases as $\Omega(\mathbf{R}_\mathrm{i})$ increases, so, there exists an optimal $\Omega(\mathbf{R}_\mathrm{i})$ for which the system evolves the fastest.

To make further progress, we now make explicit that $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ is an ergodic matrix by writing

$$T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \frac{S(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}{\int d\mathbf{R}_\mathrm{f}\, S(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})} \approx \frac{S(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}{S(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{i})\Omega(\mathbf{R}_\mathrm{i})}\,, \tag{16}$$

where $S(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ is non-zero only in the domain $D(\mathbf{R}_\mathrm{i})$. Then,

$$\frac{A(\mathbf{R}_\mathrm{f},\mathbf{R}_\mathrm{i})}{A(\mathbf{R}_\mathrm{i},\mathbf{R}_\mathrm{f})} = \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}\frac{\rho(\mathbf{R}_\mathrm{f})}{\rho(\mathbf{R}_\mathrm{i})} \approx \frac{\Omega(\mathbf{R}_\mathrm{i})}{\Omega(\mathbf{R}_\mathrm{f})}\frac{S(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{i})}{S(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{f})}\frac{S(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})}{S(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}\frac{\rho(\mathbf{R}_\mathrm{f})}{\rho(\mathbf{R}_\mathrm{i})}. \quad (17)$$

Noting that for the present purpose S should be viewed as a function of the left index only, it is apparent that the choice

$$S(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) \propto \sqrt{\rho(\mathbf{R}_\mathrm{f})/\Omega(\mathbf{R}_\mathrm{f})} \quad \text{yields} \quad A(\mathbf{R}_\mathrm{f},\mathbf{R}_\mathrm{i})/A(\mathbf{R}_\mathrm{i},\mathbf{R}_\mathrm{f}) \approx 1. \quad (18)$$

To be more precise, if the log-derivatives of $S(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ equal those of $\sqrt{\rho(\mathbf{R}_\mathrm{f})/\Omega(\mathbf{R}_\mathrm{f})}$ at $\mathbf{R}_\mathrm{f} = \mathbf{R}_\mathrm{i}$, the average acceptance goes as $1 - \mathcal{O}(\Delta^3)$, where Δ is the linear dimension of $D(\mathbf{R}_\mathrm{i})$, provided that $D(\mathbf{R}_\mathrm{i})$ is inversion symmetric about $\mathbf{R}_\mathrm{i}$.
Another good choice for $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$, motivated by the diffusion Monte Carlo algorithm discussed in Sec. 4, is

$$T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \frac{1}{(2\pi\tau)^{3N/2}}\exp\left[-\frac{\left(\mathbf{R}_\mathrm{f}-\mathbf{R}_\mathrm{i}-\mathbf{V}(\mathbf{R}_\mathrm{i})\,\tau\right)^2}{2\tau}\right], \quad (19)$$

where $\mathbf{V}(\mathbf{R}_\mathrm{i}) = \nabla\Psi(\mathbf{R})/\Psi(\mathbf{R})|_{\mathbf{R}=\mathbf{R}_\mathrm{i}}$ is called the *drift velocity* of the wave function and τ is the time step which can be adjusted so as to minimize the autocorrelation time of the local energy. This is sampled by choosing

$$\mathbf{R}_\mathrm{f} = \mathbf{R}_\mathrm{i} + \mathbf{V}(\mathbf{R}_\mathrm{i})\tau + \boldsymbol{\eta}, \quad (20)$$

where $\boldsymbol{\eta}$ is a vector of $3N$ random numbers drawn from the Gaussian distribution with average 0 and standard deviation $\sqrt{\tau}$.
We demonstrate the increase in the acceptance probability that can be achieved by using Eq. (18) or Eq. (19) with a simple one-dimensional example. We wish to sample $\Psi(\mathbf{R})^2$. The simplest choice for $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ is a uniform distribution in $\Omega(\mathbf{R}_\mathrm{i})$ specified by $|\mathbf{R}_\mathrm{f} - \mathbf{R}_\mathrm{i}| < \Delta$ and zero outside

$$T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \begin{cases} \dfrac{1}{2\Delta} & \text{if } \mathbf{R}_\mathrm{f} \in \Omega(\mathbf{R}_\mathrm{i}), \\ 0 & \text{elsewhere.} \end{cases} \quad (21)$$

Instead, our recommended prescription from Eq. (18) is

$$T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \begin{cases} \dfrac{1}{2\Delta}\left(1 + \left.\dfrac{\nabla\Psi(\mathbf{R})}{\Psi(\mathbf{R})}\right|_{\mathbf{R}=\mathbf{R}_\mathrm{i}}(\mathbf{R}_\mathrm{f}-\mathbf{R}_\mathrm{i})\right) & \text{if } R_f \in \Omega(\mathbf{R}_\mathrm{i}), \\ 0 & \text{elsewhere} \end{cases} \quad (22)$$

and the prescription of Eq. (19) becomes

$$T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \frac{1}{(2\pi\tau)^{1/2}}\exp\left[-\frac{\left(\mathbf{R}_\mathrm{f}-\mathbf{R}_\mathrm{i}-V(\mathbf{R}_\mathrm{i})\,\tau\right)^2}{2\tau}\right]. \quad (23)$$

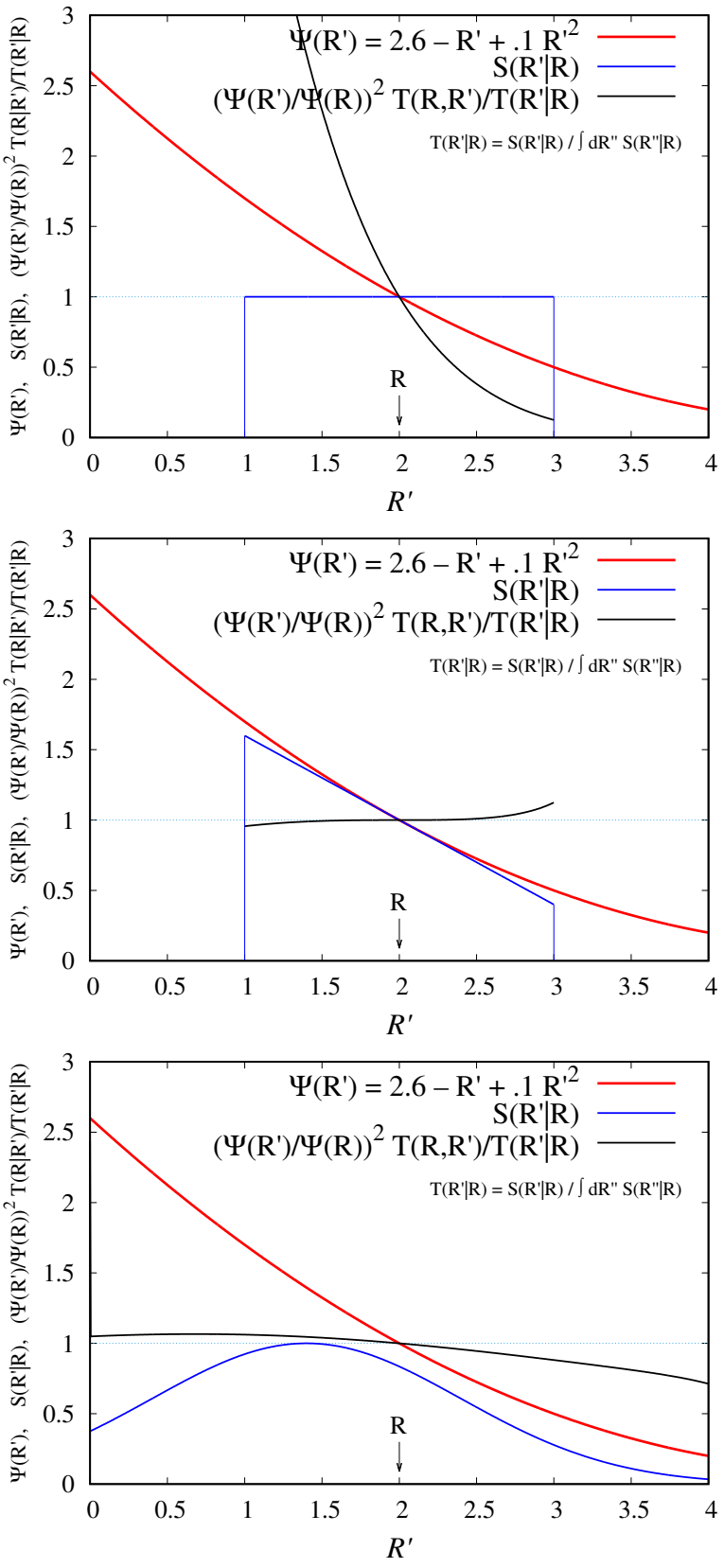

Fig. 1: Plots of $\Psi(R')$, $S(R'|R)$, $\left(\frac{\Psi(R')}{\Psi(R)}\right)^2 \frac{T(R,R')}{T(R',R)}$ for the three choices of $T(R',R)$ in Eqs. (21), (22), (23). The lower two plots have much larger average acceptances than the top.

In Fig. 1 we show plots of $\Psi(R')$, $S(R'|R)$, and $\left(\Psi(R')/\Psi(R)\right)^2 T(R,R')/T(R',R)$ for the three different choices of $T(R',R)$ in Eqs. (21), (22), and (23). In the top plot $\left(\frac{\Psi(R')}{\Psi(R)}\right)^2 \frac{T(R,R')}{T(R',R)}$ deviates linearly in $R'-R$ from 1 in the vicinity of $R'=R$ and deviates greatly from 1 over much of $\Omega(R)$, in the middle plot it deviates cubically in $R'-R$ from 1 in the vicinity of $R'=R$ and is close to 1 in all of $\Omega(R)$, and in the bottom plot it deviates linearly in $R'-R$ from 1 but stays close to 1 over a fairly wide range. Hence the choices of $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$ in Eq. (18) and (19) result in much smaller autocorrelation times than the simple choice of a symmetric $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$.

The analysis in Eqs. (16) to (19) and the examples in Eqs. (21) to (23) assume that $\rho(\mathbf{R})=\Psi(\mathbf{R})^2$ has continuous derivatives. In reality, $\Psi(\mathbf{R})^2$ has derivative discontinuities at electron-nucleus and electron-electron coincidence points. The former are much more important than the latter, since the wavefunction is large there and the velocity is directed towards the nucleus, so the electron tends to overshoot the nucleus. So, there is a high likelihood of having electrons near the nuclei and their acceptances are likely to be small. To tackle this problem, a highly efficient VMC algorithm with autocorrelation times of the local energy close to 1 is presented in Refs. [12, 13] wherein the electron moves are made in spherical coordinates centered on the nearest nucleus and the size of radial moves is proportional to the distance to the nearest nucleus. In addition, the size of the angular moves gets larger as one approaches a nucleus. This algorithm allows one to achieve, in many cases, an autocorrelation time of the local energy close to 1.

3.1.4 Moving the electrons all at once or one by one?

The accept-reject step can be performed after moving each electron, or after moving all the electrons. The former requires more computer time since the wavefunction and its gradient must be recalculated after each move. The increase in computer time is not a factor of N as one may naively expect, but more like a factor of 2. The reason is that it takes $\mathcal{O}(N^3)$ time to calculate a determinant from scratch, but only $\mathcal{O}(N^2)$ time to recalculate it, using the matrix determinant lemma and the Sherman-Morrison formula, after a single row or column has been changed. For systems with many electrons, moving the electrons one at a time leads to a more efficient algorithm because larger moves can be made for the same average acceptance, so the autocorrelation time of the local energy is smaller, more than compensating the increase of the calculation time per MC step.

4 Projector Monte Carlo

In Sec. 2 we wrote down the exact energy as a mixed estimator where the bra is the exact wavefunction and the ket the trial wavefunction, but we did not explain how one accesses the exact wavefunction, i.e., the ground state of the Hamiltonian. This is done using a projector. Projector Monte Carlo is a stochastic realization of the *power method* for finding the dominant eigenvector of a matrix. If one repeatedly multiplies an arbitrary eigenvector (not orthogonal to the dominant eigenvector) by the matrix, then one eventually gets the dominant eigenvector,

since at each step it gets multiplied by the largest in magnitude eigenvalue of the matrix. The power method is an example of an iterative eigensolver. Other iterative eigensolvers exist, such as the Lanczos method and the diagonally-preconditioned Davidson method, and they are much more efficient than the power method for deterministic calculations. So, if the dimension of the Hilbert space is not very large, say $< 10^{10}$ then one would just do a deterministic calculation using the Lanczos or Davidson methods. However, when the Hilbert space is so large (even infinite) that it is not possible to store even a few vectors of Hilbert space dimension, then PMC methods become the method of choice, since at any point in time they store only a sample of states.

The *projector* is any function of the Hamiltonian that maps the ground state eigenvalue of $\hat{\boldsymbol{H}}$ to 1, and the higher eigenvalues of $\hat{\boldsymbol{H}}$ to absolute values that are < 1 (preferably close to 0). We use the term "projector" somewhat loosely, since it is only repeated application of the projector that yields the desired state:

$$|\Psi_0\rangle \;=\; \lim_{n\to\infty} \hat{\boldsymbol{P}}^n(\tau)\;|\Psi_{\rm T}\rangle. \tag{24}$$

Possible choices for the projector are

$$\text{Exponential projector:} \qquad \hat{\boldsymbol{P}} \;=\; e^{\tau(E_T\hat{\mathbf{1}}-\hat{\boldsymbol{H}})} \tag{25}$$

$$\text{Linear projector:} \qquad \hat{\boldsymbol{P}} \;=\; \hat{\mathbf{1}}+\tau(E_T\hat{\mathbf{1}}-\hat{\boldsymbol{H}}) \qquad \big(\tau < 2/(E_{\rm max}-E_0)\big) \tag{26}$$

$$\text{Green function projector:} \qquad \hat{\boldsymbol{P}} \;=\; \frac{1}{\hat{\mathbf{1}}-\tau(E_T\hat{\mathbf{1}}-\hat{\boldsymbol{H}})} \tag{27}$$

where $E_{\rm T}$ is an estimate of the ground state energy.

4.1 Importance sampling

The projectors above enable us to sample $e_i = \langle\phi_i|\Psi_0\rangle$. However, according to Eq. (6) we want to sample from $g_i e_i = \langle\phi_i|\Psi_{\rm G}\rangle\,\langle\phi_i|\Psi_0\rangle$. Since

$$\sum_f P_{fi}e_i \;=\; e_f \tag{28}$$

the similarity transformed matrix with elements

$$\tilde{P}_{fi} = \frac{g_f P_{fi}}{g_i} \tag{29}$$

has an eigenstate with elements $g_i e_i$:

$$\sum_i \tilde{P}_{fi}(g_i e_i) \;=\; \sum_i \left(\frac{g_f P_{fi}}{g_i}\right)(g_i e_i) \;=\; g_f e_f. \tag{30}$$

$\hat{\tilde{\boldsymbol{P}}}$ is called the *importance sampled* projector and it samples $\langle\phi_i|\Psi_{\rm G}\rangle\,\langle\phi_i|\Psi_0\rangle$.

4.2 Branching random walks

Note that unlike the Markov matrix in Eq. (10) used in the Metropolis-Hastings method, the columns of the projector are not normalized to 1. We can write $\tilde{P}_{fi}$ as

$$\tilde{P}_{fi} \;=\; \frac{\tilde{P}_{fi}}{\sum_f \tilde{P}_{fi}} \sum_f \tilde{P}_{fi} \equiv T_{fi}\, W_i \tag{31}$$

where now T_{fi} is a Markov matrix and W_i are multiplicative weights. So, instead of the unweighted Monte Carlo walk that we had when doing VMC, we now have *walkers* that are specified by not just their position and but also by their weight. At each Monte Carlo step, the weight gets multiplied by W_i. If we have a single weighted walker, then a few generations of the walk will dominate the averages and the computational effort expended on the rest of the walk would be largely wasted. It is possible to have a population of walkers of fluctuating population size, with each walker having unit weight, but this leads to unnecessary birth/death events. So, it is best to have a population of walkers, such that all walkers within a generation have roughly the same weight, say within a factor of 2, and birth/death events when the weights go outside that range. Even so, the weights of different generations will vary a lot in a sufficiently long run. So, efficiency demands that we exercise *population control* to make the weights of each generation approximately the same. The population control error is proportional to the inverse of the target population size $N_{\rm walk}$. The error arises because of a negative correlation between the energy averaged over the generation and the weight of the generation. When the energy is low, the weight tends to be large and population control artificially reduces the weight and thereby creates a positive bias in the energy. Similarly, when the energy is high, the weight tends to be small and population control artificially increases the weight and this too creates a positive bias in the energy. Since the relative fluctuations in the energy and in the weight go as $1/\sqrt{N_{\rm walk}}$, the relative fluctuations in their covariance goes as $1/N_{\rm walk}$, resulting in a $\mathcal{O}(1/N_{\rm walk})$ population control bias. So, one way to reduce the population control error is to simply use a large population, and this is what most people do. If one wishes to be sure that the error is sufficiently small, plot the energy versus $1/N_{\rm walk}$ and take the limit $1/N_{\rm walk} \to 0$. But there exists a better way that allows us to estimate and remove most of the population control error within a single run [14, 15].

4.3 Taxonomy of PMC methods

The various PMC methods can be characterized by a) the form of the projector, and, b) the space in which the walk is performed, i.e., the single-particle basis and the quantization used. By second quantization we mean that walkers are characterized by only the occupancy of the single-particle states, whereas by first quantization we mean that walkers are characterized by the occupancy and by which electrons are on which states. So, walkers that are related by electron permutations in a first quantized space map onto the same walker in a second quantized space. Table 1 is a taxonomy of some PMC methods. For example, DMC uses the exponential projector (imaginary-time propagator) with a 1^{st} quantized continuous real space basis,

Method	Projector	SP Basis	Quantization
Diffusion Monte Carlo (DMC) [16, 17, 15]	$e^{\tau(E_T\hat{\mathbf{1}}-\hat{\boldsymbol{H}})}$	$\mathbf{r}$	1^{st}
GFMC [18–20]	$e^{\tau(E_T\hat{\mathbf{1}}-\hat{\boldsymbol{H}})}$ (samp. τ)	$\mathbf{r}$	1^{st}
Lattice-regularized DMC (Sorella, Casula)	$e^{\tau(E_T\hat{\mathbf{1}}-\hat{\boldsymbol{H}})}$ (samp. τ)	$\mathbf{r}_i$	1^{st}
FCIQMC [21–23]	$\hat{\mathbf{1}}+\tau(E_T\hat{\mathbf{1}}-\hat{\boldsymbol{H}})$	ϕ_i^{orthog}	2^{nd}
phaseless AFQMC [24, 5]	$e^{\tau(E_T\hat{\mathbf{1}}-\hat{\boldsymbol{H}})}$	$\phi_i^{\mathrm{nonorthog}}$	2^{nd}

Table 1: *Taxonomy of PMC methods. The methods are characterized by the projector, and the space in which the walk is performed, i.e., the single-particle basis and the quantization.*

AFQMC also uses the exponential projector but with a 2^{nd} quantized orbital basis, and FCIQMC uses a linear projector with a 2^{nd} quantized orbital basis. In AFQMC the orbitals are nonorthogonal and evolve continuously during the MC run, whereas in FCIQMC they are orthogonal and are fixed during the entire run. The linear projector has the advantage that if the Hamiltonian is known exactly in the chosen basis, so also is the projector. However, it can be used only when the spectrum of the Hamiltonian is bounded.

4.4 Diffusion Monte Carlo

We now discuss in some detail just one of the various PMC methods, the PMC method in real-space using first-quantized walkers. This method is more commonly known as diffusion Monte Carlo (DMC), and the projector is often referred to as the Green function since it is the Green function of the Fokker-Planck equation in the short-time approximation. We limit the following discussion to $\Psi_{\mathrm{G}}(\mathbf{R}) = \Psi_{\mathrm{T}}(\mathbf{R})$.
We now derive an approximate expression for

$$\left\langle \Psi_{\mathrm{T}}|\mathbf{R}'\right\rangle \langle \mathbf{R}'|e^{\tau(E_T\hat{\mathbf{1}}-\hat{\boldsymbol{H}})}|\mathbf{R}\rangle \frac{1}{\langle \mathbf{R}|\Psi_{\mathrm{T}}\rangle} \;\equiv\; \Psi_{\mathrm{T}}(\mathbf{R}')G(\mathbf{R}',\mathbf{R},\tau)\frac{1}{\Psi_{\mathrm{T}}(\mathbf{R})} \;\equiv\; \tilde{G}(\mathbf{R}',\mathbf{R},\tau).$$

We multiply the imaginary-time the Schrödinger equation

$$-\frac{1}{2}\nabla^2\Psi(\mathbf{R},t) \;+\; \big(\mathcal{V}(\mathbf{R})-E_{\mathrm{T}}\big)\Psi(\mathbf{R},t) \;=\; -\frac{\partial\Psi(\mathbf{R},t)}{\partial t} \tag{32}$$

by $\Psi_{\mathrm{T}}(\mathbf{R})$ and rearrange terms to obtain

$$-\frac{\nabla^2}{2}(\Psi\Psi_{\mathrm{T}}) \;+\; \nabla\cdot\left(\frac{\nabla\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}\Psi\Psi_{\mathrm{T}}\right) \;+\; \left(\underbrace{\frac{-\nabla^2\Psi_{\mathrm{T}}}{2\Psi_{\mathrm{T}}}+\mathcal{V}}_{E_{\mathrm{L}}(\mathbf{R})}-E_{\mathrm{T}}\right)(\Psi\Psi_{\mathrm{T}}) \;=\; -\frac{\partial(\Psi\Psi_{\mathrm{T}})}{\partial t} \tag{33}$$

defining $f(\mathbf{R},t)=\Psi(\mathbf{R},t)\Psi_{\mathrm{T}}(\mathbf{R})$, this is

$$\underbrace{-\frac{1}{2}\nabla^2 f}_{\text{diffusion}} \;+\; \underbrace{\nabla\cdot\left(\frac{\nabla\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}f\right)}_{\text{drift}} \;+\; \underbrace{\big(E_{\mathrm{L}}(\mathbf{R})-E_{\mathrm{T}}\big)f}_{\text{growth/decay}} \;=\; -\frac{\partial f}{\partial t} \tag{34}$$

Since we know the solution for each individual term on the LHS, an approximation Green function is

$$\tilde{G}(\mathbf{R}',\mathbf{R},\tau) = \frac{1}{(2\pi\tau)^{3N/2}}\, e^{-(\mathbf{R}'-\mathbf{R}-\mathbf{V}(\mathbf{R})\tau)^2/(2\tau)+\left(E_{\rm T}-\left(E_{\rm L}(\mathbf{R}')+E_{\rm L}(\mathbf{R})\right)/2\right)\tau} + \mathcal{O}(\tau^2). \qquad (35)$$

This is the same as the proposal matrix in Eq. (20) but there is now an additional growth/decay factor. So, the walker moves exactly as before, but the weight gets multiplied by this factor. Simply by turning on this multiplicative factor, one can switch between VMC and DMC.
This projector gives rather poor results—it has a very large time-step error and typically gives charge distributions for finite systems that are too diffuse. We describe next two improvements: a) take into account the singularities of the projector, and b) introduce an accept-reject step to ensure that the exact distribution is sampled in the limit that the $\Psi_{\rm T}(\mathbf{R})$ is exact.

4.4.1 Singularities of $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$

Following Ref. [15], Table 2 shows the singularities of $\tilde{G}(\mathbf{R}',\mathbf{R},\tau)$ at the nodes of $\Psi_{\rm T}(\mathbf{R})$ and the particle coincidences. $\tilde{G}(\mathbf{R}',\mathbf{R},\tau)$ is accurate if $\mathbf{V}(\mathbf{R})$ is nearly constant over the time step, τ, and if $E_{\rm L}$ changes nearly linearly between $\mathbf{R}$ and $\mathbf{R}'$. Both assumptions fail dramatically at these singularities. For exact $\Psi_{\rm T}$ the local energy $E_{\rm L}(\mathbf{R})$ is constant, but for the approximate $\Psi_{\rm T}$ used in practice, it diverges to $\pm\infty$ as the inverse of the distance to the nearest node. On the other hand the divergence of $E_{\rm L}(\mathbf{R})$ at particle coincidences is easily removed simply by imposing the electron-nucleus and electron-electron cusps on $\Psi_{\rm T}$. The velocity $\mathbf{V}(\mathbf{R})$ diverges at the nodes and has a discontinuity at particle coincidences even for exact $\Psi_{\rm T}$, so this cannot be taken care of by improving $\Psi_{\rm T}$. Instead, we must improve upon the implicit assumption in $\tilde{G}(\mathbf{R}',\mathbf{R},\tau)$ that $\mathbf{V}(\mathbf{R})$ is constant during the time step τ. When an electron is near a node, the velocity diverges as the inverse distance to the node and is directed away from the node, so the large velocity does not persist over the entire time step τ. Taking this into account and integrating over the time step we find that the average velocity over the time-step τ is:

$$\bar{\mathbf{V}} \;=\; \frac{-1+\sqrt{1+2V^2\,\tau}}{V^2\,\tau}\mathbf{V} \;\rightarrow\; \begin{cases} \mathbf{V} & \text{if } V^2\,\tau \ll 1 \\ \sqrt{2/\tau}\;\hat{\mathbf{V}} & \text{if } V^2\,\tau \gg 1 \end{cases} \qquad (36)$$

Similar improvements to the average velocity in the vicinity of electron-nucleus coincidences, and the average of $E_{\rm L}$ can also be made.

Table 2: *Singularities of* $\tilde{G}(\mathbf{R}',\mathbf{R},\tau)$.

Region	Local energy $E_{\rm L}$	Velocity $\mathbf{V}$
nodes	$E_{\rm L} \sim \pm 1/R_\perp$ for $\Psi_{\rm T}$ $E_{\rm L} = E_0$ for Ψ_0	$V \sim 1/R_\perp$ for both $\Psi_{\rm T}$ and Ψ_0
e-n and e-e coincidences	$E_{\rm L} \sim 1/x$ if cusps not imposed $E_{\rm L}$ finite if cusps are imposed $E_{\rm L} = E_0$ for Ψ_0	V has a discontinuity for both $\Psi_{\rm T}$ and Ψ_0

4.4.2 Imposing detailed balance in DMC

Since $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$ has a time-step error, it fails to sample the correct distribution even in the limit that $\Psi_{\rm T}(\mathbf{R})$ is exact. So, for large values of τ, the version of DMC presented so far can actually be less accurate than VMC. Following Refs. [17, 15] this is easily remedied as follows: If we omit the third term on the left-hand side of Eq. (34) then it can be verified that $f(\mathbf{R}) = \Psi_{\rm T}(\mathbf{R})^2$ is the solution since

$$-\frac{1}{2}\nabla^2\psi_{\rm T}^2(\mathbf{R}) \; + \; \nabla\cdot\left(\frac{\nabla\psi_{\rm T}}{\psi_{\rm T}}\psi_{\rm T}^2(\mathbf{R})\right) \; = \; 0. \tag{37}$$

However, we can sample $\Psi_{\rm T}(\mathbf{R})^2$ using Metropolis-Hastings. So, we can view the drift-diffusion part of $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$ as being the proposal matrix $T(\mathbf{R}', \mathbf{R})$ and introduce an accept-reject step after the drift and diffusion steps and before the reweighting step. With this simple modification, DMC is guaranteed to give an energy lower than the VMC energy, and in the limit of $\tau \to 0$ an energy that is variational, i.e., higher than the true energy. Finally, we account for the fact that the walker moves less far, since some of the moves are rejected, by using an effective time step $\tau_{\rm eff}$ for reweighting

$$\tau_{\rm eff} = \tau\,\frac{R^2_{\rm accep}}{R^2_{\rm prop}}\,, \tag{38}$$

where $R^2_{\rm prop}$ is the sum of the squares of all the proposed one-electron move distances and $R^2_{\rm accep}$ is the same sum, but including only the accepted moves.

4.5 Sign problem

PMC suffer from a sign problem except in a few special situations, e.g., 1-dimensional problems in real space. In all PMC methods, the sign problem occurs because an undesired state grows relative to the state of interest when the system is evolved by repeated stochastic applications of the projector. This results in a computational cost that grows exponentially with system size N. A reasonable definition [25] is that there is no sign problem if the computer time required to compute the value of an observable for an N-particle system with specified error, ϵ, scales as $T \propto N^{\delta}\epsilon^{-2}$, where δ is a small power (say, $\delta \leq 4$). (It is assumed that N is increased in some approximately homogeneous way, e.g., adding more atoms of the same species.) The details of how the sign problem manifests itself is different in the various PMC methods, and we discuss two examples.

4.5.1 Sign problem in DMC

In the absence of importance sampling (which has the important side effect of imposing the fixed-node constraint), the DMC projector of Eq. (35) becomes

$$G(\mathbf{R}', \mathbf{R}, \tau) \equiv \langle \mathbf{R}'|\hat{\boldsymbol{P}}(\tau)|\mathbf{R}\rangle \approx \frac{e^{-(\mathbf{R}'-\mathbf{R})^2/(2\tau)+\left(E_T-\left(\mathcal{V}(\mathbf{R}')+\mathcal{V}(\mathbf{R})\right)/2\right)\tau}}{(2\pi\tau)^{3N/2}}. \tag{39}$$

It is nonnegative everywhere, so there is no sign problem if one were interested in the dominant state of this projector. However, the dominant state of this projector is the Bosonic ground state whereas the state of interest is the Fermionic ground state. If one started with a positive distribution and a negative distribution such that their sum is purely Fermionic as illustrated in Fig. 2, and applied the projector deterministically, both the positive and the negative distributions would tend to the Bosonic ground state, but, their sum would yield the Fermionic ground state, though with an amplitude that gets exponentially small relative to the amplitude of the individual components with increasing MC time. However, the projection is done stochastically and the probability of positive and negative walkers landing on the same state at the same MC time step and cancelling is very small if the portion of the state space that contributes significantly to the expectation values is finite and large, and it is zero if the state space is continuous (unless the dynamics of the walkers is modified to force opposite-sign walkers to land on the same spot). Hence it is not possible to sum the positive and negative contributions to extract the Fermionic ground state. This is demonstrated in Fig. 2. Furthermore, the problem cannot be solved by using an extremely large population of walkers. This enhances the probability of cancellations, but, because of fluctuations, eventually only positive or only negative walkers will survive and so the Fermionic state will completely disappear.

Fixed-node approximation: The importance-sampled Green function of Eq. (35) is not just a similarity transform, as in Eq. (29), of the Green function of Eq. (39). A *fixed-node* constraint has sneaked in as well. The velocity in Eq. (35) diverges at nodes and is directed away from them, so the number of node crossings per unit time goes to zero in the $\tau \rightarrow 0$ limit.[2] So, the solution obtained is the solution to the Schrödinger equation with the boundary condition that it goes to zero at the nodes of $\Psi_{\rm T}(\mathbf{R})$. Since we are now adding in an artificial constraint, the resulting energy has a positive *fixed-node error*, which disappears in the limit that the nodal surface of $\Psi_{\rm T}(\mathbf{R})$ is exact. The fixed-node approximation not only enables the calculation of the Fermionic ground state by eliminating the non-Fermionic states, it also enables the calculation of Fermionic excited states by preventing a collapse to the Fermionic ground state. For the excited states one loses the upper-bound property, but nevertheless the method has been used to calculate challenging excited states accurately [26–28].

4.5.2 Sign problem in FCIQMC

It may appear from the above discussion that the sign problem can be solved by performing the MC walk in a 2^{nd}-quantized, i.e., antisymmetrized, basis. Each 2^{nd}-quantized basis state consists of all the permutations of the corresponding 1^{st}-quantized basis states. Then there are no Bosonic states or states of any symmetry other than Fermionic, so there is no possibility of getting noise from non-Fermionic states. Of course, it is well known that this does not solve the sign problem. The problem is that different paths leading from a state to another can contribute with opposite sign. If the opposite sign contributions occur at the same MC step, then the

[2]The number of node crossings per MC step goes as $\tau^{3/2}$, so the number of crossings per unit time goes to zero as $\sqrt{\tau}$.

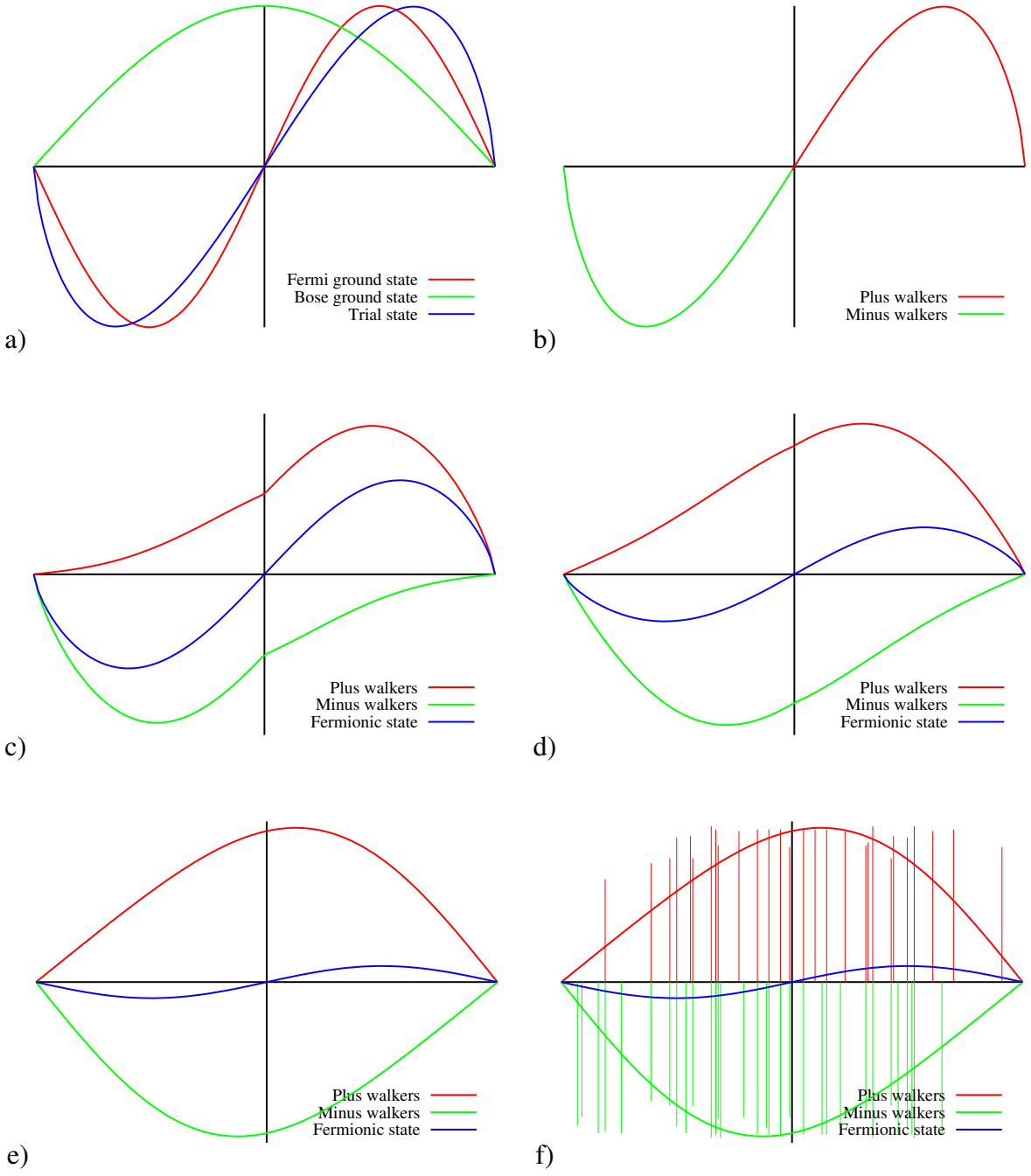

Fig. 2: *Demonstration of the sign problem in DMC. a) The green curve schematically depicts the Bosonic ground state, the red curve the Fermionic ground state, and the blue curve an approximate Fermionic wavefunction. b) The starting positive distribution is shown in red and the starting negative distribution in green. Their sum is purely Fermionic. c-f) The red and the green curves show the evolution of the positive and negative distributions. Their sum, the blue curve, converges to the Fermionic ground state. f) For a finite population, the walkers are depicted as delta functions and in a continuous space they never meet and cancel (unless they are forced to in some way). Consequently there is an exponentially vanishing "signal to noise" ratio.*

contributions cancel and yield a net contribution of smaller absolute magnitude, just as they would in a deterministic calculation. The problem occurs when opposite sign contributions occur at different MC steps. Further, since Ψ and $-\Psi$ are equally good, they are each sampled with equal probability in the course of a long MC run.

In a few special situations the sign problem is absent. The necessary and sufficient condition for there to be no sign problem is that all columns (or equivalently rows) of the projector have the same sign structure aside from an overall sign. Equivalently, there is no sign problem if it is possible to find a set of sign changes of the basis functions such that all elements of the projector are nonnegative. For example, the projector with the following sign structure

$$\begin{bmatrix} + & - & + & + \\ - & + & - & - \\ + & - & + & + \\ + & - & + & + \end{bmatrix} \tag{40}$$

does not have a sign problem, since changing the sign of the 2^{nd} basis state makes all the elements nonnegative. Note that it is not necessary to actually make these sign changes—the projectors before and after the sign changes are equally good.

Although a 2^{nd}-quantized basis does not solve the sign problem, it is advantageous to use a 2^{nd}-quantized basis when designing an exact Fermionic algorithm. First, an antisymmetrized basis is a factor of $N!$ smaller, and so the probability of opposite sign walkers meeting and cancelling is greater. Second, since each 2^{nd}-quantized basis state consists of a linear combination of 1^{st}-quantized basis states, 2^{nd}-quantized states that are connected by the projector may have multiple connections coming from several of the constituent 1^{st}-quantized states. Hence there is the possibility of internal cancellations in the 2^{nd}-quantized basis, which reduces the severity of the sign problem [29]. Third, since Bosonic and other symmetry states are eliminated it is clear that one can achieve a stable signal to noise for any large but finite basis by making the walker population very large. The limit of an infinite walker population is equivalent to doing a deterministic projection, which of course does not have a sign problem.

The FCIQMC method [21] does just this. It uses efficient algorithms for dealing with a large number of walkers and obtains essentially exact energies for very small molecules in small basis sets, using a large but manageable number of walkers. The MC walk, is done in an orbital occupation number (or equivalently determinantal) basis. For somewhat larger systems it is necessary to employ the *initiator approximation* [22] which greatly reduces the walker population needed to achieve a stable signal to noise ratio. Only states that have more walkers on them than some threshold value can spawn walkers on states that are not already occupied. The associated initiator error disappears of course in the limit of an infinite population and in practice it is possible to get accurate energies for interesting systems. However, the initiator error can be of either sign and it can be nonmonotonic, so extrapolation to the infinite population limit can be tricky. A large gain in efficiency can be gained by doing the projection on the most important states deterministically [23] since the sign-problem is present only for stochastic projection.

5 Form of trial wavefunctions

A major advantage of QMC methods is that since the integrals are done using Monte Carlo, one has a great deal of freedom in the choice of the form of the trial wavefunction $\Psi_{\rm T}(\mathbf{R})$, allowing for compact yet accurate choices. As mentioned in Sec. 2.1, both the accuracy and the efficiency of both VMC and PMC (when there is a sign problem) depend on $\Psi_{\rm T}(\mathbf{R})$. The only constraint on the choice of $\Psi_{\rm T}(\mathbf{R})$ is that it should be possible to evaluate it and its local energy quickly (in low-order polynomial in N time).

5.1 Slater-Jastrow wavefunctions

The most commonly used choice for electronic systems is the so-called Jastrow-Slater form, that is, the product of a linear combination of determinants D_n of single-particle orbitals and a Jastrow correlation factor

$$\Psi_{\rm T} = \left(\sum_n d_n \, {\rm D}_n^{\uparrow} \, {\rm D}_n^{\downarrow} \right) \times \mathcal{J}. \tag{41}$$

Note that in order to save computation time, we have replaced each determinant by a product of an up-spin and a down-spin determinant, which is not fully antisymmetric. This is legitimate because the expectation value is unchanged upon full antisymmetrization for any operator that does not have explicit spin dependence. The single-particle orbitals are usually expanded in spherical harmonics times Slater functions (monomial times an exponential in radial distance) for all-electron calculations in order to be able to impose the electron-nucleus cusps, and in spherical harmonics times gaussians or Gauss-Slater functions [30] for pseudopotential calculations. The minimal Jastrow function is a function of the electron-electron coordinates with the correct antiparallel- and parallel-spin cusps, but more typically it is written as a product of electron-nucleus, electron-electron and electron-electron-nucleus Jastrows:

$$\mathcal{J} \;=\; \prod_{\alpha i} \exp\left(A_{\alpha i}\right) \prod_{ij} \exp\left(B_{ij}\right) \prod_{\alpha ij} \exp\left(C_{\alpha ij}\right) \tag{42}$$

where α indexes the nuclei, and i and j index the electrons. Adding higher-order Jastrows, say 3 electrons and a nucleus, leads to minor gains relative to the increase in computational cost [31]. In all there are 4 kinds of parameters that can be optimized: a) the linear coefficients d_n multiplying the determinants, b) the orbital coefficients that specify the orbitals in terms of the basis functions, c) the exponents of the basis functions, and d) the Jastrow parameters. The number of basis exponents and the number of Jastrow parameters scale linearly in the number of atomic species, the number of orbital parameters scale as the number of electrons times the number of basis functions, and the number of d_n can be combinatorially large in the number of basis functions and electrons. However, in practice only a very tiny fraction of these are used. In fact one of the big advantages of QMC methods is that because of the effectiveness of the Jastrow in capturing some of the correlation, the number of determinants can be orders of magnitude smaller than in other methods for the same accuracy.

In real-space QMC, the Jastrow is particularly effective in introducing the so-called "dynamic correlation." The multi-determinantal expansion is used mostly to include "near-degeneracy" or "static" correlation, which requires relatively few determinants. Consequently the number of determinants required to obtain a given energy is often orders of magnitude less in the presence of a flexible Jastrow than in its absence. Moreover, the size of the single-particle basis needed is reduced, particularly if the exponents of the basis functions are also optimized (though this is rarely done). Note that although the Jastrow does not directly change the nodes of $\Psi_{\rm T}(\mathbf{R})$, when the wavefunction is optimized in the presence of the Jastrow it indirectly changes the nodes of $\Psi_{\rm T}(\mathbf{R})$ and thereby enables accurate fixed-node DMC energies with compact wavefunctions.

The Jastrow plays another very important role in real-space QMC. The local energy, $E_{\rm L}(\mathbf{R})$, diverges to $\pm\infty$ at electron-nucleus and electron-electron coincidences, unless cusp-conditions [32, 33] are imposed. The electron-nucleus cusps can be imposed by placing constraints on the orbitals (both intra-atom and inter-atom contributions need to be considered) but the electron-electron cusp requires the Jastrow function. Once the cusp conditions are imposed, $E_{\rm L}(\mathbf{R})$ becomes finite at particle coincidences (though there is still a finite discontinuity in the limit that two electrons approach a nucleus [34]). This greatly reduces the fluctuations of the local energy and improves the efficiency of both VMC and DMC.

The multideterminant expansion is typically chosen by performing a small complete active space self consistent field (CASSCF) calculation and keeping the most important determinants. However, for challenging molecules there are several determinants outside the CAS space of affordable size that are more important than some of the determinants that are included from within the CAS space. Consequently, as the number of included determinants is increased, convergence of the energy is observed but to a spurious value. This problem is solved by selecting the determinants from a configuration interaction calculation [35, 36], which selects the most important determinants from the entire Hilbert space.

5.2 Symmetry-projected broken-symmetry mean-field wavefunctions

Recently there has been remarkable increase in the number of determinants that can be efficiently included in the multideterminant expansion [37–39]. Nevertheless, since the number of determinants grows combinatorially in the size of the single-particle basis and the number of electrons (though with a much reduced prefactor because of the Jastrow) there is considerable interest in using more flexible mean-field states than a single determinant, namely the antisymmetrical geminal power (AGP) [40] and Pfaffian [41,42] as the antisymmetric part of QMC trial wavefunctions [43]. Recently these ideas have been extended to use, in QMC trial wavefunctions, symmetry-projected broken-symmetry mean-field states [43], first employed in traditional quantum chemistry calculations [44]. The symmetries that are broken and restored are combinations of particle-number, S^2, S_z and complex conjugation. The most flexible of these breaks bonds correctly and yields remarkably good potential energy curves [43], considering that the computational cost is only marginally greater than that for a single determinant.

5.3 Backflow wavefunctions

Another direction (which can be combined with the above) is to use backflow wavefunctions, first introduced by Feynman to describe correlations in liquid Helium. The orbitals are evaluated at backflow-transformed coordinates that depend on the positions of all the electrons. These give considerably improved energies and fluctuations [45–48], but they incur a factor of N increase in the computational cost since the determinant lemma and Sherman-Morrison formula cannot be used achieve an $\mathcal{O}(N^2)$ cost for updating determinants and their derivatives when a single electron is moved.

5.4 Wavefunctions in orbital-space QMC

Although most VMC and PMC calculations for electronic systems have been done in real space, recently there has been considerable interest in orbital-space QMC [49–51]. The orbital-space Jastrow plays a rather different role than its real-space counterpart—its most important contribution is to suppress double occupancy of orbitals and so it is effective in describing static correlations. The straightforward approach has a computational cost that scales as $\mathcal{O}(N^4)$ for constant error per electron, but ideas borrowed from the semistochastic heatbath configuration interaction (SHCI) method [52], reduce this cost to $\mathcal{O}(N^2)$ [51].

6 Optimization of trial wavefunctions

Accurate variational wavefunctions typically have a large number of linear and nonlinear parameters, that have to be optimized. As many as several hundred thousand have been used [53, 39]. One of the interesting features of QMC methods is that sometimes tiny changes in the algorithm, that may go unnoticed, can make a dramatic improvement to its efficiency. It is very helpful to think about ideal situations where the variance becomes zero; although this may never be achieved in practice, it is helpful for designing algorithms with low variance. This is particularly true in the case of wavefunction optimization algorithms. At the present time the 3 most used optimization algorithms are the stochastic reconfiguration method [54, 40], the Newton method [55], and the linear method [56–58]. We will present the Newton method in some detail to illustrate the sort of small algorithmic changes that can provide large efficiency improvements, but will mention the other two methods only cursorily.

Optimizing the wavefunctions is important for several reasons. We enumerate below the various errors in QMC calculations that are reduced by optimizing the wavefunction:

1. Statistical error (both the rms fluctuations of E_{L} and the autocorrelation time)
2. Variational error in E_{VMC}
3. Fixed-node error in E_{DMC}
4. Time-step error in DMC
5. Population control error in DMC
6. Pseudopotential locality error in DMC when using nonlocal pseudopotentials
7. Error of mixed estimates of observables that do not commute with the Hamiltonian in DMC

In fact all errors, aside from the finite-size errors present in calculations of periodic systems, benefit from wavefunction optimization.

The next question for optimizing wavefunctions is: precisely what quantity do we want to optimize? Possible choices are:

$$1)\ \text{minimize}\ \ E_{\rm VMC} = \frac{\langle\Psi_{\rm T}|H|\Psi_{\rm T}\rangle}{\langle\Psi_{\rm T}|\Psi_{\rm T}\rangle} = \langle E_{\rm L}\rangle_{\Psi_{\rm T}^2} \tag{43}$$

$$2)\ \text{minimize}\ \ \sigma^2_{\rm VMC} = \frac{\langle\Psi_{\rm T}|(H-E_{\rm T})^2|\Psi_{\rm T}\rangle}{\langle\Psi_{\rm T}|\Psi_{\rm T}\rangle} = \left\langle E^2_{\rm L}(\mathbf{R}_i)\right\rangle_{\Psi_{\rm T}^2} - \langle E_{\rm L}(\mathbf{R}_i)\rangle^2_{\Psi_{\rm T}^2} \tag{44}$$

$$3)\ \text{maximize}\ \ \Omega^2 = \frac{|\,\langle\Psi_{{\rm F}N}|\Psi_{\rm T}\rangle\,|^2}{\langle\Psi_{{\rm F}N}|\Psi_{{\rm F}N}\rangle\,\langle\Psi_{\rm T}|\Psi_{\rm T}\rangle} = \frac{\left\langle \frac{\Psi_{{\rm F}N}}{\Psi_{\rm T}}\right\rangle^2_{\Psi_{\rm T}^2}}{\left\langle \left|\frac{\Psi_{{\rm F}N}}{\Psi_{\rm T}}\right|^2\right\rangle_{\Psi_{\rm T}^2}} \tag{45}$$

$$4)\ \text{minimize}\ \ E_{\rm DMC} = \frac{\langle\Psi_{{\rm F}N}|H|\Psi_{\rm T}\rangle}{\langle\Psi_{{\rm F}N}|\Psi_{\rm T}\rangle} = \langle E_{\rm L}\rangle_{|\Psi_{{\rm F}N}\Psi_{\rm T}|} \tag{46}$$

In fact all of the above have been studied to some extent in published and unpublished work. For an infinitely flexible wave function all these optimization criteria will yield the exact wavefunction (except that minimizing σ could yield an excited state) but for the imperfect functional forms used in practice they differ. Since the final energy is obtained from $E_{\rm DMC}$ rather than $E_{\rm VMC}$, the most desirable option is the last one. However, the very limited experience gained so far indicates that minimizing $E_{\rm VMC}$ with flexible wavefunctions results in approximately minimizing also $E_{\rm DMC}$, so the additional effort of minimizing $E_{\rm DMC}$ is not worthwhile. Hence the common practice is to minimize $E_{\rm VMC}$ or a linear combination of $E_{\rm VMC}$ and $\sigma^2_{\rm VMC}$, with most of the weight on $E_{\rm VMC}$.

Early work on optimizing wavefunctions used variance minimization [59, 60] because early attempts at minimizing $E_{\rm VMC}$ required many more Monte Carlo samples than for minimizing $\sigma^2_{\rm VMC}$. This is explained in Fig. 3. In green we schematically show the local energies for some sampled points. Since the wavefunction is not exact these energies have a spread. In red we show how these energies change as one attempts to minimize $E_{\rm VMC}$. Most of the energies go down and some may go up, but the average on the sample goes down. As the wavefunction is made more flexible the average may go down well below the true energy. Now if one draws a fresh sample, then we find that the average on this sample has actually gone up rather than down! Once one realizes the problem with energy minimization it becomes clear that the way around it is to minimize the variance of the local energies as shown in blue. In that case local energies move closer to each other and the average energy also tends to go down, less so than for energy minimization, but this gain in energy is genuine—an independent sample shows the same effect. Of course energy minimization will work if one has a sufficiently large number of samples, but the point of the above thought experiment is that the number of samples needed is smaller for variance minimization than for naive energy minimization. We next discuss more clever ways to minimize the energy that overcome this problem by minimizing the expectation value of the energy without minimizing the energy of the chosen sample.

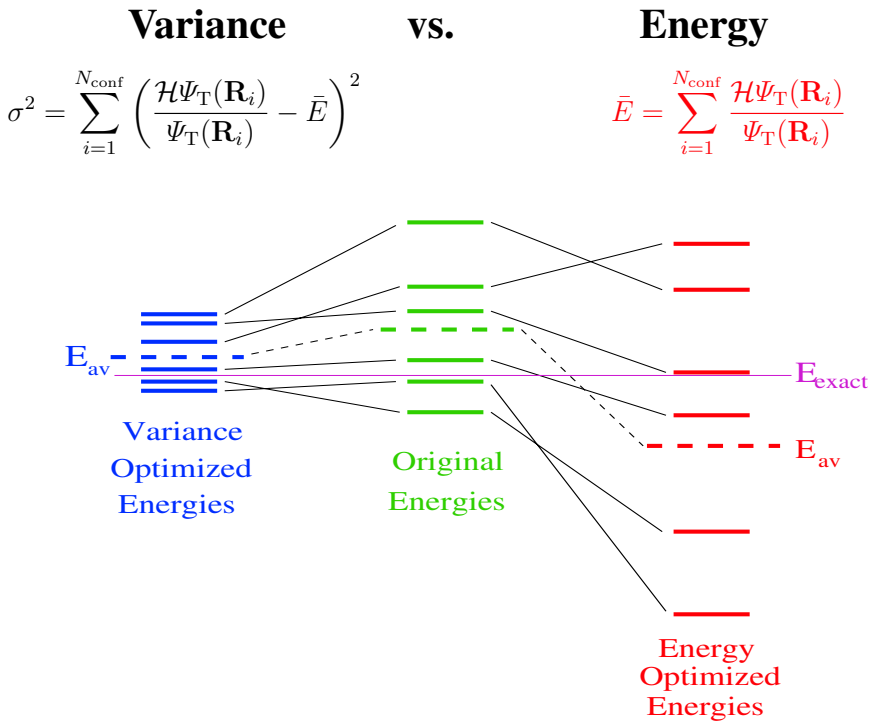

Fig. 3: *Why variance minimization requires fewer samples than naive energy minimization.*

6.1 Newton method

6.1.1 Minimization of $E_{\rm VMC}$

In the Newton method, the parameter changes, $\delta\mathbf{p}$, are obtained by solving linear equations

$$\mathbf{h}\,\delta\mathbf{p} = -\mathbf{g}\,, \tag{47}$$

where $\mathbf{h}$ is the Hessian and $\mathbf{g}$ the gradient of $E_{\rm VMC}$ with respect to the variational parameters. In the rest of this section $\Psi_{\rm T}$ and $E_{\rm VMC}$ are the only wavefunction and energy of relevance, so in the interest of notational brevity we replace these by Ψ and $\bar{E}$ respectively.

$$\bar{E} = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \langle E_{\rm L}\rangle_{|\Psi|^2}\,; \qquad E_{\rm L}(\mathbf{R}) = \frac{H\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \tag{48}$$

where the notation $\langle\cdots\rangle_{|\Psi|^2}$ denotes a Monte Carlo average over samples drawn from $|\Psi|^2$. Following Ref. [55] the energy gradient wrt parameter p_i, denoted by $\bar{E}_i$, is

$$\bar{E}_i = \frac{\langle\Psi_i|H\Psi\rangle + \langle\Psi|H\Psi_i\rangle}{\langle\Psi|\Psi\rangle} - 2\frac{\bar{E}\,\langle\Psi|\Psi_i\rangle}{\langle\Psi|\Psi\rangle} = 2\frac{\langle\Psi_i|H\Psi\rangle - \bar{E}\,\langle\Psi|\Psi_i\rangle}{\langle\Psi|\Psi\rangle} \tag{49}$$

$$\approx \left\langle \frac{\Psi_i}{\Psi}E_{\rm L} + \frac{H\Psi_i}{\Psi} - 2\bar{E}\frac{\Psi_i}{\Psi} \right\rangle_{\psi^2} \approx 2\left\langle \frac{\Psi_i}{\Psi}(E_{\rm L} - \bar{E}) \right\rangle_{\psi^2} \tag{50}$$

In Eq. (49) we use the Hermiticity of the Hamiltonian to go from the expressions on the left to that on the right. The expressions in Eq. (50) are the MC estimates of the corresponding

expressions in Eq. (49). Note however, that the expressions on the left and right of Eq. (50) become precisely equal only in the limit of an infinite sample. For a finite sample, the expression on the right has much smaller fluctuations than the one on the left for sufficiently good trial wavefunctions[3] because it has zero variance in the limit that Ψ is exact.
Rewriting the gradient as

$$\bar{E}_i \;=\; 2\frac{\langle \Psi_i|H\Psi\rangle - \bar{E}\,\langle\Psi|\Psi_i\rangle}{\langle\Psi|\Psi\rangle} \;=\; 2\frac{\langle \Psi_i\Psi(E_{\rm L}-\bar{E})\rangle}{\langle|\Psi|^2\rangle} \tag{51}$$

and taking the derivative wrt p_j, we obtain the following expression for the hessian:

$$\bar{E}_{ij} = 2\left[\frac{\langle(\Psi_{ij}\Psi+\Psi_i\Psi_j)(E_{\rm L}-\bar{E})\rangle + \langle\Psi_i\Psi(E_{{\rm L},j}-\bar{E}_j)\rangle - \bar{E}_i\,\langle\Psi\Psi_j\rangle}{\langle|\Psi|^2\rangle}\right] \tag{52}$$

$$= 2\left[\left\langle\left(\frac{\Psi_{ij}}{\Psi}+\frac{\Psi_i\Psi_j}{|\Psi|^2}\right)(E_{\rm L}-\bar{E})\right\rangle_{\psi^2} - \left\langle\frac{\Psi_i}{\Psi}\right\rangle_{\psi^2}\bar{E}_j - \left\langle\frac{\Psi_j}{\Psi}\right\rangle_{\psi^2}\bar{E}_i + \left\langle\frac{\Psi_i}{\Psi}E_{{\rm L},j}\right\rangle_{\psi^2}\right]. \tag{53}$$

What can be done to reduce the variance of this expression? The last term is not symmetric in i and j because we started from the right hand rather than the left hand expression in Eq. (50), so we can symmetrize it, but that does not affect the variance appreciably.
Next note that $\langle E_{{\rm L},i}\rangle_{|\Psi|^2} = \frac{\langle|\Psi|^2\left(\frac{H\Psi}{\Psi}\right)_i\rangle}{\langle|\Psi|^2\rangle} = \frac{\left\langle|\Psi|^2\left(\frac{H\Psi_i}{\Psi}-\frac{\Psi_i}{|\Psi|^2}H\Psi\right)\right\rangle}{\langle|\Psi|^2\rangle} = \frac{\langle\Psi H\Psi_i-\Psi_i H\Psi\rangle}{\langle|\Psi|^2\rangle} = 0$, so we are free to add terms such as $\left\langle\frac{\Psi_j}{\Psi}\right\rangle_{\psi^2}\langle E_{{\rm L},i}\rangle_{\psi^2}$. Now, the fluctuations of the covariance $\langle ab\rangle - \langle a\rangle\langle b\rangle$ are smaller than those of the product $\langle ab\rangle$, when $\sqrt{\langle a^2\rangle - \langle a\rangle^2} \ll |\langle a\rangle|$ and $\langle b\rangle$ is small. ($\langle E_{{\rm L},i}\rangle_{\psi^2}$ is 0 on an infinite sample and small on a finite sample.) Hence we make the replacement

$$\left\langle\frac{\Psi_i}{\Psi}E_{{\rm L},j}\right\rangle_{\psi^2} \to \frac{1}{2}\left(\left\langle\frac{\Psi_i}{\Psi}E_{{\rm L},j}\right\rangle_{\psi^2} - \left\langle\frac{\Psi_i}{\Psi}\right\rangle_{\psi^2}\langle E_{{\rm L},j}\rangle_{\psi^2} + \left\langle\frac{\Psi_j}{\Psi}E_{{\rm L},i}\right\rangle_{\psi^2} - \left\langle\frac{\Psi_j}{\Psi}\right\rangle_{\psi^2}\langle E_{{\rm L},i}\rangle_{\psi^2}\right).$$

The resulting expression is

$$\begin{aligned}\bar{E}_{ij} = 2&\left[\left\langle\left(\frac{\Psi_{ij}}{\Psi}+\frac{\Psi_i\Psi_j}{|\Psi|^2}\right)(E_{\rm L}-\bar{E})\right\rangle_{\psi^2} - \left\langle\frac{\Psi_i}{\Psi}\right\rangle_{\psi^2}\bar{E}_j - \left\langle\frac{\Psi_j}{\Psi}\right\rangle_{\psi^2}\bar{E}_i\right] \\ &+\left\langle\frac{\Psi_i}{\Psi}E_{{\rm L},j}\right\rangle_{\psi^2} - \left\langle\frac{\Psi_i}{\Psi}\right\rangle_{\psi^2}\langle E_{{\rm L},j}\rangle_{\psi^2} + \left\langle\frac{\Psi_j}{\Psi}E_{{\rm L},i}\right\rangle_{\psi^2} - \left\langle\frac{\Psi_j}{\Psi}\right\rangle_{\psi^2}\langle E_{{\rm L},i}\rangle_{\psi^2} \qquad (54)\\ = 2&\left[\left\langle\left(\frac{\Psi_{ij}}{\Psi}-\frac{\Psi_i\Psi_j}{|\Psi|^2}\right)(E_{\rm L}-\bar{E})\right\rangle_{\psi^2} \quad (0 \text{ for } p_i \text{ linear in exponent})\right. \\ &\left.+2\left\langle\left(\frac{\Psi_i}{\Psi}-\left\langle\frac{\Psi_i}{\Psi}\right\rangle_{\psi^2}\right)\left(\frac{\Psi_j}{\Psi}-\left\langle\frac{\Psi_j}{\Psi}\right\rangle_{\psi^2}\right)(E_{\rm L}-\bar{E})\right\rangle_{\psi^2}\right] \\ &+\left\langle\frac{\Psi_i}{\Psi}E_{{\rm L},j}\right\rangle_{\psi^2} - \left\langle\frac{\Psi_i}{\Psi}\right\rangle_{\psi^2}\langle E_{{\rm L},j}\rangle_{\psi^2} + \left\langle\frac{\Psi_j}{\Psi}E_{{\rm L},i}\right\rangle_{\psi^2} - \left\langle\frac{\Psi_j}{\Psi}\right\rangle_{\psi^2}\langle E_{{\rm L},i}\rangle_{\psi^2}. \qquad (55)\end{aligned}$$

[3] A subtle point is that the zero-variance estimator on the RHS of Eq. (50) is also an infinite-variance estimator when $\Psi_{\rm T}$ is not exact and the parameters being optimized can change the nodes of the wavefunction. So, for poor wavefunctions it may be preferable to use a known alternative expression that has finite variance for approximate $\Psi_{\rm T}$.

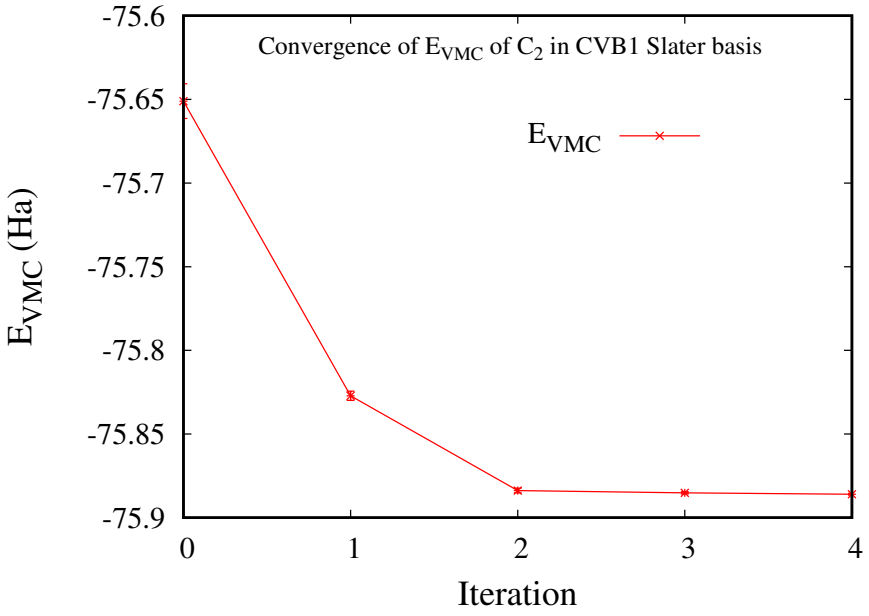

Fig. 4: *Energy convergence of C_2 using a Slater CVB1 basis with 48 basis functions, optimizing 24 Jastrow, 164 CSF and 90 orbital parameters. If the basis exponents are optimized then with 44 basis functions, a considerably more compact wavefunction with just 13 CSFs gives an energy better than this by 1 mHa (not shown). However, converging the wavefunction takes many more optimization iterations.*

The expressions in the first two lines have zero variance in the limit of an exact trial wavefunction and the last line has a much reduced variance compared to our starting point, Eq. (53). Finally, we note that the individual summands in Eq. (55) have a leading 3^{rd}-order divergence near the nodes of the trial wavefunction for parameters that change wavefunction nodes, but the leading divergences cancel each other, leaving a 2^{nd}-order divergence which gives a finite expectation value but infinite variance. Despite having infinite variance estimators for the gradient and the hessian, the method works remarkably well for small systems even for parameters that move the nodes of Ψ_{T}, as shown in Fig. 4. If finite variance estimators are needed they are obtained by using nodeless Ψ_{G} with $\Psi_{\mathrm{G}} \approx |\Psi_{\mathrm{T}}|$ except near the nodes of Ψ_{T} [61].

6.1.2 Minimization of σ_{VMC}

As mentioned earlier, another option is to minimize the variance of the energy

$$\sigma^2 \;=\; \frac{\int \mathrm{d}^{3N}R\, |\Psi|^2 (E_{\mathrm{L}} - \bar{E})^2}{\int \mathrm{d}^{3N}R\, |\Psi|^2}\,. \tag{56}$$

The exact gradient and hessian have been derived and used but the following simpler option works just as well. When the parameters are changed, the variance changes both because E_{L} changes and because the distribution of sampled points changes. If we ignore the latter, which means that we are computing the variance on a fixed set of Monte Carlo configurations, we get

$$(\sigma^2)_i \;=\; 2\left\langle E_{\mathrm{L},i}(E_{\mathrm{L}} - \bar{E})\right\rangle = 2\left\langle (E_{\mathrm{L},i} - \bar{E}_i)(E_{\mathrm{L}} - \bar{E})\right\rangle. \tag{57}$$

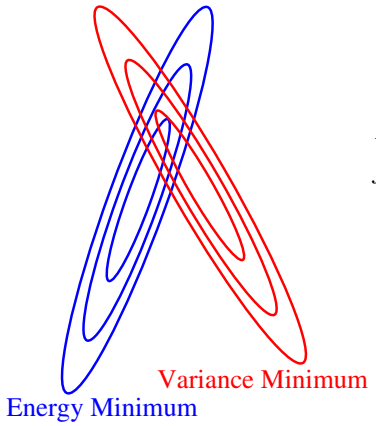

Fig. 5: *Schematic of energy and variance contours. Adding a small fraction of variance to the energy in the function to be minimized can reduce the variance while raising the energy only very slightly.*

In the case of energy minimization we added a term with zero expectation value to reduce the variance of the energy. Similarly, in the right hand expression above we have added a term to minimize the variance of the variance.
Then the (positive definite) Levenberg-Marquardt approximation to the Hessian matrix is

$$(\sigma^2)_{ij} \;=\; 2\left\langle (E_{\mathrm{L},i}-\bar{E}_i)(E_{\mathrm{L},j}-\bar{E}_j)\right\rangle . \tag{58}$$

6.1.3 Mixed minimization

A linear combination of the energy and variance can be optimized simply by using the same linear combination for the gradient and the hessian. The reason for possibly wanting to use a linear combination is that by adding in small fraction of variance minimization (say 0.01-0.05) to energy minimization, it may be possible to reduce the variance without appreciably raising the energy. To explain why, we show in Fig. 5 schematic contours for the energy and variance. If only the energy is minimized then the parameters may lie anywhere close to the bottom of the energy well, but if a small fraction of the variance is added to the mix then the portion of the energy well closer to the bottom of the variance minimum is favored.
Fig. 6 has convergence plots for energy, variance, and mixed minimization for the NO_2 molecule using a pseudopotential. The optimization takes only a few iterations. We see that mixed minimization lowers the variance while raising the energy only slightly. However, energy minimization has the smallest autocorrelation time, T_{corr}, so the benefit from doing a mixed minimization rather than just an energy minimization is small.

6.1.4 Stabilizing the optimization

If we add a positive constant a_{diag} to the diagonal of the hessian matrix, i.e., $\mathbf{h} \rightarrow \mathbf{h} + a_{\mathrm{diag}}\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, then the proposed moves get smaller and rotate from the Newtonian direction to the steepest descent direction. We use this to stabilize the optimization. At each optimization iteration, we perform two kinds of MC runs. First we do a run to compute the gradient and the hessian. Second, we do a 10-times shorter MC run, that does 3 correlated sampling calculations with parameters obtained from this gradient and hessian and 3 values of a_{diag} that differ from each other by factors of 10 to get the corresponding 3 values of the

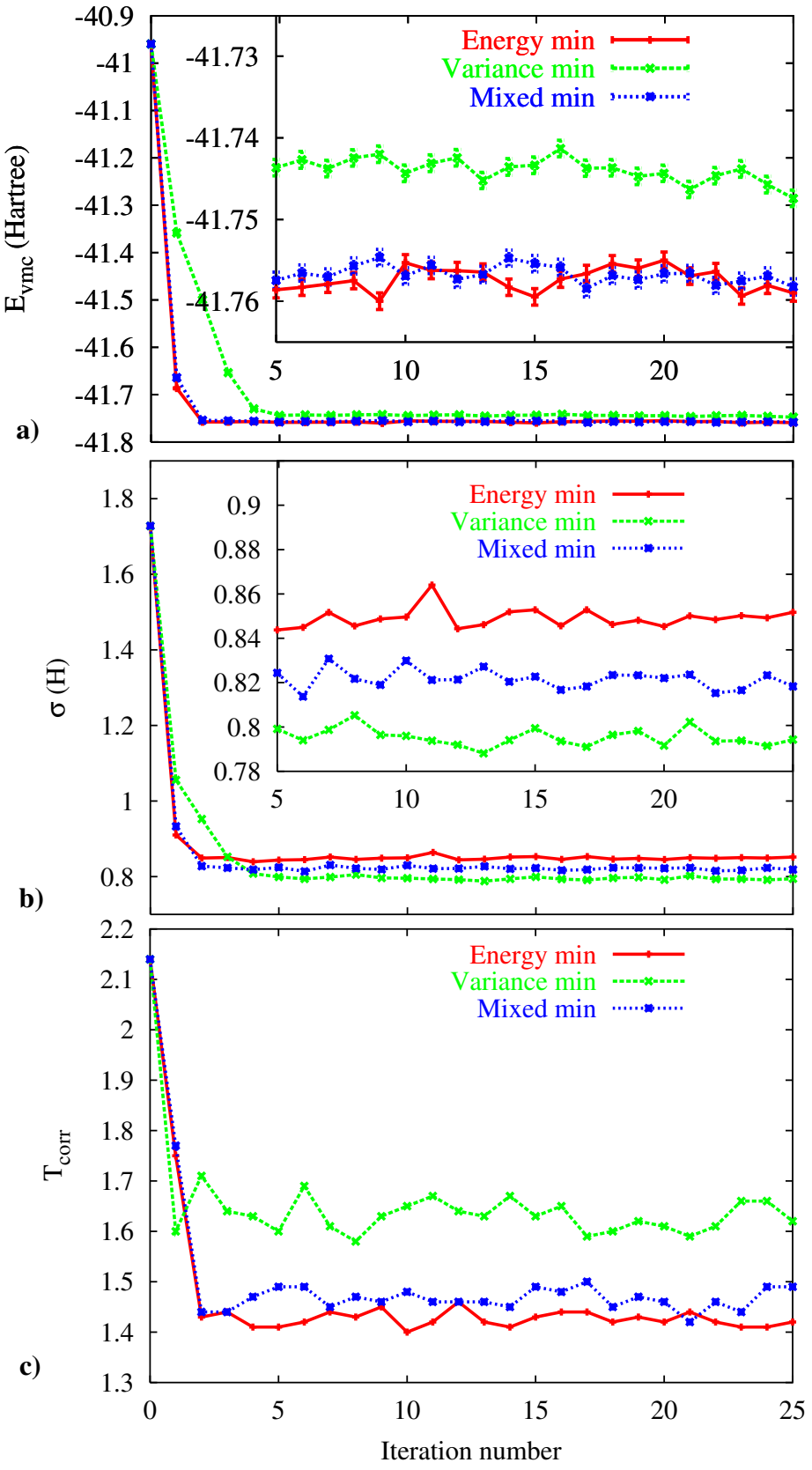

Fig. 6: *Optimization of Jastrow parameters for an NO_2 molecule using energy minimization, variance minimization, and mixed energy and variance minimization. Convergence is achieved in just a few iterations. Mixed minimization lowers σ while raising the energy only slightly. T_{corr} is smallest for energy minimization.*

energy (or whatever linear combination of the energy and variance that we are optimizing). The optimal value of $a_{\rm diag}$ is then predicted from a parabolic fit of the 3 energy values with some bounds imposed, provided the 2^{nd} derivative of the fit is positive; else the $a_{\rm diag}$ value that gives the lowest energy is chosen. In addition, $a_{\rm diag}$ is forced to increase if the parameter variations exceed some chosen thresholds, or if some parameters exit their allowed domain of variation (e.g., if a basis exponent becomes negative). Despite these precautions, the energy occasionally goes up a lot, in which case one goes back to the earlier gradient and hessian and proposes new parameter changes with a larger value of $a_{\rm diag}$. This only happens for the larger systems that we study; for small systems such as NO_2 shown in Fig. 6 the optimization converges within a few iterations with no need to go back to an earlier iteration. Note that even when it is necessary to go back, the entire procedure is automatic; there is no need for human intervention. Finally we note that we do not necessarily choose the parameters from the last optimization iteration as the final best parameters. Instead we choose the parameters from the iteration for which $E_{\rm VMC} + 3\sigma_{\rm VMC}$ is lowest. In fact, this often is the last iteration because in order to save time, we use a small number of MC steps in the first iteration and gradually increase the number of MC steps with each new iteration (upto some maximum), so even after $E_{\rm VMC}$ has converged, $\sigma_{\rm VMC}$ continues to go down with iteration number.

6.2 Linear method

The linear optimization method is probably at the present time the most commonly used optimization method. It was originally developed for linear parameters [62], but was extended to nonlinear parameters [56–58]. It has the advantage that it does not require calculating the 2^{nd} derivatives of $\Psi_{\rm T}(\mathbf{R})$ and it converges just as quickly as the Newton method. Similar to what we have described for the Newton method, there are far from obvious changes that need to be made to the straightforward version of the method in order to make it efficient. The details can be found in the original literature [62, 56–58].

6.3 Stochastic reconfiguration method

The stochastic reconfiguration method [54, 40] can be viewed as an approximation to the Newton method. Of the 3 methods mentioned in this lecture, it requires the least computation per optimization iteration, but it typically takes several times as many iterations to converge. Although it converges more slowly, when applied to heavy systems it can sometimes oscillate less than the other methods because it requires fewer derivatives and suffers less from infinite-variance estimators.

Optimization of many parameters: When the number of parameters to be optimized is large (more than a few thousand) storage of the relevant matrices (hessian for Newton method, Hamiltonian and overlap for the linear method, overlap for the stochastic reconfiguration method) becomes a problem. However, they need not be stored if iterative methods are used to solve the relevant equations. Then it becomes practical to optimize on the order of 10^5–10^6 parameters. Details can be found in Refs. [53, 8].

7 Outlook

I hope this lecture has given you some flavor of the sort of thinking that goes into designing accurate and efficient QMC algorithms. QMC methods have been worked upon less that some other electronic structure methods. Hence, I think there is still considerable room for improvement. An example of that would be the development of a continuous real-space PMC method that does not have a time-step error. In fact one of the earliest PMC methods invented [18–20] does not have a time-step error. However, despite this major advantage, it is not used anymore because it is much less efficient that DMC. Further, there exist, in a discrete space, efficient PMC methods that use the exponential projector with no time-step error. It seems possible that one could invent an efficient time-step error free algorithm for continuous real-space PMC as well.

Acknowledgments

I acknowledge valuable discussions with Sandeep Sharma. This work was supported by the AFOSR through grant FA9550-18-1-0095.

References

[1] D.M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995)

[2] M.P. Nightingale and C.J. Umrigar (eds.): *Quantum Monte Carlo Methods in Physics and Chemistry*, NATO ASI Ser. C 525 (Kluwer, Dordrecht, 1999)

[3] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001)

[4] J. Kolorenc and L. Mitas, Rep. Prog. Phys. **74**, 026502 (2011)

[5] S. Zhang, in E. Pavarini, E. Koch, and U. Schollwöck (eds.) *Emergent Phenomena in Correlated Matter*, Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013)

[6] J. Toulouse, R. Assaraf, and C.J. Umrigar, p. 285 in P. Hoggan and T. Özdoğan (eds.): *Advances in Quantum Chemistry*, Vol. 73 (2015)

[7] F. Becca and S. Sorella: *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, 2017)

[8] J. Feldt and C. Filippi, in L. Gonzaález and R. Lindh (eds.): *Quantum Chemistry and Dynamics of Excited States: Methods and Applications* (Wiley-VCH, 2019)

[9] C. Everett and E. Cashwell.: *Third Monte Carlo sampler. Revision and extension of samplers I and II*. Tech. Rep., Los Alamos National Laboratory (1983)

[10] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, J. Chem. Phys. **21**, 1087 (1953)

[11] W.K. Hastings, Biometrika **57**, 97 (1970)

[12] C.J. Umrigar, Phys. Rev. Lett. **71**, 408 (1993)

[13] C.J. Umrigar, p. 129 in [2]

[14] M.P. Nightingale and H. Blote, Phys. Rev. Lett. **60**, 1562 (1988)

[15] C.J. Umrigar, M.P. Nightingale, and K.J. Runge, J. Chem. Phys. **99**, 2865 (1993)

[16] J.B. Anderson, J. Chem. Phys. **65**, 4121 (1976)

[17] P.J. Reynolds, D.M. Ceperley, B.J. Alder, and W.A. Lester, J. Chem. Phys. **77**, 5593 (1982)

[18] M.H. Kalos, D. Levesque, and L. Verlet, Phys. Rev. A **9**, 2178 (1974)

[19] D.M. Ceperley and M.H. Kalos, pp. 145–194 in K. Binder (ed.): *Monte Carlo Methods in Statistical Physics* (Springer, Berlin, 1979)

[20] K.E. Schmidt and M.H. Kalos, p. 125 in K. Binder (ed.): *Applications of the Monte Carlo Method in Statistical Physics*, Topics in Current Physics, Vol. 36 (Springer, 1984)

[21] G.H. Booth, A.J.W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009)

[22] D. Cleland, G.H. Booth, and A. Alavi, J. Chem. Phys. **132**, 041103 (2010)

[23] F.R. Petruzielo, A.A. Holmes, H.J. Changlani, M.P. Nightingale, and C.J. Umrigar, Phys. Rev. Lett. **109**, 230201 (2012)

[24] S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003)

[25] D.M. Ceperley and L. Mitas, Advances in Chemical Physics, Vol. 93, 1–38 (John Wiley & Sons, 1996)

[26] F. Schautz and C. Filippi, J. Chem. Phys. **120**, 10931 (2004)

[27] F. Schautz, F. Buda, and C. Filippi, J. Chem. Phys. **121**, 5836 (2004)

[28] H. Zulfikri, C. Amovilli, and C. Filippi, J. Chem. Theory Comput. **12**, 1157 (2016)

[29] M.H. Kolodrubetz, J.S. Spencer, B.K. Clark, and W.M.C. Foulkes, J. Chem. Phys. **138**, 024110 (2013)

[30] F.R. Petruzielo, J. Toulouse, and C.J. Umrigar, J. Chem. Phys. **134**, 064104 (2011)

[31] C.J. Huang, C.J. Umrigar, and M.P. Nightingale, J. Chem. Phys. **107**, 3007 (1997)

[32] T. Kato, Comm. Pure Appl. Math. **10**, 151 (1957)

[33] R.T. Pack and W. Byers-Brown, J. Chem. Phys. **45**, 556 (1966)

[34] C.R. Myers, C.J. Umrigar, J.P. Sethna, and J. Morgan, Phys. Rev. A **44**, 5537 (1991)

[35] A. Scemama, A. Benali, D. Jacquemin, M. Caffarel, and P.F. Loos, J. Chem. Phys. **149**, 034108 (2018)

[36] M. Dash, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **14**, 4176 (2018)

[37] B.K. Clark, M.A. Morales, J. McMinis, J. Kim, and G.E. Scuseria, J. Chem. Phys. **135**, 244105 (2011)

[38] C. Filippi, R. Assaraf, and S. Moroni, J. Chem. Phys. **144**, 194105 (2016)

[39] R. Assaraf, S. Moroni, and C. Filippi, J. Chem. Theory Comput. **13**, 5273 (2017)

[40] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004)

[41] M. Bajdich, L. Mitas, L.K. Wagner, and K.E. Schmidt, Phys. Rev. B **77**, 115112 (2008)

[42] M. Bajdich, L. Mitas, G. Drobny, L. Wagner, and K. Schmidt, Phys. Rev. Lett. **96**, 130201 (2006)

[43] A. Mahajan and S. Sharma, J. Phys. Chem. A **123**, 3911 (2019)

[44] G.E. Scuseria, C.A. Jiménez-Hoyos, and T.M. Henderson, J. Chem. Phys. **135**, 124108 (2011)

[45] K.E. Schmidt, M.A. Lee, M.H. Kalos, and G.V. Chester, Phys. Rev. Lett. **47**, 807 (1981)

[46] M.D. Brown and J.R. Trail and P. Lopez Rios and R.J. Needs, J. Chem. Phys. **126**, 224110 (2007)

[47] P. Seth, P.L. Rios, and R.J. Needs, J. Chem. Phys. **134**, 084105 (2011)

[48] M. Holzmann and S. Moroni, Phys. Rev. B **99**, 085121 (2019)

[49] E. Neuscamman, Mol. Phys. **114**, 577 (2016)

[50] H. Wei and E. Neuscamman, J. Chem. Phys. **149**, 184106 (2018)

[51] I. Sabzeyari and S. Sharma, J. Chem. Theory Comput. **14**, 6276 (2018)

[52] A.A. Holmes, N.M. Tubman, and C.J. Umrigar, J. Chem. Theory Comput. **12**, 3674 (2016)

[53] E. Neuscamman, C.J. Umrigar, and G.K.L. Chan, Phys. Rev. B **85**, 045103 (2012)

[54] S. Sorella, Phys. Rev. B **64**, 024512 (2001)

[55] C.J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005)

[56] C.J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R.G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007)

[57] J. Toulouse and C.J. Umrigar, J. Chem. Phys. **126**, 084102 (2007)

[58] J. Toulouse and C.J. Umrigar, J. Chem. Phys. **128**, 174101 (2008)

[59] R.L. Coldwell, Int. J. Quant. Chem. **11**, 215 (1977)

[60] C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988)

[61] C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501 (2008)

[62] M.P. Nightingale and Vilen Melik-Alaverdian, Phys. Rev. Lett. **87**, 043041 (2001)