

TUTORIAL | DECEMBER 30 2024

A brief introduction to the diffusion Monte Carlo method and the fixed-node approximation

Alfonso Annarelli ; Dario Alfè ; Andrea Zen  



J. Chem. Phys. 161, 241501 (2024)

<https://doi.org/10.1063/5.0232424>



Articles You May Be Interested In

A particle-based ellipsoidal statistical Bhatnagar–Gross–Krook solver with variable weights for the simulation of large density gradients in micro- and nano-nozzles

Physics of Fluids (November 2020)

Multiple structural transformations in Lennard-Jones clusters: Generic versus size-specific behavior

J. Chem. Phys. (May 2006)

Entanglement of four-qubit systems: A geometric atlas with polynomial compass II (the tame world)

J. Math. Phys. (February 2017)



The Journal of Chemical Physics

Special Topics Open for Submissions

[Learn More](#)

A brief introduction to the diffusion Monte Carlo method and the fixed-node approximation

Cite as: J. Chem. Phys. 161, 241501 (2024); doi: 10.1063/5.0232424

Submitted: 7 August 2024 • Accepted: 11 December 2024 •

Published Online: 30 December 2024



View Online



Export Citation



CrossMark

Alfonso Annarelli,^{1,a)}  Dario Alfè,^{1,2,3,b)}  and Andrea Zen^{1,2,c)} 

AFFILIATIONS

¹Dipartimento di Fisica Ettore Pancini, Università di Napoli Federico II, Monte S. Angelo, I-80126 Napoli, Italy

²Department of Earth Sciences, University College London, Gower Street, London WC1E 6BT, United Kingdom

³Thomas Young Centre and London Centre for Nanotechnology, 17-19 Gordon Street, London WC1H 0AH, United Kingdom

^{a)}alfonso.annarelli@gmail.com

^{b)}dario.alfè@unina.it

^{c)}Author to whom correspondence should be addressed: andrea.zen@unina.it

ABSTRACT

Quantum Monte Carlo (QMC) methods represent a powerful family of computational techniques for tackling complex quantum many-body problems and performing calculations of stationary state properties. QMC is among the most accurate and powerful approaches to the study of electronic structure, but its application is often hindered by a steep learning curve; hence it is rarely addressed in undergraduate and post-graduate classes. This tutorial is a step toward filling this gap. We offer an introduction to the diffusion Monte Carlo (DMC) method, which aims to solve the imaginary time Schrödinger equation through stochastic sampling of the configuration space. Starting from the theoretical foundations, the discussion leads naturally to the formulation of a step-by-step algorithm. To illustrate how the method works in simplified scenarios, examples such as the harmonic oscillator and the hydrogen atom are provided. The discussion extends to the fixed-node approximation, a crucial approach for addressing the fermionic sign problem in multi-electron systems. In particular, we examine the influence of trial wave function nodal surfaces on the accuracy of DMC energy by evaluating results from a non-interacting two-fermion system. Extending the method to excited states is feasible in principle, but some additional considerations are needed, supported by practical insights. By addressing the fundamental concepts from a hands-on perspective, we hope this tutorial will serve as a valuable guide for researchers and students approaching DMC for the first time.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0232424>

I. INTRODUCTION

The development of modern physics, begun during the first decades of the 20th century with the introduction of quantum mechanics, provides a theoretical framework to describe the properties of materials in terms of their elementary constituents.¹ What these *átομα* are depends on the scale of the energy involved. For most applications relevant to materials modeling, they are the nuclei and the electrons,² with the latter being responsible for most material properties, including their shape and appearance, strength, electrical and thermal conductivity, response to external perturbations, chemical reactions, etc.³⁻⁵

These properties are described by the Hamiltonian of the system, $\hat{H} = \hat{T} + \hat{V}$, where \hat{T} and \hat{V} are the kinetic and the potential energy operators, respectively. Of particular importance is the full set of eigenstates $\{|\psi_n\rangle\}$ and eigenvalues $\{E_n\}$ of the Hamiltonian, which in the non-relativistic limit are the solutions of the time independent Schrödinger equation

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle. \quad (1)$$

Among all the eigenstates, the one corresponding to the lowest eigenvalue, the so-called *ground state*, often plays a very special role. This is because although it is the state describing the system at zero

absolute temperature, it usually remains a good descriptor of the properties of a material also at moderately high temperature (such as ambient).^{6,82–84} However, apart from some special cases (e.g., the hydrogen atom), in general, Eq. (1) is impossible to solve analytically, and this remains so far one of the biggest challenges of chemistry and materials science.

The purpose of this tutorial is to introduce a practical method, i.e., the diffusion Monte Carlo (DMC),^{7–10} that provides a route to obtain the ground state $|\psi_0\rangle$, albeit in a stochastic manner. It holds a prominent position among electronic structure methods due to its extreme accuracy, as well as an important historical role. Ceperley and Alder's pioneering Quantum Monte Carlo (QMC) calculations of the electron gas¹¹ serve as the foundation for common approximations to exchange-correlation functionals in Density Functional Theory (DFT),^{12–14} thereby playing a crucial role in enabling it. Applications of the DMC method are many and varied, and it is challenging to report a chronological and exhaustive list of them here. In no particular order, it is worth mentioning calculations on diatomic molecules,¹⁵ graphene,¹⁶ solid hydrogen¹⁷ and Compton scattering profiles.¹⁸ Furthermore, results from QMC simulations are increasingly being used to generate datasets for training machine learning models of interatomic potentials. Significant examples come from Kapil *et al.*¹⁹ and Niu *et al.*²⁰ where the phase diagrams of monolayer nanoconfined water and high-pressure hydrogen, respectively, were predicted successfully using machine learning potentials with QMC as the reference method. This approach is particularly valid for the computation of phase diagrams or similar complex problems, as little inaccuracies in the evaluation of the interaction energy (say, less than the “chemical accuracy,” which is 1 kcal/mol) can lead to dramatically wrong predictions of stable phases.

Despite being very promising, quantum Monte Carlo techniques are still niche methods due to their complexity and the fact that they cannot be used as black-box tools, making them hardly accessible to non-experts. Hence, this work is primarily aimed at students and beginners, but we also hope to provide a fresh new perspective on the topic for experienced users. Compared to previous reviews, we showcase the techniques' advantages and difficulties via toy models and simplified systems where the effects of the method's technicalities and implementation choices can be easily analyzed. These are made even clearer by exploring our example codes, available on GitHub.²¹ They focused on simplicity (much shorter than modern production codes) and clarity (making use of a few elementary functions) at the expense of efficiency. In fact, we preferred the uniform sampling algorithm to the more efficient importance sampling one,^{11,15} as the latter requires more in-depth examination of subtle technical details, which we chose to discuss at a later stage. In addition, the codes can be easily adjusted by the user to test the DMC algorithm on any kind of single-particle one-dimensional or three-dimensional radial potential, making them a versatile teaching tool.

Systematic errors affecting the DMC, e.g., the *time step* and the *population control bias*,²² are also analyzed in this work, with an emphasis on practical examples and the different behavior between ground state and excited state calculations. Above all, the *fixed-node approximation*,^{23,24} which is essential for dealing with multi-fermion systems, leads to a *trial node error* on the energy, whose functional dependence can be straightforwardly explained thanks to the simplicity of the toy models.

II. THE SCHRÖDINGER EQUATION IN IMAGINARY TIME

Let us consider the time-dependent Schrödinger equation for a system with Hamiltonian \hat{H} ,²⁵

$$i\frac{\partial|\phi(t)\rangle}{\partial t} = (\hat{H} - E_T)|\phi(t)\rangle, \quad (2)$$

where $|\phi(t)\rangle$ is the wave function of the system at time t and E_T is an arbitrary energy offset, which obviously does not change the physical behavior of the system. Throughout this tutorial we will use atomic units, whereby the mass of the electron m_e , its charge e , and the reduced Planck's constant \hbar are all equal to one. Let us rewrite Eq. (2) in terms of the imaginary time $\tau = it$,^{24,26}

$$-\frac{\partial|\phi(\tau)\rangle}{\partial\tau} = (\hat{H} - E_T)|\phi(\tau)\rangle. \quad (3)$$

If the Hamiltonian does not depend on time, then the imaginary time evolution of the wave function is

$$|\phi(\tau)\rangle = e^{-\tau(\hat{H} - E_T)}|\phi(0)\rangle. \quad (4)$$

The meaning of the exponential operator $e^{\hat{A}}$ can be clarified by its action on the set of the eigenstates of \hat{A} , being

$$e^{\hat{A}}|\alpha_n\rangle = e^{a_n}|\alpha_n\rangle, \quad (5)$$

with $|\alpha_n\rangle$ the n th eigenvector of \hat{A} and a_n the corresponding eigenvalue. To represent the wave function, we can express it in terms of a complete basis set, such as the totality of the eigenstates of any operator, again, for example, the eigenstates $|\psi_n\rangle$ of the Hamiltonian. We can, therefore, write

$$|\phi(0)\rangle = \sum_n c_n |\psi_n\rangle, \quad (6)$$

where the coefficients c_n in the expansion are given by the projections of the wave function $|\phi(0)\rangle$ onto the corresponding eigenstate $|\psi_n\rangle$: $c_n \equiv \langle\psi_n|\phi(0)\rangle$. An equivalent way to express Eq. (6) is to note that the identity operator (i.e., the operator that leaves any element of the Hilbert state unchanged by its action on it) can be written as $\mathbb{1} = \sum_n |\psi_n\rangle\langle\psi_n|$, where the sum runs over all eigenstates $|\psi_n\rangle$. If the operator has a continuous spectrum, such as the position operator, for example, then the sum is replaced by an integral, and we would write $\mathbb{1} = \int d\mathbf{R}|\mathbf{R}\rangle\langle\mathbf{R}|$. These expansions are usually referred to as *resolutions of the identity*.

If we now insert Eq. (6) into Eq. (4), we obtain

$$|\phi(\tau)\rangle = \sum_n c_n |\psi_n\rangle e^{-\tau(E_n - E_T)}, \quad (7)$$

where we have assumed $E_0 < E_1 \leq E_2 \leq \dots$. If we set $E_T = E_0$, we see that, provided $\langle\psi_0|\phi(0)\rangle = c_0 \neq 0$, in the limit of long imaginary time τ , the wave function approaches the ground state of the Hamiltonian

$$\lim_{\tau \rightarrow \infty} |\phi(\tau)\rangle = c_0 |\psi_0\rangle. \quad (8)$$

This seems useful, but of course E_0 is precisely what we are trying to calculate. Furthermore, how do we realize this imaginary time evolution?

Let us consider the resolution of the identity written in terms of the eigenstates of the position operator, $\int d\mathbf{R}' |\mathbf{R}'\rangle \langle \mathbf{R}'|$, and introduce it in the r.h.s. of Eq. (4) between the exponential operator $e^{-\tau(\hat{H}-E_T)}$ and the wave function $|\phi(0)\rangle$. If we also project the wave function $|\phi(0)\rangle$ onto $\langle \mathbf{R}|$, which provides a representation of the wave function in terms of a specific eigenstate of the position operator, or its *real space representation*, we obtain

$$\phi(\mathbf{R}, \tau) = \int d\mathbf{R}' G(\mathbf{R}, \mathbf{R}', \tau) \phi(\mathbf{R}', 0), \quad (9)$$

where we have written $\phi(\mathbf{R}, \tau) \equiv \langle \mathbf{R} | \phi(\tau) \rangle$. The quantity $G(\mathbf{R}, \mathbf{R}', \tau) \equiv \langle \mathbf{R} | e^{-\tau(\hat{H}-E_T)} | \mathbf{R}' \rangle$ is called the *Green function*,^{27,28} and Eq. (9) provides a recipe to evolve the wave function $\phi(\mathbf{R}, \tau)$ with imaginary time. We see that it depends on the value of the wave function at a previous time *everywhere* in space—i.e., on the projection of the wave function onto *every* eigenstate of the position operator—and on the value of the Green function. If we cannot exactly solve the Schrödinger equation, we cannot exactly solve Eq. (9) either, the reason being that we are in general unable to obtain an analytic expression of $G(\mathbf{R}, \mathbf{R}', \tau)$.

There are some special cases, however, where $G(\mathbf{R}, \mathbf{R}', \tau)$ is known analytically. One such case is that of the Hamiltonian without the kinetic term. The corresponding real space representation of the imaginary time Schrödinger equation becomes identical to a rate equation, for example, with the wave function $\phi(\mathbf{R}, \tau)$ representing the density of a population of bacteria in some aqueous medium at position \mathbf{R} and time τ . The Green function for this process is

$$G(\mathbf{R}, \mathbf{R}', \tau) \equiv \langle \mathbf{R} | e^{-\tau(\hat{V}-E_T)} | \mathbf{R}' \rangle = e^{-\tau(V(\mathbf{R}')-E_T)} \delta(\mathbf{R}-\mathbf{R}'), \quad (10)$$

where we have used $\langle \mathbf{R} | \mathbf{R}' \rangle = \delta(\mathbf{R}-\mathbf{R}')$ and δ is the delta function. Applying this to Eq. (9), we obtain

$$\begin{aligned} \phi(\mathbf{R}, \tau) &= \int d\mathbf{R}' e^{-\tau(V(\mathbf{R}')-E_T)} \delta(\mathbf{R}-\mathbf{R}') \phi(\mathbf{R}', 0) \\ &= e^{-\tau(V(\mathbf{R})-E_T)} \phi(\mathbf{R}, 0), \end{aligned} \quad (11)$$

which shows that the process is local. Equation (11) also shows that even if $\phi(\mathbf{R}, 0)$ is normalized, $\phi(\mathbf{R}, \tau)$ may not be, in general. The population of bacteria will grow by replication in regions where $V(\mathbf{R}) \leq E_T$ and decrease where $V(\mathbf{R}) \geq E_T$, and unless E_T is adjusted appropriately, the overall population may not remain constant. Indeed, this population fluctuation provides a feedback mechanism to adjust E_T if one insists on normalization.

A second case for which the Green function can be obtained analytically is that for which the Hamiltonian only contains the kinetic operator, $\hat{T} = -\frac{1}{2}\nabla^2$, and so the Green function is

$$G(\mathbf{R}, \mathbf{R}', \tau) = \langle \mathbf{R} | e^{\tau \frac{1}{2} \nabla^2} | \mathbf{R}' \rangle. \quad (12)$$

The real space representation of the imaginary time Schrödinger equation in this case is identical to a diffusion equation, for example, describing the diffusion of bacteria in some aqueous medium, and again the wave function can be interpreted as the density of these diffusing bacteria at position \mathbf{R} and time τ . To obtain the analytic expression of the Green function, it is useful to insert a resolution of

the identity in terms of the eigenstates of the momentum operator $\hat{p} = -i\nabla$ so that we have $\hat{T} = \frac{1}{2}\hat{p}^2$ and,

$$\begin{aligned} G(\mathbf{R}, \mathbf{R}', \tau) &= \int d\mathbf{p} \langle \mathbf{R} | e^{\tau \frac{1}{2} \nabla^2} | \mathbf{p} \rangle \langle \mathbf{p} | \mathbf{R}' \rangle \\ &= \int d\mathbf{p} \langle \mathbf{R} | e^{-\tau \frac{1}{2} p^2} | \mathbf{p} \rangle \langle \mathbf{p} | \mathbf{R}' \rangle. \end{aligned} \quad (13)$$

The real three-dimensional space representation of the eigenstates of the momentum operator is²⁹

$$\langle \mathbf{R} | \mathbf{p} \rangle = \frac{1}{(2\pi)^{\frac{3}{2}}} e^{i\mathbf{p}\cdot\mathbf{R}}, \quad (14)$$

and, therefore,

$$G(\mathbf{R}, \mathbf{R}', \tau) = \int d\mathbf{p} \frac{1}{(2\pi)^{\frac{3}{2}}} e^{i\mathbf{p}\cdot(\mathbf{R}-\mathbf{R}')} e^{-\tau \frac{1}{2} p^2}. \quad (15)$$

To compute this integral,³⁰ first of all, we recognize that it can be written as the product of the integrals in each Cartesian direction, which are all equal. By making the change of variables $t = p_\alpha \sqrt{\tau/2}$ and $\omega_\alpha = (\alpha - \alpha')/\sqrt{\tau/2}$, with $\alpha = x, y, \text{ or } z$, we have

$$G(\mathbf{R}, \mathbf{R}', \tau) = \left(\frac{2}{\tau}\right)^{\frac{3}{2}} \chi(\omega_x) \chi(\omega_y) \chi(\omega_z), \quad (16)$$

where the χ 's are one-dimensional integrals of the type

$$\chi(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-t^2} e^{it\omega}, \quad (17)$$

which can be written as

$$\begin{aligned} \chi(\omega) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-t^2 + 2\frac{i\omega}{2}t - \frac{\omega^2}{4} + \frac{\omega^2}{4}} \\ &= \frac{1}{2\pi} e^{-\frac{\omega^2}{4}} \int_{-\infty}^{+\infty} dt e^{-(t-\frac{i\omega}{2})^2}. \end{aligned} \quad (18)$$

We now make the change of variable $x' = t - \frac{i\omega}{2}$, and the integral becomes

$$\chi(\omega) = \frac{1}{2\pi} e^{-\frac{\omega^2}{4}} \int_{-\infty - \frac{i\omega}{2}}^{+\infty - \frac{i\omega}{2}} dx' e^{-x'^2}, \quad (19)$$

which is a line integral in the complex plane. To evaluate it, let us introduce the closed loop integral (see Fig. 1),

$$I = \frac{1}{2\pi} \left\{ \int_{-c - \frac{i\omega}{2}}^{+c - \frac{i\omega}{2}} + \int_{+c - \frac{i\omega}{2}}^{+c} + \int_{+c}^{-c} + \int_{-c}^{-c - \frac{i\omega}{2}} \right\} dx' e^{-x'^2} = 0, \quad (20)$$

which is equal to zero because there are no poles (divergences) of $e^{-x'^2}$ inside the loop. This is true for any value of c .

If we now let $c \rightarrow \infty$, the two vertical segments in the loop give vanishing contributions to the integral, and we are, therefore, left with

$$\begin{aligned} \chi(\omega) &= \frac{1}{2\pi} e^{-\frac{\omega^2}{4}} \int_{-\infty - \frac{i\omega}{2}}^{+\infty - \frac{i\omega}{2}} dx' e^{-x'^2} \\ &= \frac{1}{2\pi} e^{-\frac{\omega^2}{4}} \int_{-\infty}^{+\infty} dx' e^{-x'^2} = \frac{1}{2\sqrt{\pi}} e^{-\frac{\omega^2}{4}}. \end{aligned} \quad (21)$$

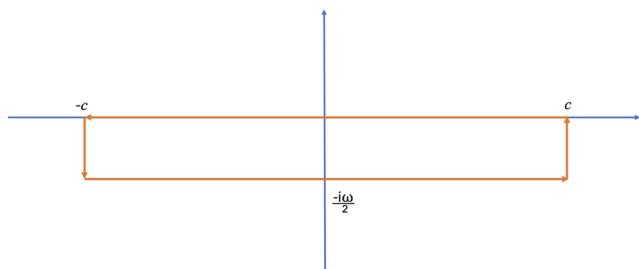


FIG. 1. The closed loop in the complex plane to evaluate the integral I.

From Eqs. (16) and (21), we get

$$G(\mathbf{R}, \mathbf{R}', \tau) = \left(\frac{1}{2\pi\tau} \right)^{\frac{3}{2}} e^{-\frac{|\mathbf{R}-\mathbf{R}'|^2}{2\tau}}. \quad (22)$$

Notice that the Green function (22) implies that the displacement $|\mathbf{R} - \mathbf{R}'|$ is normally distributed, with a standard deviation σ equal to the square root of the propagation time τ : $\sigma = \sqrt{\tau}$. Contrary to the rate case described previously, the diffusion process is not local. The number of bacteria—which we will more generally call *walkers* from now on—at one place depends on how they diffuse from everywhere else. Note that the Green function is normalized, indicating that no walkers are destroyed or created in the diffusion process. It is also symmetric; $G(\mathbf{R}, \mathbf{R}', \tau) = G(\mathbf{R}', \mathbf{R}, \tau)$, which ensures *detailed balance*³¹ and eventually allows the establishment of an equilibrium distribution. If this was not the case, the diffusion process would amplify density inhomogeneities and prevent convergence to the correct distribution.

We start to see a picture developing. When both the potential and the kinetic term are present in the Hamiltonian, they will both make their effect felt. Initially, the density of walkers will depend on time, depending on their initial distribution. The potential term will cause an exponential growth of walkers in regions of space where the potential is low, and the diffusion process will move these walkers from where the density is large to regions where the density is low. After some equilibration time, if the energy offset E_T is appropriately chosen so that overall the number of walkers is on average constant, these two processes start to balance each other, and a dynamic equilibrium is established; the density of walkers becomes time independent, and their distribution represents the ground state wave function of the Hamiltonian.

This seems to have solved the problem; however, there is still a small obstacle in the way, which has to do with the fact that the diffusion and the rate processes are not interchangeable. That is, if we let the walkers diffuse for some time τ_d first and then we let them replicate (or die) for some other time τ_r , we obtain a distribution $\phi(\mathbf{R}, \tau_d + \tau_r)$ that is in general different from what we would obtain if we swapped the order of the two processes. In the language of quantum mechanics, this is expressed by noting that the kinetic and the potential operator do not commute; it matters which one comes first.

One may think that the two operators need to be applied simultaneously, or perhaps that applying them in short succession may reduce the difference (and so the error) caused by which one is applied first. Indeed, the latter is the strategy employed in

practice to address this problem. We introduce the Trotter–Suzuki approximation^{32,33} to express the exponential of the sum of two generic operators, \hat{A} and \hat{B} , bounded from below:

$$e^{-\delta\tau(\hat{A}+\hat{B})} = e^{-\frac{1}{2}\delta\tau\hat{B}} e^{-\delta\tau\hat{A}} e^{-\frac{1}{2}\delta\tau\hat{B}} + \mathcal{O}(\delta\tau^3), \quad (23)$$

which shows that in the limit of short $\delta\tau$ it does not matter which operator is applied first. By taking $\hat{A} = -\frac{1}{2}\nabla^2$ and $\hat{B} = \hat{V} - E_T$, we obtain^{24,31}

$$\begin{aligned} G(\mathbf{R}, \mathbf{R}', \delta\tau) &= \left(\frac{1}{2\pi\delta\tau} \right)^{\frac{3}{2}} e^{-\frac{\delta\tau}{2}(V(\mathbf{R})-E_T)} e^{-\frac{|\mathbf{R}-\mathbf{R}'|^2}{2\delta\tau}} \\ &\times e^{-\frac{\delta\tau}{2}(V(\mathbf{R}')-E_T)} + \mathcal{O}(\delta\tau^3) \\ &= G_d(\mathbf{R}, \mathbf{R}', \delta\tau)G_r(\mathbf{R}, \mathbf{R}', \delta\tau) + \mathcal{O}(\delta\tau^3), \end{aligned} \quad (24)$$

where the

$$G_d(\mathbf{R}, \mathbf{R}', \delta\tau) = \left(\frac{1}{2\pi\delta\tau} \right)^{\frac{3}{2}} e^{-\frac{|\mathbf{R}-\mathbf{R}'|^2}{2\delta\tau}} \quad (25)$$

is the Green function describing the diffusion process and

$$G_r(\mathbf{R}, \mathbf{R}', \delta\tau) = e^{-\delta\tau\left(\frac{V(\mathbf{R})+V(\mathbf{R}')}{2}-E_T\right)}, \quad (26)$$

the one describing the rate process.

The evolution of the wave function for a length of time τ , as in Eq. (9), can then be split into $n = \tau/\delta\tau$ steps,

$$\begin{aligned} \phi(\mathbf{R}, n\delta\tau) &= \int d\mathbf{R}' G(\mathbf{R}, \mathbf{R}', \delta\tau)\phi(\mathbf{R}', (n-1)\delta\tau), \\ \phi(\mathbf{R}, (n-1)\delta\tau) &= \int d\mathbf{R}'' G(\mathbf{R}, \mathbf{R}'', \delta\tau)\phi(\mathbf{R}'', (n-2)\delta\tau), \\ &\dots\dots\dots \\ \phi(\mathbf{R}, 2\delta\tau) &= \int d\mathbf{R}' G(\mathbf{R}, \mathbf{R}', \delta\tau)\phi(\mathbf{R}', \delta\tau), \\ \phi(\mathbf{R}, \delta\tau) &= \int d\mathbf{R}' G(\mathbf{R}, \mathbf{R}', \delta\tau)\phi(\mathbf{R}', 0), \end{aligned} \quad (27)$$

which can be written in a more compact form as

$$\begin{aligned} \phi(\mathbf{R}, \tau) &= \int d\mathbf{R}_{n-1} \int d\mathbf{R}_{n-2} \dots \int d\mathbf{R}_1 \int d\mathbf{R}_0 G(\mathbf{R}, \mathbf{R}_{n-1}, \delta\tau) \\ &G(\mathbf{R}_{n-1}, \mathbf{R}_{n-2}, \delta\tau) \dots G(\mathbf{R}_2, \mathbf{R}_1, \delta\tau)G(\mathbf{R}_1, \mathbf{R}_0, \delta\tau)\phi(\mathbf{R}_0, 0), \end{aligned} \quad (28)$$

leading to a total error proportional to $\delta\tau^2$.

III. DIFFUSION MONTE CARLO

The steps outlined in Eqs. (27) or (28) are at the basis of the diffusion Monte Carlo (DMC) method. The term *diffusion* indicates the diffusion process, described by the Green function G_d , and *Monte Carlo* indicates the stochastic approach to evaluate the integrals³⁴ in Eqs. (27) or (28) that we will now outline.

The wave function $\phi(\mathbf{R}, \tau)$ represents the density of walkers at position \mathbf{R} and time τ . If there is only one walker and we

use its position to sample $\phi(\mathbf{R}, 0)$, if the walker is initially at \mathbf{R}_0 , we have

$$\phi(\mathbf{R}, 0) = \delta(\mathbf{R} - \mathbf{R}_0). \quad (29)$$

Inserting (29) into (27), we obtain

$$\phi(\mathbf{R}, \delta\tau) = G(\mathbf{R}, \mathbf{R}_0, \delta\tau) = G_d(\mathbf{R}, \mathbf{R}_0, \delta\tau)G_r(\mathbf{R}, \mathbf{R}_0, \delta\tau). \quad (30)$$

Since G_d is normalized, if there was no G_r , Eq. (30) would give us the probability of finding the walker at position \mathbf{R} at time $\delta\tau$, which means that this diffusion process can be simulated by drawing a random position \mathbf{R} extracted from the probability distribution $G_d(\mathbf{R}, \mathbf{R}_0, \delta\tau)$. This walker will then have a chance to be destroyed or to be replicated depending on the value of $G_r(\mathbf{R}, \mathbf{R}_0, \delta\tau)$. Note that G_r depends on our choice of E_T (high E_T means more replication, low E_T means more destruction), and so we can affect the chances of the walker dying or being replicated by adjusting E_T .

If we repeat this process with N walkers, each of them will evolve according to the same process, and after a time $\delta\tau$ they will have diffused to new positions randomly extracted from G_d in each case, and they will be eliminated or replicated according to G_r . Those walkers that find themselves in regions of space where the potential $V(\mathbf{R})$ is low will have high chances of being replicated, and conversely, those exploring regions of high $V(\mathbf{R})$ will most likely be eliminated. After this *branching* process, the overall number of walkers will decrease if E_T is too low, or vice versa, it will increase if E_T is too high, and so the fluctuating number of walkers provides a natural mechanism to adjust E_T so that, on average, their number remains constant.

After a sufficiently large number of time steps, n_{equil} , the distribution of walkers reaches a dynamic equilibrium: at every position \mathbf{R} in space, the excess replication/elimination of walkers into an infinitesimal volume $d\mathbf{R}$ centered at \mathbf{R} is balanced by diffusion out of/in to the volume $d\mathbf{R}$. When this happens, the distribution of the walkers resembles the ground state of the Hamiltonian $\psi_0(\mathbf{R})$, and E_T becomes an estimate of the ground state energy E_0 . Of course, since the distribution is only sampled by the position of the walkers, there will be a statistical fluctuation associated with it. Similarly, the estimate of E_0 based on E_T will also be affected by a statistical error. This error can be made as small as required either by increasing the number of walkers or, under the assumption of ergodicity, by continuing the simulation for a sufficiently large number of time steps n and estimating E_0 as the average $\langle E_T^j \rangle$ over the steps j after the equilibration time. This estimator of the ground state energy is called the *growth estimator*.²²

A. The algorithm

The following is a simple pseudo-code to summarize the process (see also the diagram in Fig. 2) for a three-dimensional system:

1. Initialize:

- Choose a number N_{target} of walkers, choose $\delta\tau$, and choose the total number of steps n of imaginary time evolution.

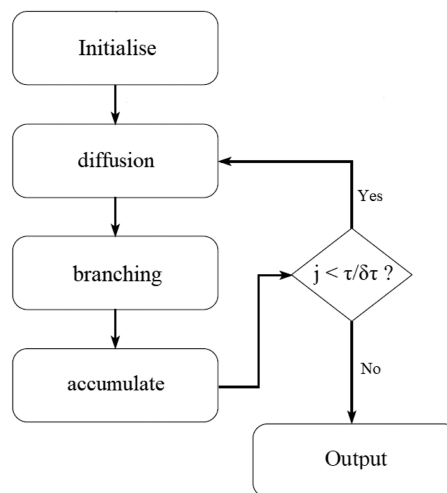


FIG. 2. Flow diagram of the DMC algorithm.

- Distribute the $N = N_{\text{target}}$ walkers at positions $(\mathbf{R}_1, \dots, \mathbf{R}_N)$ so that $\phi(\mathbf{R}, 0) = \sum_{i=1}^N \delta(\mathbf{R} - \mathbf{R}_i)$.
- Give a guess value for E_T .

2. Cycle over $j = 1, n$ time steps:

- Cycle over the N_j walkers (where N_j is the number of walkers at time step j):
 - Move (*diffusion*) each walker i from position \mathbf{R}_i to new position \mathbf{R}'_i with probability $\left(\frac{1}{2\pi\delta\tau}\right)^{\frac{3}{2}} e^{-\frac{|\mathbf{R}_i - \mathbf{R}'_i|^2}{2\delta\tau}}$.
 - Evaluate $p_i = e^{-\delta\tau\left(\frac{V(\mathbf{R}_i) + V(\mathbf{R}'_i)}{2} - E_T\right)}$.
 - Evaluate the *branching term*: $m \equiv \text{integer}(p_i + \eta)$, where η is a random number uniformly distributed between 0 and 1 and $\text{integer}(x)$ is the function returning the integer part of x .
 - If $m = 0$, eliminate the walker from the simulation; do nothing if $m = 1$; otherwise, add $m - 1$ copies.
 - If the new total number of walkers $N_{j+1} > (<) N_{\text{target}}$ reduce (increase) E_T .
- If $j > n_{\text{equil}}$, *accumulate* ground state properties, e.g., the energy.

The feedback mechanism outlined in the last step has to be implemented with care, because if it is too aggressive, it can cause wild fluctuations in the value of E_T . What is typically performed is to make the update using a slow varying function of the value of E_T averaged over a chosen number of previous steps N_{block} ,

$$E_T = E_{\text{block}} - C \log\left(\frac{N_j}{N_{\text{target}}}\right) E_{\text{block}} = \frac{1}{N_{\text{block}}} \sum_{k=1}^{N_{\text{block}}} E_T^k, \quad (31)$$

where E_T^k is the value of E_T evaluated k steps in the past and C is some appropriate energy constant that can be adjusted to modify the speed of the readjustment of the number of walkers.

To decide how many walkers are desirable to have in a simulation, it is useful to discuss one point first. The cost of moving all walkers by one time step is clearly proportional to the number of walkers N , but statistical sampling efficiency is also proportional to N ; that is, if we halve the number of walkers and double the number of simulation steps, we still obtain the same number of samples, so this would suggest that the cost of the simulation is independent of N . However, only the steps $j > n_{\text{equil}}$ are useful to compute ground state properties, with the first n_{equil} steps usually discarded. The cost of these initial n_{equil} steps is also proportional to N , and so a small N would make the simulation more efficient. On the other hand, DMC simulations are usually run on parallel computers, by distributing the walkers on different processors and, therefore, the more walkers are available, the more processors can be used. Load balance³⁵ may also become an issue if there are not enough walkers on each node, as the simulation will proceed at the speed of the slowest processor, which is the one with the largest number of walkers. For this reason, walkers are redistributed as evenly as possible after every replication/destruction, but unless they are an exact multiple of the number of processors, there will be an inevitable imbalance, with some processors having N_{proc} walkers and others having $N_{\text{proc}} + 1$. For example, in the extreme limiting case of $N_{\text{proc}} = 1$, the processors with two walkers will run twice as slow as those who have one walker. Finally, the population control mechanism of Eq. (31) may bias the simulation if the number of walkers is small (more details in Sec. II B). In conclusion, we will usually aim for a large number of walkers, large enough to avoid population control biases and load balance issues, but not unnecessarily large, so as to keep the equilibration time as small as possible fraction of the overall simulation.

Another estimate of the ground state energy E_0 can be obtained by writing it as the expectation value of the Hamiltonian over the ground state $|\psi_0\rangle$:

$$E_0 = \frac{\langle \psi_0 | \hat{H} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} = \lim_{\tau \rightarrow \infty} \frac{\langle e^{-\tau(\hat{H}-E_0)} \phi | \hat{H} | e^{-\tau(\hat{H}-E_0)} \phi \rangle}{\langle e^{-\tau(\hat{H}-E_0)} \phi | e^{-\tau(\hat{H}-E_0)} \phi \rangle} \\ = \lim_{\tau \rightarrow \infty} \frac{\langle e^{-2\tau(\hat{H}-E_0)} \phi | \hat{H} | \phi \rangle}{\langle e^{-2\tau(\hat{H}-E_0)} \phi | \phi \rangle}, \quad (32)$$

where ϕ can be any function not orthogonal to ψ_0 (i.e., $\langle \phi | \psi_0 \rangle \neq 0$). We choose ϕ to be a constant, and we replace the unknown value of E_0 in the r.h.s. of Eq. (32) with its estimator E_T . Of course ϕ cannot be exactly constant, because it would not be possible to normalize it. However, we can choose it to be constant in a sufficiently large region of space and make it go to zero smoothly at the edges, say within a volume V_{smooth} . There will be a kinetic energy contribution from the wave function in V_{smooth} , given by $\int_{V_{\text{smooth}}} -\phi(\mathbf{R}, 2\tau) \frac{1}{2} \nabla^2 \phi(\mathbf{R}) d\mathbf{R} / \int \phi(\mathbf{R}, 2\tau) d\mathbf{R}$, but by taking the edges sufficiently far, this contribution can be made as small as wanted, as

both $\phi(\mathbf{R})$ and $\phi(\mathbf{R}, 2\tau)$ must decrease to zero at large distances. A real space representation of Eq. (32) reads

$$E_0 \approx \lim_{\tau \rightarrow \infty} \frac{\int \phi(\mathbf{R}, 2\tau) V(\mathbf{R}) d\mathbf{R}}{\int \phi(\mathbf{R}, 2\tau) d\mathbf{R}} \\ \approx \frac{1}{n - n_{\text{equil}}} \sum_{j=1}^{n-n_{\text{equil}}} \frac{1}{N_j} \sum_{i=1}^{N_j} V(\mathbf{R}_{ij}) \\ = \frac{1}{n - n_{\text{equil}}} \sum_{j=1}^{n-n_{\text{equil}}} E_{\text{step}}^j \equiv \langle E_{\text{step}} \rangle, \quad (33)$$

with \mathbf{R}_{ij} the position of walker i at time step $n_{\text{equil}} + j$, and where we have replaced the ensemble average over the distribution $\phi(\mathbf{R}, 2\tau)$ in the limit of long τ with the time average over the simulation [which is performed accordingly to Eq. (28)], in which we have assumed that after n_{equil} times steps the distribution of walkers given by $\phi(\mathbf{R}, n_{\text{equil}}\delta\tau)$ has become stationary and proportional to the ground state of the Hamiltonian. The equivalence between the ensemble and the time averages obviously relies on an ergodic assumption.³⁶ The resulting estimator will be presented later (Sec. V) in a more general form, referred to as the *mixed estimator*. In Eq. (33) we have introduced the average of the potential over the instantaneous distribution of walkers, $E_{\text{step}}^j = \frac{1}{N_j} \sum_{i=1}^{N_j} V(\mathbf{R}_{ij})$. Note that the first \approx sign in Eq. (33) is due to the approximations entering the evaluation of $\phi(\mathbf{R}, 2\tau)$ according to Eq. (28), which are the short time step $\delta\tau$ approximation of Eq. (23) and the replacement of E_0 with E_T in the branching term, and the second \approx sign also includes the approximation of the evaluation of the integral as an average over the simulation, so it will be affected by a statistical error,

$$\sigma_{\langle E_{\text{step}} \rangle} = \sqrt{\frac{n_c}{n - n_{\text{equil}}} \left(\langle E_{\text{step}}^2 \rangle - (\langle E_{\text{step}} \rangle)^2 \right)}, \quad (34)$$

where n_c is the effective number of steps that we typically need to wait to obtain two statistically independent samples, i.e., the *correlation length*, and it can be obtained by standard *re-blocking* procedures.³⁷⁻³⁹ Notice that also the growth estimator $\langle E_T^j \rangle$ has a stochastic error $\sigma_{\langle E_T^j \rangle}$ that can be estimated with an equation analogous to (34).

We will now show two simple applications of the methods described earlier, the harmonic oscillator and the hydrogen atom. Codes (written in C) are available on GitHub.²¹

B. The harmonic oscillator

In this section, we will apply the techniques outlined in Sec. II, the introduction of Sec. III, and Sec. III A to the one-dimensional harmonic oscillator, for which the potential energy is

$$V(x) = \frac{1}{2} kx^2, \quad (35)$$

and we will chose $k = \omega^2 = 1$, for simplicity. The ground state wave function of this system is known analytically, and apart from a normalization constant, it is equal to

$$\psi_0(x) \propto e^{-\frac{x^2}{2}}, \quad (36)$$

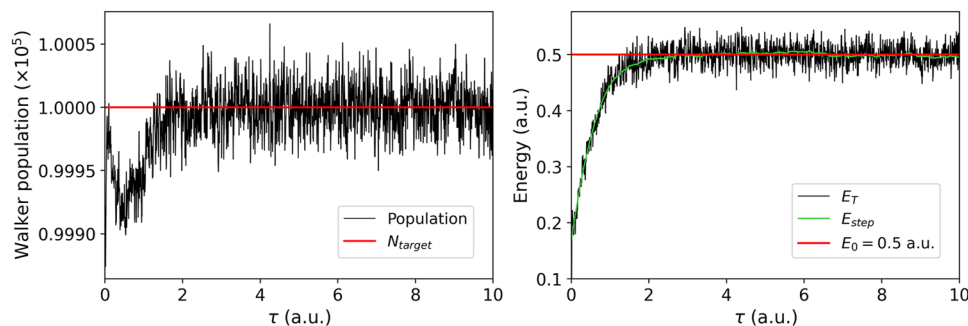


FIG. 3. Left panel: The population of walkers as a function of imaginary time for the harmonic oscillator simulation described in the text. The red line is the target population chosen to be $N_{\text{target}} = 10^5$. Right panel: The trial energy E_T (black), the step energy E_{step} (green), and the exact ground state energy (red).

and the ground state energy is $E_0 = 0.5$ a.u. As our first illustration, we perform a DMC simulation with $N_{\text{target}} = 10^5$ and $\delta\tau = 0.005$ a.u. For the energy constant of Eq. (31), we set $C = 1/\delta\tau$, and we also set $N_{\text{block}} = 100$. To initialize the distribution of walkers, we arbitrarily choose a flat function for $-1 \leq x \leq 1$ and zero outside this range (see Fig. 4), and we initially set $E_T = 0$. Both choices are clearly significantly different from the exact ground state of the Hamiltonian. In the left panel of Fig. 3, we show the number of walkers as a function of imaginary time, and in the right panel the value of E_T , together with the average value of the potential energy, E_{step} , still as a function of imaginary time. The variation of the number of walkers and of E_T in the initial part of the simulation depends on the choice for the initial distribution of walkers; however, we see that after $\tau \approx 2$ a.u., E_T stabilizes around 0.5 a.u., and also the population of walkers starts to oscillate around the target value. In Fig. 4, we show the instantaneous distribution of walkers at $\tau = 0, 0.5, 1, \text{ and } 2$ a.u. It changes with time, eventually becoming indistinguishable after $\tau \approx 2$ a.u. from the exact distribution, which matches the ground state wave function ψ_0 , as stated previously in Sec. II. It is important to emphasize that this is not the general case for real system calculations, where the widely employed *importance sampling* algorithm (Sec. V) does not directly

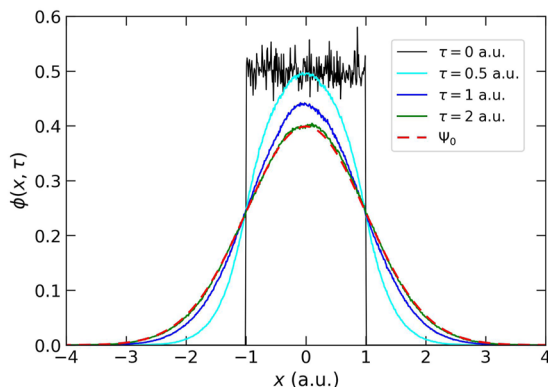


FIG. 4. The instantaneous distribution of walkers $\phi(x, \tau)$ at different imaginary times τ for the harmonic oscillator simulation described in the text, compared to the ground state wave function $\psi_0(x) = e^{-x^2/2}/\sqrt{2\pi}$ (here the normalization constant has been chosen such that $\int \psi_0(x) dx = 1$).

sample the ground state wave function, as it is biased by an arbitrary trial wave function selected by the user.

To obtain an estimate of the ground state energy, we can now average either E_T or E_{step} over the course of the simulation, after discarding an initial equilibration time. Given that in a diffusion Monte Carlo simulation the configurations between successive steps often exhibit substantial *correlation*, the statistical error on these energy estimators must be evaluated as in Eq. (34). The aforementioned re-blocking procedure involves partitioning the data into equal-sized blocks and generating new independent variables by averaging the measurements within each block. If the block averages are all statistically independent of each other, which is only true when $\# \text{steps per block} > n_c$, then the estimator of the standard deviation on the average reaches the appropriate value. Plotting $\sigma_{(E_T)}$ as a function of reblock iterations (each of which doubles the size of blocks at the previous stage) in Fig. 5 yields the best estimate for the standard error as the value at the plateau, while the block length at the plateau onset returns an estimate of n_c .

Since the Trotter–Suzuki approximation for the Green function, Eq. (23), is only exact in the limit $\delta\tau \rightarrow 0$, we need to perform a series of simulations with different values of $\delta\tau$, which we report in

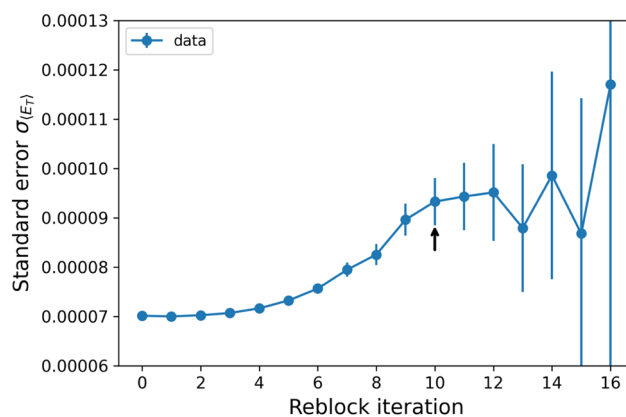


FIG. 5. $\sigma_{(E_T)}$ (computed from block averages) as a function of the reblocking iteration number, made with the Python module *pyblock*. The black arrow points out the optimal iteration number. Since the number of blocks decreases with every step, we expect a higher statistical uncertainty on the last points in the graph. DMC run performed with $\delta\tau = 0.005$ a.u., $\tau = 10^3$ a.u., and $N_{\text{target}} = 10^5$.

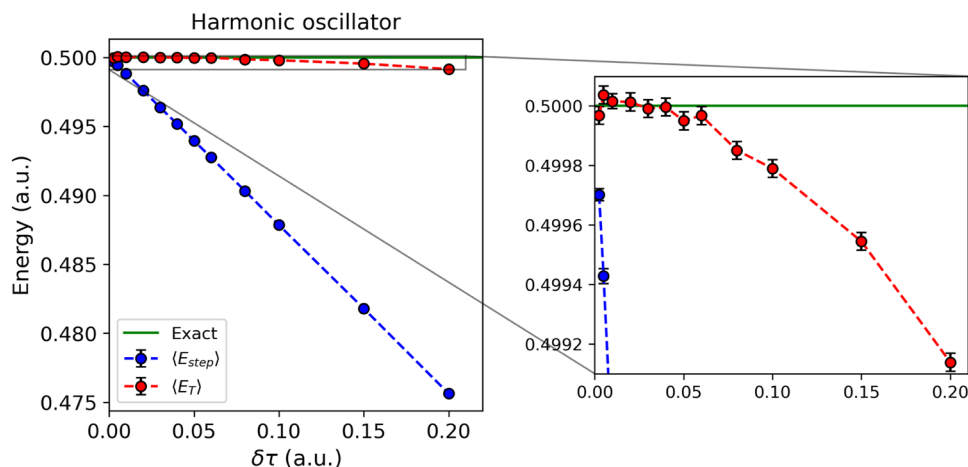


FIG. 6. Average values $\langle E_T \rangle$ (red) and $\langle E_{step} \rangle$ (blue) over the course of the simulations as a function of time step $\delta\tau$, for the harmonic oscillator simulations described in the text, compared to the exact value for the ground state energy of 0.5 a.u. (green line). The total length of each simulation is 10^9 a.u. of imaginary time.

Fig. 6, in all cases for a total imaginary time of 10^3 a.u. We see that in the limit $\delta\tau \rightarrow 0$, both energy estimators converge to the correct value 0.5, and we note that for this system the convergence is faster for the growth estimator $\langle E_T \rangle$. Notice that decreasing the time step at fixed total imaginary time implies the use of a larger number of steps and then a higher computational cost, which ideally scale as $1/\delta\tau$.

A little residual discrepancy between the growth estimator at $\delta\tau \rightarrow 0$ and the exact value of the energy may sometimes occur as a consequence of a finite discrete population. Even though each generation of walkers on average evolves by construction from the previous one, the feedback of the number of walkers into E_T^j produces a *population control bias*.^{22,40,41} This is caused by the correlation of the fluctuations in the average energy of a generation j of walkers and the corresponding DMC branching factor (containing E_T^j). These fluctuations and, therefore, the error, vanish in the ideal limit $N_{target} \rightarrow \infty$. Hence, it is tested by a sequence of simulations varying the parameter N_{target} , keeping the time step fixed. For this test, we choose $\delta\tau = 0.05$ a.u. because it returns a valid result for the energy, statistically equivalent to any lower time step's one. In addition, the overall number of samplings must be the same for all simulations, i.e., the product $(N_{target} \cdot n)$, where $n = \tau/\delta\tau$ is the number of steps, needs to be a constant, set equal to 10^9 in the case reported in Fig. 7. So the simulations with a higher population target are shorter than the others in order to maintain the same statistical accuracy on each energy estimate and consequently the same computational cost. It is evident from Fig. 7 that the growth estimator converges to the exact value for $N_{target} \gtrsim 3000$, when the population control bias becomes lower than the statistical error.

C. The hydrogen atom

It is useful to describe now a more realistic example of a Coulomb interaction, where the potential has a divergence (i.e., it corresponds to an unbounded operator), such as that of the hydrogen atom,

$$V(r) = -\frac{1}{r}, \quad r = \sqrt{x^2 + y^2 + z^2}. \quad (37)$$

This can become a problem when one or more walkers start to explore a region of space that is close enough to the divergence.⁴² Here, for any finite value of the time step, the branching term can exceed unity by a large amount, and when this happens, the Trotter–Suzuki approximation is not justified anymore. Indeed, finite time steps can take the walker accidentally very close to the nucleus, which can cause uncontrolled spikes in the population that might even be larger than the total population of walkers. These large imbalances do not typically occur very often, but when they do, they may throw the simulation off-course and make the algorithm impractical to use. One possible solution to avoid these population “explosions” is to artificially limit the value of the branching term, m ; for example, never allow it to go beyond 2 (a similar approach has been employed, for example, in Ref. 42). This modification would bias the weight, but it would do so when the Trotter–Suzuki approximations cannot be used anyway. In the end, this bias would disappear in the limit of zero time step, because in this limit the branching term would always converge to one, i.e., the Trotter–Suzuki approximation is restored, and so this approach

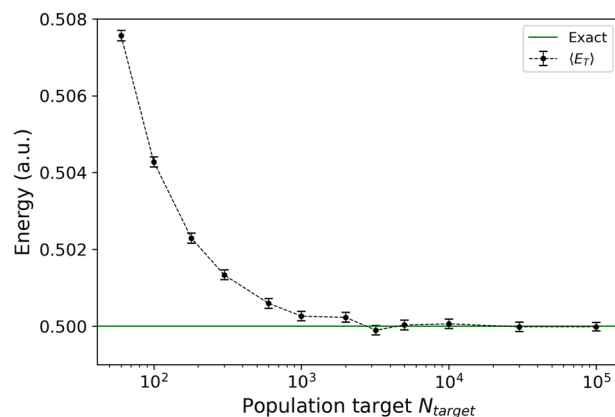


FIG. 7. Energy estimator $\langle E_T \rangle$ as a function of the target population number N_{target} compared to the exact harmonic ground state energy (green line). $\delta\tau = 0.05$ a.u., $N_{target} \cdot n = 10^9$.

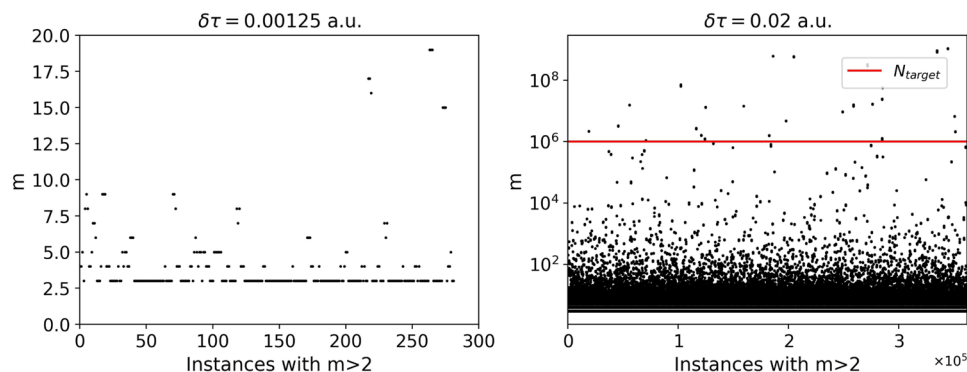


FIG. 8. Value of the branching factor m , plotted for instances in which it would exceed the value of 2. The left panel is for a simulation with $\delta\tau = 0.00125$ a.u. and 1.6×10^6 total time steps, and the right panel is for a simulation with $\delta\tau = 0.02$ a.u. and 0.5×10^6 total time steps. The red line on the right panel shows the target population.

can be controlled by studying the simulation results in this limit. To illustrate this point, we show in Fig. 8 what the value of the branching term would be if we did not limit it to 2 during the course of two simulations, one with $\delta\tau = 0.00125$ a.u. (left panel), which is run for 1.6×10^6 steps, and one with $\delta\tau = 0.02$ a.u. (right panel), which is run for 0.5×10^6 steps. In both cases we are using 10^6 walkers. The value of the branching factor is only reported when it would exceed 2. We see that for the simulation with the short time step, this happens less than 300 times, with the value of the branching factor only reaching a maximum value of 19, which would have a negligible effect on the simulation, and we could also safely avoid the constraint. However, for the simulation with the larger time step, the value 2 is exceeded 3.6×10^5 times and, more importantly, reaching maximum values of over 10^9 , which is more than 10^3 times the target population. These spikes in the population would effectively be impossible to manage, and as a result, a simulation with this value of the time step could not be run. By contrast, imposing the constraint $m \leq 2$, we can safely run simulations also with large values of $\delta\tau$, which we report in Fig. 9. Later, we will present an improved algorithm that allows for quicker convergence in the time step without the need to impose such an artificial constraint.

Then we performed a new calculation with an initial rectangular distribution of walkers in all three Cartesian coordinates and a sufficiently small time step ($\delta\tau = 0.001$ a.u.). As a result, let us finally plot in Fig. 10 the equilibrium radial distribution $F_{eq}(r) \equiv 4\pi r^2 \phi(r, \tau \rightarrow \infty)$, where $\phi(r, \tau \rightarrow \infty)$ is the projected wave function estimating the exact ground state wave function $\psi_0(r) = \exp(-r)/\sqrt{\pi}$. To plot it, we arrange a large number of spherical bins around the nucleus position $r = 0$ and then count how many walkers there are in each bin. It is clear that with this arbitrary choice, the starting distribution of walkers is very different from the ground state of the Hamiltonian, but still the algorithm converges to a suitable solution.

IV. FERMIONIC SYSTEMS AND SIGN PROBLEM

A. Many particles systems

For a many particles system, the wave function of the ground state of any Hamiltonian is symmetric under exchange of particles. This means that the techniques described in Secs. II and III always result in a bosonic solution, which is characterized by a real and positive-definite wave function (*no-node theorem*^{43,44}). If one

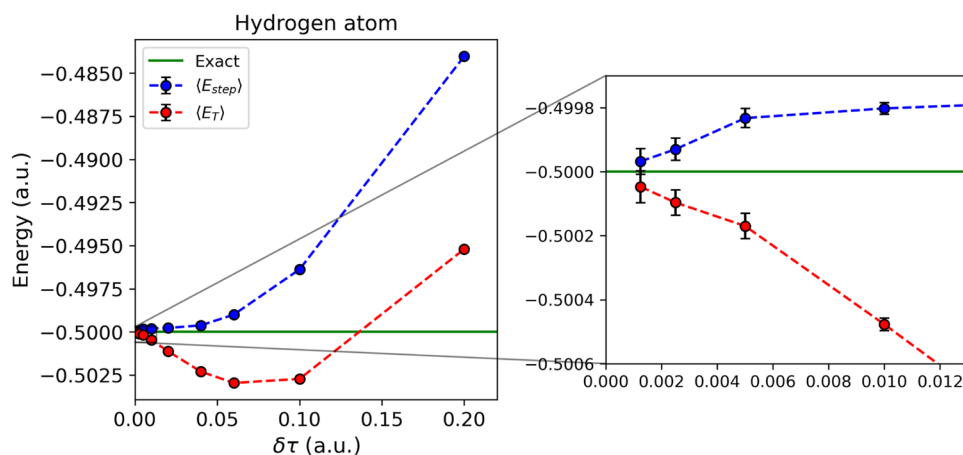


FIG. 9. Average values of the trial energy $\langle E_r \rangle$ and of $\langle E_{step} \rangle$ over the course of the simulations as a function of time step for the hydrogen atom, compared with the exact energy -0.5 a.u. (green line). Error bars are smaller than the size of the symbols. The graph on the right shows a zoom-in of the data for a smaller range of time steps.

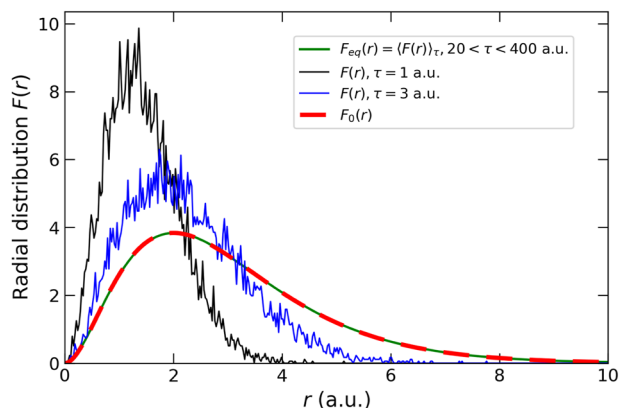


FIG. 10. Equilibrium walkers' radial distribution (green solid line) compared to the exact radial distribution $F_0(r) = 4\pi r^2 \psi_0(r)$ of the hydrogen atom (red dashed line). Since we set $\delta\tau = 0.001$ a.u. and $n_{\text{equiv}} \delta\tau = 20$ a.u., the equilibrium profile has been obtained by averaging over 380 000 configurations. The instantaneous radial distributions at $\tau = 1$ a.u. and $\tau = 3$ a.u. are also shown as representative of the equilibration phase (black and blue solid lines, respectively). $N_{\text{target}} = 10^4$.

is interested in a system of fermions, such as any electronic system, then there is no immediate access to the corresponding lowest energy solution. That poses a significant challenge, known in the literature as the *sign problem*.

One possibility could be to build an anti-symmetric (fermionic) solution as a difference of semi-positive functions, say ϕ_A and ϕ_B , performing two independent DMC simulations with different starting conditions, and sampling the difference $\phi = \phi_A - \phi_B$.³¹ The same arguments developed in Secs. II and III also apply to ϕ , namely, the fermionic eigenstates of the Hamiltonian with higher energy would decay exponentially faster than the ground state. The problem with this approach, however, is that both ϕ_A and ϕ_B converge to the bosonic ground state ψ_0 , and so the fermionic signal in their difference, ϕ , becomes smaller and smaller, and eventually it is overwhelmed by statistical noise. This makes it difficult, if not impossible, to accumulate statistics for the fermionic solution.

A common alternative approach is to introduce what is known as the fixed node approximation (FNA),²³ whereby a *nodal surface* constraint is imposed, for example, by introducing adsorbing walls. The nodal surface of the function ϕ is the hypersurface defined by the points \mathbf{R} where $\phi(\mathbf{R}) = 0$. A trial nodal surface, determined prior to the simulation, is generally extracted from an appropriate trial wave function ψ , which accurately approximates the ground state of the fermionic system. In order to implement such an approximation for any finite value of $\delta\tau$, a further step must be added to the usual algorithm in Sec. II A: a walker moving across a node from \mathbf{R} to \mathbf{R}' , such that $\psi(\mathbf{R})\psi(\mathbf{R}') < 0$, is deleted by setting its branching term m equal to 0, or alternatively the move is rejected by updating $\mathbf{R}' = \mathbf{R}$.

Of course, unless the nodal surface is identical to that of ψ_0 , the constraint increases the energy of the system because of the variational principle.¹⁵ In other words, the ground state energy estimate, which is normally a smooth function of the trial nodes' position,

has a minimum equal to E_0 and then a *second-order dependence* on the nodal surface error.^{31,45} Let us stress that the previous considerations hold because a fermionic nodal surface (i.e., the one associated with an antisymmetric trial wave function) automatically enforces the orthogonality to every other symmetric state, including the lowest-energy nodeless ground state of \hat{H} . This upper bound can be improved by improving the nodal surface with a better choice of ψ . The FNA usually only introduces a small error, as typically ψ is taken from Hartree–Fock or density functional theory calculations, which in most cases do provide good quality nodal surfaces, but it is an approximation that makes DMC non-exact.

For the sake of completeness, let us now introduce some general properties related to nodal surfaces.^{45–49} For a system with d spatial dimensions and N electrons, the nodal surface is formally a $(dN - 1)$ -dimensional hypersurface. Due to the Pauli principle, the wave function must vanish whenever any two electrons coincide. This defines $(dN - d)$ -dimensional constraints, known as *coincidence planes*, which determine completely the nodal surface only in the case $d = 1$. For higher dimensionality, they represent just a scaffolding through which the nodes must pass, and usually no general arguments can provide further information. Certainly, the nodes should exhibit the symmetries inherent in the ground-state wave function. For instance, in the case of a translationally invariant Hamiltonian, the nodes must share translational invariance. It is worth noting that this imposes only a d -dimensional limitation on the nodes, so the constraints are not overly restrictive.

In addition, for the ground-state nodal surface of Hamiltonian with a local potential, a *tiling property* has been proved.^{24,45} Let us first define a *nodal region* or *pocket*, namely a set of points in the dN -dimensional space that can be connected without crossing the nodes. Hence the walkers' evolution progresses independently in each region, where the algorithm returns the lowest-energy nodeless wave function vanishing on the boundary. In principle, a pocket α may yield a lower energy estimate E_0^α (i.e., *pocket eigenvalue*) than the others, and the gradual population correction to E_T increases the walker density in the former while emptying the latter. This happens for some excited state DMC calculations, but not for the ground state ones. In fact, the aforementioned tiling theorem states that all the nodal pockets of the ground state of some local Hamiltonian \hat{H} are equivalent by exchange symmetry, i.e., the energy estimate is the same in every nodal region. The tiling theorem holds even when the ground state is degenerate, in which case every possible real linear combination of the degenerate ground states possesses the tiling property. It can be extended to cases where additional discrete symmetries are present. For instance, if one aims for the lowest antisymmetric state exhibiting odd parity under the inversion operator $\hat{\Pi}$, then the ground state will have the tiling property with respect to the combined action of \hat{P} (particle permutation) and $\hat{\Pi}$.

Two main implications of the tiling theorem emerge in a fixed-node Monte Carlo calculation. First, the trial wave function ψ and, therefore, its trial nodal surface, need to satisfy the tiling property in order to achieve an accurate result. In general, wave functions derived from the solution to a mean field equation, such as the local density functional approximation, are satisfactory. Second, due to the equivalence between the nodal regions, the DMC energy estimate is expected to be independent of the initial walkers distribution.

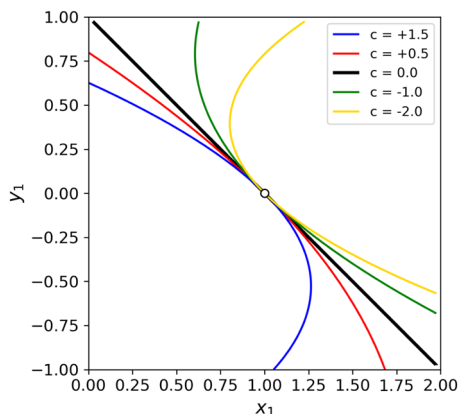


FIG. 11. 2D cross section of the trial nodal surfaces in Eq. (44) for a two-fermion system with harmonic potential. They depend on the real parameter c , and the exact one is obtained for $c = 0$ (black solid line). The fermion 2 is fixed at the position indicated by the open circle ($x_2 = 1, y_2 = 0$).

B. A two-fermion system

The toy model we are going to analyze in this section consists of two identical spinless fermions in a two-dimensional harmonic potential with angular frequencies $\omega_x = \omega_y = \omega$. Since it is a multi-particle system, its antisymmetric ground state exhibits a (three-dimensional) nodal surface, which needs to be fixed *a priori* into the algorithm. It is described by the following non-interacting Hamiltonian (in atomic units):

$$\hat{H} = \hat{H}_1 + \hat{H}_2, \quad (38)$$

where

$$\hat{H}_i = \frac{\hat{p}_{x,i}^2}{2} + \frac{\hat{p}_{y,i}^2}{2} + \frac{1}{2}\omega^2 \hat{x}_i^2 + \frac{1}{2}\omega^2 \hat{y}_i^2, \quad i = 1, 2, \quad (39)$$

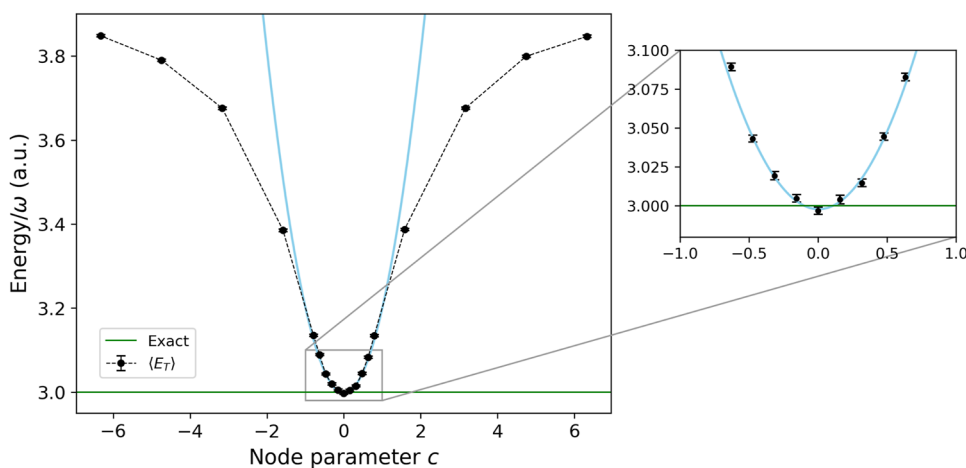


FIG. 12. DMC energy against the node parameter c for a two-fermion system. The inset shows the quadratic behavior of the energy close to the minimum at $c = 0$. $\delta\tau = 5 \times 10^{-5}$ a.u., $\tau = 315$ a.u., $N_{\text{target}} = 10^4$.

whose eigenstates can be factorized into the single-particle harmonic oscillator eigenfunctions,

$$\psi_{n_1, m_1, n_2, m_2}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{n_1}(x_1)\psi_{m_1}(y_1)\psi_{n_2}(x_2)\psi_{m_2}(y_2), \quad (40)$$

with eigenvalues $E_{n_1, m_1, n_2, m_2}/\omega = n_1 + m_1 + n_2 + m_2 + 2$, where

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{\omega}{\pi}\right)^{1/4} e^{-\frac{\omega}{2}x^2} H_n(\sqrt{\omega}x), \quad n = 0, 1, 2, \dots,$$

and the function H_n is the Hermite polynomial of degree n . Since the Hamiltonian exhibits a trivial symmetry when exchanging the x and y coordinates of the same particle, the fermionic ground state is degenerate,

$$\begin{aligned} \psi_0^F(\mathbf{r}_1, \mathbf{r}_2) = & \alpha \frac{\psi_{1,0,0,0}(\mathbf{r}_1, \mathbf{r}_2) - \psi_{1,0,0,0}(\mathbf{r}_2, \mathbf{r}_1)}{\sqrt{2}} \\ & + \beta \frac{\psi_{0,1,0,0}(\mathbf{r}_1, \mathbf{r}_2) - \psi_{0,1,0,0}(\mathbf{r}_2, \mathbf{r}_1)}{\sqrt{2}}, \end{aligned} \quad (41)$$

where the coefficients α and β are arbitrary real numbers such that $|\alpha|^2 + |\beta|^2 = 1$.

We chose $\omega = 0.4$ a.u. and $\alpha = \beta$ for simplicity. Hence, the exact nodal surface arises from the implicit equation $\psi_0^F(\mathbf{r}_1, \mathbf{r}_2) = 0$, namely,

$$x_1 - x_2 + y_1 - y_2 = 0, \quad (42)$$

which is clearly invariant under particle exchange. The *coincidence plane*

$$\begin{cases} x_1 = x_2, \\ y_1 = y_2, \end{cases} \quad (43)$$

does not provide in this case ($d = 2$) enough information to identify the entire nodal surface. In order to test the behavior of the trial node error, we can then devise an arbitrary trial nodal surface, for example,

$$x_1 - x_2 + y_1 - y_2 + c\sqrt{\omega}(y_1^2 - y_2^2) = 0, \quad (44)$$

still passing through the coincidence plane in Eq. (43) for every value of the real parameter c (Fig. 11), and perform fixed-node calculations while varying it. They have been implemented by killing walkers that cross the nodes in a simple sampling framework, starting with a uniform walker distribution. Once the time step has been fixed to a sufficiently small value such that the energy bias is negligible with respect to the statistical error, the calculations return the energy profile as in Fig. 12. Although it strictly depends on the particular form of the trial nodal surfaces, the presence of a minimum at $c = 0$ and the quadratic behavior close to it clearly stand out, returning a good estimate for the ground state energy: $\langle E_T \rangle / \omega = 2.997 \pm 0.002$ a.u., to be compared with $E_0^F / \omega = 3$ a.u. Let it be noted that the trial nodal surface presented in Eq. (44) could be rejected *a priori* because it fails to satisfy the symmetries of the Hamiltonian (except for $c = 0$), making it incompatible with a non-degenerate exact eigenstate calculation. Nevertheless, this does not impact the energy behavior, and the example still serves the purpose of demonstrating how an incorrect trial node can influence a fixed-node DMC solution. The same applies for the trial node tested later in Sec. IV B. Despite appearing straightforward within our example, this optimization procedure for the nodal surface is rarely employed in large-scale real systems because the number of parameters increases rapidly with the system size and dimensionality, and the calculations become very expensive.⁵⁰

V. EXCITED STATES

Excited states, both vibrational and electronic, are pivotal in the understanding of the physics and chemistry of atoms, molecules, and condensed matter. As discussed earlier, the fermion ground state is essentially an excited state; it represents the lowest antisymmetric state of a system, and its energy is greater than the boson one. This seems to suggest that the fixed-node approximation could be generalized to the calculation of an excited state ψ_i by employing a trial nodal surface from a proper trial wave function $\tilde{\psi}_i$. According to the variational principle,⁵¹ for the DMC solution to avoid collapsing into lower-energy states, $\tilde{\psi}_i$ should be orthogonal to all the lower eigenstates ψ_j , $j < i$.^{52,53} Therefore, orthogonality can often be achieved by imposing symmetry conditions on $\tilde{\psi}_i$.⁵⁴ Indeed, if the selected trial wave function has a defined symmetry that meets some specific mathematical conditions⁵⁵ (for example, the *gerade/ungerade* states in centrosymmetric molecules), the variational principle ensures that the computed energy is either equal to or greater than the eigenvalue of the lowest exact eigenstate with that particular symmetry.⁵⁶ However, such considerations usually do not completely specify the exact solution. For excited states that are energetically not the lowest in their symmetry, neither the variational principle nor the tiling property is guaranteed, and one may expect a strong dependence of the result on both the trial nodal surface and the initial distribution of walkers.^{53,56–58} In particular, the energy estimator from Eq. (33) for the eigenstate ψ_i will be expressed as a linear combination of energies that include the lower states,

$$\tilde{E}_i = \langle V \rangle_{\phi(\mathbf{R}, \tau \rightarrow \infty)}^{\{\tilde{\psi}_i(\mathbf{R})=0\}} = \sum_{j=0}^{\infty} c_{ij} E_j, \quad (45)$$

where the average of the potential energy has been computed by sampling the equilibrium distribution $\phi(\mathbf{R}, \tau \rightarrow \infty)$ with nodes fixed

by the implicit equation $\tilde{\psi}_i(\mathbf{R}) = 0$. The coefficients c_{ij} are related to the overlap integral $(\tilde{\psi}_i | \psi_j)$. A similar expression applies also to the growth estimator. The incorrect placement of nodes leads to a mixing of energies and, therefore, to a deviation from the ideal situation $c_{ij} = \delta_{ij}$.

In case the trial nodal surface does not possess the tiling property, the most favorable nodal pocket, namely the one with the lowest energy of all those initially occupied, becomes the only populated region in the large τ limit.⁵⁶ In other words, the DMC energy converges to $\tilde{E}_i = \min_{\alpha} \{E_i^{\alpha}\}$, where the index α runs over all the filled nodal regions. Everything introduced so far will be clarified with an example in Sec. V A.

A. First excited state of the harmonic oscillator

Let us investigate again the single-particle system with a 1D harmonic potential, but now we focus on the first excited state. It is described by the wave function,

$$\psi_1(x) = \left(\frac{\omega}{\pi}\right)^{\frac{1}{4}} \sqrt{2\omega} x \exp\left(-\frac{1}{2}\omega x^2\right), \quad (46)$$

with energy $E_1/\omega = 3/2$ a.u. We set again $\omega = 0.4$ a.u. in the following calculations. The nodal surface of such a simple system is exactly known; it is just the point $x = 0$, which divides the 1D space into two equivalent nodal regions, and it needs to be enforced with the usual fixed-node constraint. Check the evolution of the walkers distribution along the simulation in Fig. 13. The initial position of each walker is generated from a uniform distribution with a non-zero component in both nodal regions. The population rapidly converges to the exact distribution profile [except for statistical noise, Fig. 13(b)], but since the node plays the role of an infinite barrier in $x = 0$, the system reveals in the long τ limit a spontaneous symmetry breaking following a random fluctuation in the walkers distribution, which becomes more and more concentrated in just one region. Then the $\tau \rightarrow \infty$ distribution outlines the corresponding nodal pocket eigenstate [Fig. 13(d)], which is zero outside and still proportional inside the same nodal region to $|\psi_1(x)|$ apart from a normalization factor. Although only one pocket is sampled in the long time limit, the DMC energy comes to be accurate as long as the trial node corresponds to the exact one.

Adding a node constraint also has some important effects on the time step bias. After fixing the target population at a reasonable value (e.g., $N_{\text{target}} \sim 10^4$, in order to have negligible population bias), a series of simulations have been performed varying $\delta\tau$, and the resulting energies are plotted in Fig. 14, along with the analogous data for the ground state.

A slower convergent behavior is apparent for the excited state; in particular, an accurate energy estimate is achieved only with $\delta\tau = 10^{-5} - 10^{-4}$ a.u., whereas for the corresponding nodeless ground state, the same is attained with a way larger time step, namely $\delta\tau = 10^{-2} - 10^{-1}$ a.u. Hence, this results in expensive simulations, which are commonly optimized by employing the importance sampling method (see Sec. V) in every practical application. Such differences in convergence efficiency can be fully understood in light of the power law behavior of the time step bias. It can be proved^{59,60} that, for nodeless states and adopting the symmetric form of the Trotter–Suzuki approximation as in Eq. (23), the trial energy converges quadratically to its exact value as $\delta\tau$ approaches 0,

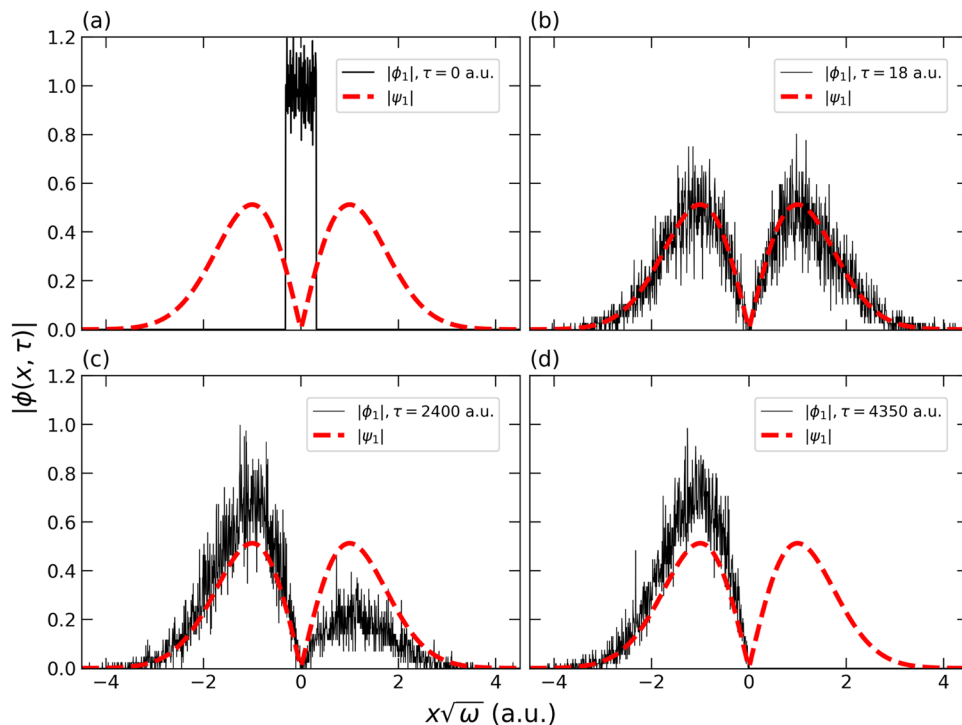


FIG. 13. Evolution of the walkers distribution $|\phi_1(x)|$ (black line) compared to the exact first excited state distribution $|\psi_1(x)|$ (red dashed line). Simulation performed with $\delta\tau = 10^{-4}$ a.u. and $N_{\text{target}} = 10^4$.

i.e., $\langle E_T \rangle(\delta\tau) - E_0 \sim \delta\tau^2$. Conversely, the act of killing walkers that cross nodes in the simple DMC algorithm leads to a $\delta\tau^{1/2}$ term in the same formula and, therefore, to a slower convergence, as supported by qualitative arguments presented in Refs. 15 and 22. This square root dependence of the growth estimator makes accurate extrapolation to $\delta\tau = 0$ difficult since $\sqrt{\delta\tau}$ has an infinite slope here.

The fitting procedure returns the best parameters as in Table I. Both extrapolated energies predict correctly the exact eigenvalues

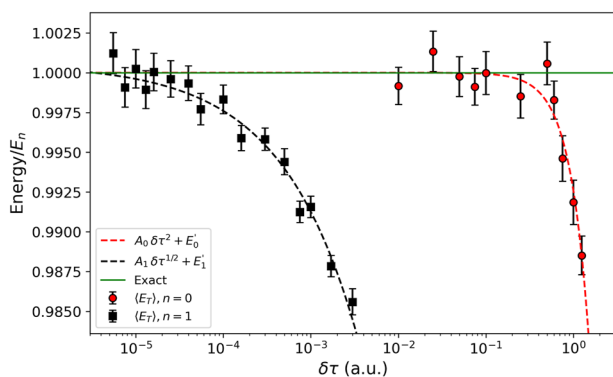


FIG. 14. The growth estimator $\langle E_T \rangle$ as a function of the time step $\delta\tau$ has been modeled according to a power law with distinct exponents for the ground state (2) and the first excited state (1/2). The energy axis has been rescaled such that both curves have the same asymptotic value at $\delta\tau = 0$. Simulations performed with $\tau = 410$ a.u. and $N_{\text{target}} = 10^4$.

with almost the same accuracy, but let us stress again that the calculation for E'_1 is much more expensive, as it is necessary to compute many low-time step points.

B. Effects of wrong nodes on excited states

We show the impact of the trial node error on the growth estimator for an excited state. The 1D harmonic oscillator Hamiltonian is invariant under inversion $\hat{\Pi}$, i.e., $V(-x) = V(x)$, so the same must be true for the nodal surface of every excited state. Clearly, the point $x = 0$ is mapped into itself by an inversion transformation, and no other single point can be. In a more formal way, only a wave function with a single node in $x = 0$ fulfills the tiling property with respect to $\hat{\Pi}$. If we fix a node at any point x_{node} other than 0, for example, $x_{\text{node}} > 0$, we expect to split the 1D domain into two non-equivalent nodal regions and get a DMC energy lower than E_1 . In particular, we start the simulation with a uniform distribution of walkers centered at $x = 0$ and large enough such that each nodal region is sampled during the initial phase. As proved for the $2s$ excited state of the hydrogen atom by Foulkes *et al.*⁵⁶ using a variational argument, the favorable pocket is the one enclosing the exact nodal point, i.e., the

TABLE I. Fitting parameters (in units of ω) for the time step bias calculation.

	Ground state	First excited state
	$E(\delta\tau) = A_0 \delta\tau^2 + E'_0$	$E(\delta\tau) = A_1 \delta\tau^{1/2} + E'_1$
A_i	-0.0037 ± 0.0004	-0.440 ± 0.015
E'_i	0.5000 ± 0.0002	1.5008 ± 0.0005

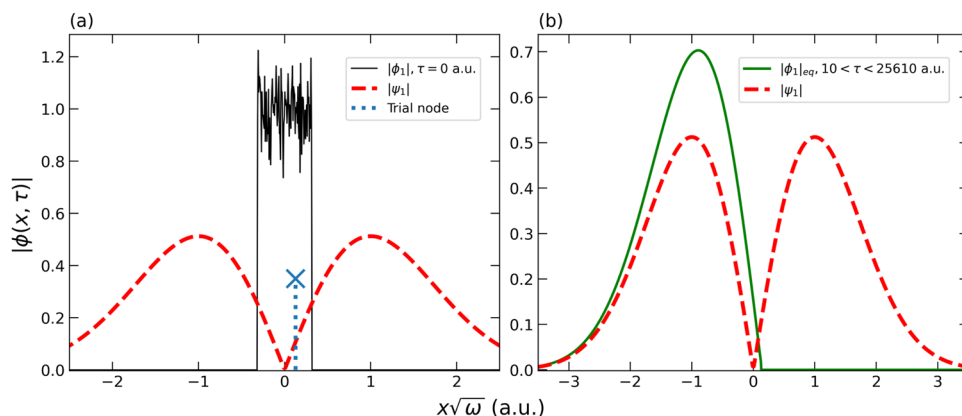


FIG. 15. Fixing the node at $x_{node} > 0$ and (a) populating both nodal regions as initial conditions lead to (b) a final (averaged) distribution of walkers entirely enclosed into the most favorable pocket $\{x < x_{node}\}$. $x_{node} \cdot \sqrt{\omega} = 0.2 \sqrt{0.4} \approx 0.13$ a.u., $\delta\tau = 0.001$ a.u., $N_{target} = 10^4$.

region $x < x_{node}$ in Fig. 15. Indeed, it is the only populated one in the high τ limit. This leads to a linear dependence of the energy on small node displacement⁶¹ and, due to the inversion symmetry of the system, to a slope discontinuity at $x_{node} = 0$ (Fig. 16). For large node displacement, instead, we observe how the growth estimator stabilizes around a minimum value equal to $E_0 = 0.5$ a.u. (Fig. 17). Although the variational principle does not hold for the first excited state, i.e., $\langle E_T \rangle(x_{node}) \leq E_1$, it is still valid for the ground state. Then a large value of x_{node} implies a large overlap between the walkers distribution and the nodeless ground state ψ_0 and an energy estimate close to (but never below) E_0 , according to Eq. (45).

In addition, the absence of the tiling property leads to a strong dependence on the initial condition. If we generate all the walkers inside the unfavorable pocket, the dynamics is constrained there due to the fixed node, and eventually the distribution cannot sample the lowest of the pocket eigenvalues. Accordingly, the populated region in Fig. 18(b) is where $x > x_{node}$ and the energy estimate is higher than E_1 . Its behavior is still linear for small node displacement, but now with a positive slope for $x_{node} > 0$ equal to $b_{>0} = 1.10 \pm 0.02$, as derived from data fitting in Fig. 19. Not surprisingly, the (negative)

slope from Fig. 16 is equal to $b_{<0} = -1.08 \pm 0.01$, sharing the same absolute value but different sign with $b_{>0}$ within the error.

For the excited state of a multi-particle system in two (or more) dimensions, the above-mentioned analysis is not so simple. Even if its nodal surface is exactly known, which is not usually the case, it is difficult to parameterize in a systematic fashion, and no intuitive representation is available. In any case, we expect the same features: the possible lack of variationality and a linear dependence of the DMC energy on the fixed-node error.

VI. DIFFUSION MONTE CARLO WITH IMPORTANCE SAMPLING

Throughout this tutorial, we have used, for the sake of clarity, the simplest DMC algorithm, which is easier to implement, yet limited in capability, and which nevertheless allows the main aspects of DMC to be grasped. However, production codes for the study of the electronic structure of molecules, solids, and surfaces employ a

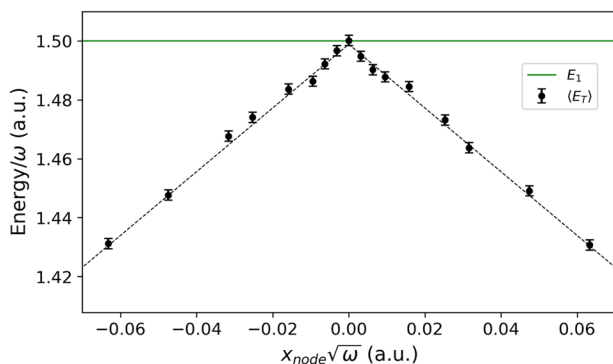


FIG. 16. Error in the DMC energy of an excited state is first order in the node displacement. Due to inversion symmetry, the data obtained with $x_{node} > 0$ are consistent within the error bars with the $x_{node} < 0$ ones. Simulations were performed with $\delta\tau = 1.6 \times 10^{-5}$ a.u., $\tau = 410$ a.u., and $N_{target} = 10^4$.

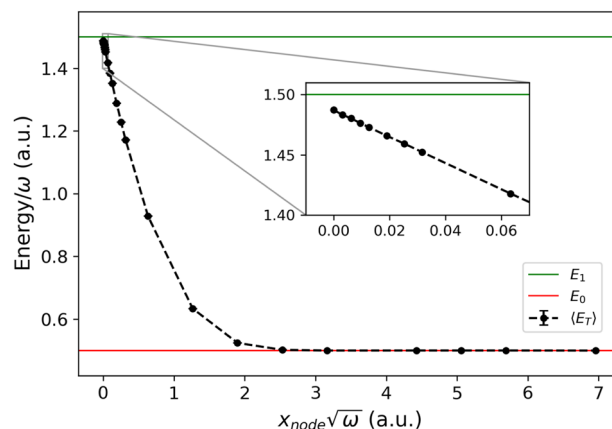


FIG. 17. Fixed-node DMC energy is linear in the zero limit (inset window) and gradually converges to the ground state energy E_0 (red line) as a larger x_{node} is enforced. The discrepancy with E_1 (green line) at $x_{node} = 0$ is due to the time step bias, which does not affect the overall qualitative behavior. Error bars are smaller than the size of the symbols. Simulations were performed with $\delta\tau = 0.001$ a.u., $\tau = 25610$ a.u., and $N_{target} = 10^4$.

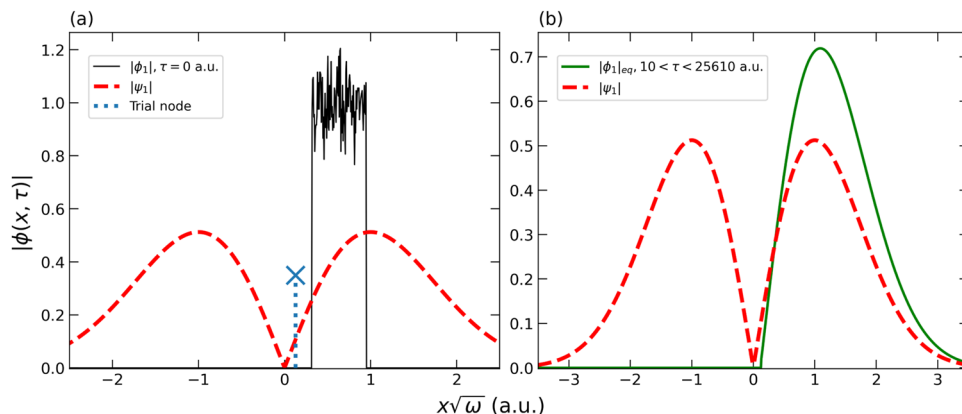


FIG. 18. Fixing the node at $x_{node} > 0$ and (a) populating just the unfavorable nodal pocket as an initial condition leads to (b) a final (averaged) distribution of walkers entirely enclosed into the $\{x > x_{node}\}$ region and a DMC energy larger than E_1 . $x_{node} \cdot \sqrt{\omega} = 0.2 \sqrt{0.4} \approx 0.13$ a.u., $\delta\tau = 0.001$ a.u., and $N_{target} = 10^4$.

more efficient but also more elaborate version of it, which ultimately becomes the best approach in all practical cases. The procedure explained in Secs. 1–V is completely general and does not require any prior knowledge of the ground state wave function of the Hamiltonian; however, if some knowledge exists, then this can be used to make the algorithm more efficient, increasing the sampling where the wave function is expected to be large.^{11,15,62–65} If knowledge of the ground state of the system is coded into a *trial* wave function ψ , this differential sampling is achieved by modifying the original algorithm as follows.

Recall the Schrödinger equation in imaginary time, Eq. (3), which we rewrite as

$$\frac{\partial \phi}{\partial \tau} = \frac{1}{2} \nabla^2 \phi + (E_T - V)\phi. \quad (47)$$

We rearrange it now in terms of the product $f(\tau) = \psi\phi(\tau)$. To do that, we need expressions for $\partial f/\partial \tau$ and the Laplacian of ϕ in terms of the gradient and the Laplacian of f . Since ψ does not depend on time, we immediately have

$$\frac{\partial f}{\partial \tau} = \psi \frac{\partial \phi}{\partial \tau}. \quad (48)$$

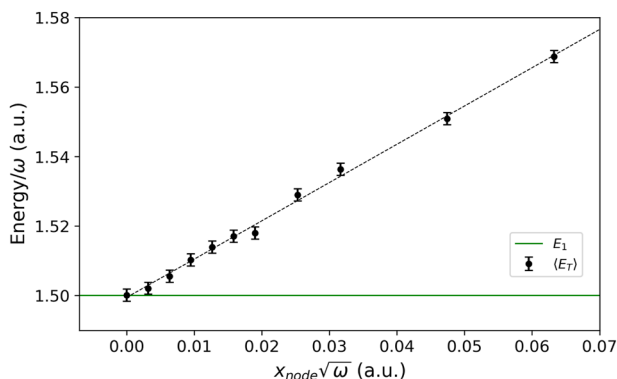


FIG. 19. If the population samples only the unfavorable nodal pocket, the error in the fixed-node DMC energy is still first order in the node displacement, but now $\langle E_T \rangle(x_{node}) \geq E_1$. $\delta\tau = 1.6 \times 10^{-5}$ a.u., $\tau = 410$ a.u., $N_{target} = 10^4$.

We then have

$$\begin{aligned} \nabla \phi &= \frac{\nabla f}{\psi} - f \frac{\nabla \psi}{\psi^2} = \frac{1}{\psi} \left[\nabla f - f \frac{\nabla \psi}{\psi} \right], \\ \nabla^2 \phi &= -\frac{\nabla \psi}{\psi^2} \cdot \left[\nabla f - f \frac{\nabla \psi}{\psi} \right] + \frac{1}{\psi} \nabla \cdot \left[\nabla f - f \frac{\nabla \psi}{\psi} \right], \\ \psi \nabla^2 \phi &= -\frac{\nabla \psi}{\psi} \cdot \nabla f + f \left(\frac{\nabla \psi}{\psi} \right)^2 + \nabla^2 f \\ &\quad - \nabla f \cdot \frac{\nabla \psi}{\psi} - f \frac{\nabla^2 \psi}{\psi} + f \left(\frac{\nabla \psi}{\psi} \right)^2 \\ &= \nabla^2 f - f \frac{\nabla^2 \psi}{\psi} + 2f \left(\frac{\nabla \psi}{\psi} \right)^2 - 2 \frac{\nabla \psi}{\psi} \cdot \nabla f. \end{aligned} \quad (49)$$

Combining Eqs. (47)–(49), we obtain

$$\frac{\partial f}{\partial \tau} = \frac{1}{2} \nabla^2 f - \frac{1}{2} f \frac{\nabla^2 \psi}{\psi} + f \left(\frac{\nabla \psi}{\psi} \right)^2 - \frac{\nabla \psi}{\psi} \cdot \nabla f + (E_T - V)f. \quad (50)$$

It is useful to write $-\frac{1}{2} f \frac{\nabla^2 \psi}{\psi} = \frac{1}{2} f \frac{\nabla^2 \psi}{\psi} - f \frac{\nabla^2 \psi}{\psi}$ and define the *local energy* $E_L(\mathbf{R}) \equiv \hat{H}\psi(\mathbf{R})/\psi(\mathbf{R})$ so that Eq. (50) becomes

$$\frac{\partial f}{\partial \tau} = \frac{1}{2} \nabla^2 f - f \frac{\nabla^2 \psi}{\psi} + f \left(\frac{\nabla \psi}{\psi} \right)^2 - \frac{\nabla \psi}{\psi} \cdot \nabla f + (E_T - E_L)f. \quad (51)$$

Finally, we define the *drift velocity* $\mathbf{v}(\mathbf{R}) \equiv \nabla \psi(\mathbf{R})/\psi(\mathbf{R})$ (or *local gradient*), and noting

$$\nabla(\mathbf{v}f) = \nabla f \cdot \frac{\nabla \psi}{\psi} + f \frac{\nabla^2 \psi}{\psi} - f \left(\frac{\nabla \psi}{\psi} \right)^2, \quad (52)$$

Equation (51) becomes

$$\frac{\partial f}{\partial \tau} = \frac{1}{2} \nabla^2 f - \nabla(\mathbf{v}f) + (E_T - E_L)f. \quad (53)$$

Equation (53) is similar to Eq. (47), but with two important differences. The first is that the potential energy V is replaced by the local energy E_L , which for a good trial wave function ψ can be much smoother. For instance, it could satisfy the *cuspl condition*^{66,67} on the

divergence of the Coulomb potential and not diverge there. In fact, a cusp in the exact wave function, i.e., a discontinuity in its first derivative, is required to cancel out the singularity of the potential with a corresponding term in the local kinetic energy. Therefore, in the limit of ψ being an eigenstate ψ_n of the Hamiltonian with eigenvalue E_n , $E_L(\mathbf{R})$ would be equal to E_n for every value of \mathbf{R} . Since $E_L(\mathbf{R})$ is replacing $V(\mathbf{R})$ in the branching factor of the Green function, $G_r(\mathbf{R}, \mathbf{R}', \delta\tau) = e^{-\delta\tau\left(\frac{E_L(\mathbf{R})+E_L(\mathbf{R}')}{2}-E_T\right)}$, this results in much weaker fluctuations and a more stable population of walkers.

The second difference comes from the extra $\nabla(\mathbf{v}f)$ term, which is not present in Eq. (47). To understand the effect of this term, it is useful to introduce an approximation and assume that for short enough time steps, \mathbf{v} is constant over an evolution of time $\delta\tau$, and so $\nabla(\mathbf{v}f) \approx \mathbf{v} \cdot \nabla f$. This allows us to write $\frac{1}{2}\nabla^2 f - \mathbf{v} \cdot \nabla f = (-\hat{p}^2/2 - i\mathbf{v} \cdot \hat{p})f$, and so the diffusion part of the Green function can be written as

$$G_d(\mathbf{R}, \mathbf{R}', \delta\tau) = \langle \mathbf{R} | e^{-\delta\tau\left(\frac{\hat{p}^2}{2} - i\delta\tau\mathbf{v} \cdot \hat{p}\right)} | \mathbf{R}' \rangle. \quad (54)$$

By inserting a resolution of the identity and working in a similar way as in Sec. I, we arrive at the following expression for a *drift-diffusion* Green function:

$$G_d(\mathbf{R}, \mathbf{R}', \tau) = \left(\frac{1}{2\pi\tau}\right)^{\frac{3}{2}} e^{-\frac{|\mathbf{R}-\mathbf{R}'-\delta\tau\mathbf{v}|^2}{2\tau}}. \quad (55)$$

We see that in addition to the diffusion process there is a *drift* caused by the term $\mathbf{v} = \nabla\psi/\psi$. This quantity is large where ψ is small and decreasing/increasing, for example, near the nodal surface of the wave function. The $\delta\tau\mathbf{v}$ term then pushes the walker away from the nodes, and it is proportional to $1/R_\perp$ close to them, where R_\perp is the distance in the direction normal to the nodal surface. In other words, the drift term, which diverges on the nodal surface, helps to keep the density of walkers low where ψ is small. Hence, the importance sampling technique provides a natural way to implement the fixed-node approximation, because in the limit of zero time step, the multiplication by ψ ensures that the steady state distribution has the same nodal surface of ψ .

The growth estimator $\langle E_T \rangle$ works also in the importance sampling formulation, whereas the one in Eq. (32) needs to be adapted by replacing the arbitrary function ϕ with the trial wave function ψ and considering that $|\psi_0\rangle = \lim_{\tau \rightarrow \infty} |e^{-\tau(\hat{H}-E_0)}\psi\rangle$,

$$\begin{aligned} E_0 &= \frac{\langle \psi_0 | \hat{H} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} = \lim_{\tau \rightarrow \infty} \frac{\langle e^{-2\tau(\hat{H}-E_0)}\psi | \hat{H} | \psi \rangle}{\langle e^{-2\tau(\hat{H}-E_0)}\psi | \psi \rangle} \\ &= \frac{\langle \psi_0 | \hat{H} | \psi \rangle}{\langle \psi_0 | \psi \rangle}. \end{aligned} \quad (56)$$

Applying the local energy definition $E_L(\mathbf{R}) = \hat{H}\psi(\mathbf{R})/\psi(\mathbf{R})$ and recalling the walkers distribution $f(\mathbf{R}, \tau) = \psi(\mathbf{R})\phi(\mathbf{R}, \tau)$, then

$$\begin{aligned} E_0 &\approx \lim_{\tau \rightarrow \infty} \frac{\int f(\mathbf{R}, 2\tau)E_L(\mathbf{R}) d\mathbf{R}}{\int f(\mathbf{R}, 2\tau) d\mathbf{R}} = \langle E_L \rangle \\ &\approx \frac{1}{n - n_{\text{equil}}} \sum_{j=1}^{n-n_{\text{equil}}} \frac{1}{N_j} \sum_{i=1}^{N_j} E_L(\mathbf{R}_{ij}), \end{aligned} \quad (57)$$

where the average is computed over a simulation with n steps, each of them evolving N_j walkers at a time, and the associated stochastic error is

$$\sigma_{\langle E_L \rangle} = \sqrt{\frac{n_c}{n - n_{\text{equil}}}} \sqrt{\langle E_L^2 \rangle - \langle E_L \rangle^2}. \quad (58)$$

In the literature, this is known as the *mixed estimator*.⁶⁸ Let us stress that the walkers population no longer samples the ground state wave function, but rather the *mixed* distribution $f(\mathbf{R}, \tau \rightarrow \infty) = \psi(\mathbf{R})\psi_0(\mathbf{R})$. As a result, the mixed estimator is inherently biased by the chosen trial wave function ψ . From Eq. (56) it is straightforward to verify that the mixed estimator can compute exact values in the same way for every operator that commutes with the Hamiltonian, e.g., every function of the Hamiltonian operator, such as atomization energy, ionization potential, and electron affinity, simply obtained as a difference of energies. However, for every other operator \hat{A} (for example, the electric dipole operator), the mixed estimator does not yield the correct expectation value A_0 , i.e.,

$$A_0 \equiv \frac{\langle \psi_0 | \hat{A} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} \neq \frac{\langle \psi_0 | \hat{A} | \psi \rangle}{\langle \psi_0 | \psi \rangle}. \quad (59)$$

Therefore, both exact and approximated methods to sample from the “pure” ψ_0^2 distribution have been developed. In that way the expectation value A_0 becomes accessible from QMC calculations. For example, it can be proved that, while both the mixed estimator A_m and the *variational estimator*,

$$A_v \equiv \frac{\langle \psi | \hat{A} | \psi \rangle}{\langle \psi | \psi \rangle}, \quad (60)$$

poorly approximate A_0 , they can be combined to provide a better approximation, namely,

$$2A_m - A_v = A_0 + O(\Delta^2), \quad (61)$$

where $\Delta \equiv \psi_0 - \psi$.^{31,69}

The challenges in implementing this enhanced algorithm from scratch consist mainly of constructing and optimizing a trial wave function specific to each quantum system. In general, a good trial wave function is required to strike a balance between accuracy and ease of evaluation. A compact and widespread representation typically employs *Slater-Jastrow* (SJ) functions. A SJ trial wave function consists of a single Slater determinant multiplied by a symmetric non-negative function of the electronic coordinates, which is dubbed the Jastrow factor. The Jastrow factor accounts for the electronic correlation and satisfies the aforementioned electron–electron and electron–nuclei *cusp conditions*. The orbitals within the Slater determinant can be derived from precise DFT or Hartree–Fock (HF) calculations or even from more sophisticated approaches depending on the required level of accuracy.⁷⁰ The Jastrow factor can have several functional forms and crucially depends on the optimization of its parameters. For an N particles system, a SJ wave function can be written as

$$\psi_{\text{SJ}}(\mathbf{X}) = e^{J(\mathbf{X})} D(\mathbf{X}), \quad (62)$$

where $D(\mathbf{X})$ is the Slater determinant, $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ and $\mathbf{x}_i = \{\mathbf{r}_i, \sigma_i\}$, i.e., spatial and spin coordinates. Usually the Jastrow factor has the following form:

$$J(\mathbf{X}) = \sum_{i=1}^N \chi(\mathbf{x}_i) - \frac{1}{2} \sum_{i \neq j}^N u(\mathbf{x}_i, \mathbf{x}_j). \quad (63)$$

Here, the χ terms describe the electron–nuclei correlation, while the u terms describe the electron–electron correlation. Depending on the investigated system, those functions can be defined in several ways, and even three-body correlation terms could be included. One of the most diffused forms for the u function in the Jastrow factor is the following:

$$u(\mathbf{x}_i, \mathbf{x}_j) = \frac{A_{\sigma_i, \sigma_j}}{r_{ij}} (1 - e^{-r_{ij}/B_{\sigma_i, \sigma_j}}), \quad (64)$$

which was used in some of the earliest work on the homogeneous electron gas.¹¹ Here, $r_{ij} = |\mathbf{r}_{ij}| = |\mathbf{r}_i - \mathbf{r}_j|$, and the parameters A_{σ_i, σ_j} and B_{σ_i, σ_j} are fixed by imposing both the cusp conditions and the proper long-range behavior.

VII. FURTHER READING

In this section, we aim to supply the reader with a comprehensive list of sources to delve into the more advanced features of the QMC methods and eventually approach the recent literature.

A general overview, with a focus on the theoretical aspects of the diffusion Monte Carlo, is provided by Foulkes *et al.*'s review,²⁴ which we have referred to extensively during the production of this tutorial. The books by Hammond *et al.*³¹ and Becca and Sorella⁶¹ also introduce the fundamentals of statistics and Monte Carlo sampling (e.g., Markov chain theory and the Metropolis algorithm) as essential tools for the reader. Great attention is paid to the different types of quantum Monte Carlo methods other than the diffusion one, such as the variational method or the exact Green's function method.

The practical implementation of QMC, as well as all aspects related to the efficiency, reliability, and limitations of the algorithm, are usually presented in the papers or guides supporting recent versions of QMC production codes. Among the most popular ones, it is worth mentioning CASINO,^{71,72} TurboRVB,⁷³ QMCPACK,⁷⁴ and CHAMP.⁷⁵ As already stated, they all employ the importance sampling together with the enhanced version of the algorithm from Umrigar *et al.*²² in order to achieve faster and more accurate calculations. Since trial wave functions and their nodal surfaces play a significant role in simulations of fermionic systems, they still represent an active research topic at present, with a main focus on their optimization. We recommend the review by Austin, Zubarev, and Lester³⁵ as it offers an insightful overview of this field.

Coming to the applications, QMC has made significant progress in several areas of electronic structure theory thanks to its favorable properties, especially in those where electron correlation is a key or dominant factor. QMC accuracy is suitable particularly for systems with weak intermolecular interactions, where a subtle balance in bonding arises from long-range dynamical correlations, i.e., dispersion effects. A comprehensive review on non-covalent interactions addressed by quantum Monte Carlo methods is that of Dubecký *et al.*⁷⁶ Furthermore, QMC is increasingly employed in

conjunction with quantum chemistry methods to provide systematic benchmark quality results for molecules. A nice discussion on this topic can be found in Al-Hamdani and Tkatchenko,⁷⁷ where the DMC is regarded as a worthy competitor to the coupled cluster with single, double, and perturbative triple [CCSD(T)] excitations, often considered the “gold standard” for many chemistry applications.

VIII. CONCLUSIONS

In this tutorial, we have provided an essential explanation of the diffusion Monte Carlo method, designed to simultaneously calculate the ground state energy and wave function of any quantum system. We developed a straightforward numerical algorithm and implemented a computer program to determine the ground state of a few simple systems of educational interest. The codes, inputs, and outputs of the examples illustrated here are available on GitHub.²¹ In this tutorial we both describe the simple and importance sampling version of the DMC algorithm. For the sake of simplicity, in the illustrated examples we employed the simple algorithm, which is less efficient than the importance sampling algorithm but way easier to implement.

The typical systematic errors affecting a DMC calculation, i.e., the time step bias and the finite population bias, have been studied, and the extrapolation to $\delta\tau \rightarrow 0$, $N_{\text{target}} \rightarrow \infty$ returned accurate energy results (e.g., the statistical error on energies in Table I is below a thousandth of ω and could be further improved with longer runs). The hydrogen atom turned out to be computationally demanding due to its three-dimensionality and the presence of a singularity in the Coulomb potential, which has been taken under control by imposing a threshold on the walker replication. The problem of instability in DMC simulations is mitigated by the importance sampling algorithm but not completely removed, so a control of instability by limiting the population size is a general feature of all DMC implementations.^{22,78} In order to apply the DMC method to systems of interacting fermions, one has to treat, as mentioned earlier, the “sign problem” due to the antisymmetry property of the many-fermion wave function. Here, we have implemented the fixed-node approximation by the killing procedure, that is, walkers crossing the nodal surface are deleted. It has been shown that the time step error follows a power law, and that this is quadratic for the nodeless ground state, but only proportional to $\delta\tau^{1/2}$ within our algorithm for any higher-energy state (including fermionic ones).

We also artificially imposed trial node errors and showed the implications. By slightly moving the trial nodal surface away from its exact position, we have shown that the energy increases quadratically in ground state calculations, while it may decrease linearly and exhibit a slope discontinuity when dealing with excited states. This last situation arises from the violation of the variational principle and the tiling property, which may occur depending on the chosen trial nodal surface. Indeed, for the first excited state of the harmonic oscillator, the exact node will be fixed completely by the inversion symmetry of the Hamiltonian. In more complex problems, symmetry arguments do not define the exact topology of the nodal surface but can constrain our guess and even restore the variational principle when imposing orthogonality to every lower-energy state.

In summary, DMC is a versatile and accurate method for simulations of quantum systems, ranging from atoms,⁷⁹ molecules, and clusters⁸⁰ to condensed matter.⁸¹ Despite not directly addressing real

systems in this work, we take a step toward democratizing DMC and lowering the learning barrier by providing a beginner-friendly discussion of both the physics and the algorithmic complexities behind this powerful method.

ACKNOWLEDGMENTS

A.A., D.A., and A.Z. acknowledge support from the European Union under the Next Generation EU (Project Nos. 20222FXZ33 and P2022MC742). D.A. and A.Z. also acknowledge support from Leverhulme Grant No. RPG-2020-038. The authors acknowledge the use of the IBISCO cluster (funded by project code PIR01_00011 "IBISCO," PON 2014–2020).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Alfonso Annarelli: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Dario Alfè:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Andrea Zen:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article.

REFERENCES

- 1 D. Paul, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1930).
- 2 N. Bohr, "I. On the constitution of atoms and molecules," *London, Edinburgh Dublin Philos. Mag. J. Sci.* **26**(151), 1–25 (1913).
- 3 A. H. Wilson, "The theory of metals. —I," *Proc. R. Soc. London, Ser. A* **138**(836), 594–606 (1932).
- 4 L. Pauling, "The principles determining the structure of complex ionic crystals," *J. Am. Chem. Soc.* **51**(4), 1010–1026 (1929).
- 5 L. Néel, "Some theoretical aspects of rock-magnetism," *Adv. Phys.* **4**(14), 191–243 (1955).
- 6 See for example the ideal Fermi gas theory as discussed in most statistical mechanics books.^{82–84}
- 7 G. W. King, "Monte-Carlo method for solving diffusion problems," *Ind. amp; Eng. Chem.* **43**(11), 2475–2478 (1951).
- 8 J. B. Anderson, "A random-walk simulation of the Schrödinger equation: H_3^+ ," *J. Chem. Phys.* **63**(4), 1499–1503 (1975).
- 9 D. Ceperley and B. Alder, "Quantum Monte Carlo," *Science* **231**(4738), 555–560 (1986).
- 10 J. B. Anderson, "Quantum Monte Carlo: Atoms, molecules, clusters, liquids, and solids," in *Reviews in Computational Chemistry* (John Wiley & Sons, Ltd., 1999), Vol. 13, pp. 133–182.

¹¹D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method," *Phys. Rev. Lett.* **45**, 566–569 (1980).

¹²P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.* **136**, B864–B871 (1964).

¹³W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.* **140**, A1133–A1138 (1965).

¹⁴W. Kohn, "Nobel lecture: Electronic structure of matter—Wave functions and density functionals," *Rev. Mod. Phys.* **71**, 1253–1266 (1999).

¹⁵P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, Jr., "Fixed-node quantum Monte Carlo for molecules," *J. Chem. Phys.* **77**(11), 5593–5603 (1982).

¹⁶E. Mostaani, N. D. Drummond, and V. I. Fal'ko, "Quantum Monte Carlo calculation of the binding energy of bilayer graphene," *Phys. Rev. Lett.* **115**, 115501 (2015).

¹⁷D. M. Ceperley and B. J. Alder, "Ground state of solid hydrogen at high pressures," *Phys. Rev. B* **36**, 2092–2106 (1987).

¹⁸C. Filippi and D. M. Ceperley, "Quantum Monte Carlo calculation of Compton profiles of solid lithium," *Phys. Rev. B* **59**, 7907–7916 (1999).

¹⁹V. Kapil, C. Schran, A. Zen, J. Chen, C. J. Pickard, and A. Michaelides, "The first-principles phase diagram of monolayer nanoconfined water," *Nature* **609**(7927), 512–516 (2022).

²⁰H. Niu, Y. Yang, S. Jensen, M. Holzmann, C. Pierleoni, and D. M. Ceperley, "Stable solid molecular hydrogen above 900 K from a machine-learned potential trained with diffusion quantum Monte Carlo," *Phys. Rev. Lett.* **130**, 076102 (2023).

²¹See examples for the uniform sampling dmc codes and the jupyter notebook files to generate the figures in the article available on GitHub at <https://github.com/zenandrea/QMC-tutorial>.

²²C. J. Umrigar, M. P. Nightingale, and K. J. Runge, "A diffusion Monte Carlo algorithm with very small time-step errors," *J. Chem. Phys.* **99**(4), 2865–2890 (1993).

²³J. B. Anderson, "Quantum chemistry by random walk. H^2P , H_3^+ , D_{3h} , $^1A_1'$, H_2 , $^3\Sigma_u^+$, H_4 , $^1\Sigma_g^+$, Be^1S ," *J. Chem. Phys.* **65**(10), 4121–4127 (1976).

²⁴W. M. C. Foulkes, L. Mitás, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," *Rev. Mod. Phys.* **73**, 33–83 (2001).

²⁵L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory, Course of Theoretical Physics Vol. 3* (Elsevier Science, 1991).

²⁶I. Kosztin, B. Faber, and K. Schulten, "Introduction to the diffusion Monte Carlo method," *Am. J. Phys.* **64**(5), 633–644 (1996).

²⁷P. J. Reynolds, J. Tobochnik, and H. Gould, "Diffusion quantum Monte Carlo," *Comput. Phys.* **4**(6), 662–668 (1990).

²⁸A. Lüchow, "Quantum Monte Carlo methods," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **1**(3), 388–402 (2011).

²⁹A. Messiah, *Quantum Mechanics* (Courier Corporation, 2014).

³⁰F. W. Byron and R. W. Fuller, *Mathematics of Classical and Quantum Physics, Dover Books on Physics and Chemistry* (Dover Publications, 1992).

³¹B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry, World Scientific Lecture and Course Notes in Chemistry Vol. 1* (World Scientific, Singapore, 1994).

³²M. Suzuki, "Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems," *Commun. Math. Phys.* **51**, 183–190 (1976).

³³H. De Raedt and B. De Raedt, "Applications of the generalized Trotter formula," *Phys. Rev. A* **28**, 3575–3580 (1983).

³⁴F. James, "Monte Carlo theory and practice," *Rep. Prog. Phys.* **43**(9), 1145 (1980).

³⁵B. M. Austin, D. Y. Zubarev, and W. A. Lester, Jr., "Quantum Monte Carlo and related approaches," *Chem. Rev.* **112**(1), 263–288 (2012).

³⁶J. v. Neumann, "Physical applications of the Ergodic hypothesis," *Proc. Natl. Acad. Sci. U. S. A.* **18**(3), 263–266 (1932).

³⁷H. Flyvbjerg and H. G. Petersen, "Error estimates on averages of correlated data," *J. Chem. Phys.* **91**, 461–466 (1989).

³⁸M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, 2017).

- ³⁹R. M. Lee, G. J. Conduit, N. Nemec, P. López Ríos, and N. D. Drummond, “Strategies for improving the efficiency of quantum Monte Carlo calculations,” *Phys. Rev. E* **83**, 066706 (2011).
- ⁴⁰N. D. Drummond, Z. Radnai, J. R. Trail, M. D. Towler, and R. J. Needs, “Diffusion quantum Monte Carlo study of three-dimensional Wigner crystals,” *Phys. Rev. B* **69**, 085116 (2004).
- ⁴¹N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes, “Finite-size errors in continuum quantum Monte Carlo calculations,” *Phys. Rev. B* **78**, 125106 (2008).
- ⁴²R. M. Martin, L. Reining, and D. M. Ceperley, *Interacting Electrons: Theory and Computational Approaches* (Cambridge University Press, 2016).
- ⁴³R. P. Feynman, *Statistical Mechanics: A Set of Lectures, Advanced Books Classics* (Avalon Publishing, 1998).
- ⁴⁴C. Wu, “Unconventional Bose-Einstein condensations beyond the ‘no-node’ theorem,” *Mod. Phys. Lett. B* **23**(01), 1–24 (2009).
- ⁴⁵D. M. Ceperley, “Fermion nodes,” *J. Stat. Phys.* **63**, 1237–1267 (1991).
- ⁴⁶D. J. Klein and H. M. Pickett, “Nodal hypersurfaces and Anderson’s random-walk simulation of the Schrödinger equation,” *J. Chem. Phys.* **64**(11), 4811–4812 (1976).
- ⁴⁷H. J. Korsch, “On the nodal behaviour of eigenfunctions,” *Phys. Lett. A* **97**(3), 77–80 (1983).
- ⁴⁸M. Caffarel and P. Claverie, “Development of a pure diffusion quantum Monte Carlo method using a full generalized Feynman–Kac formula. I. Formalism,” *J. Chem. Phys.* **88**(2), 1088–1099 (1988).
- ⁴⁹W. A. Glauser, W. R. Brown, W. A. Lester, Jr., D. Bressanini, B. L. Hammond, M. L. Koszykowski, and M. L. Koszykowski, “Random-walk approach to mapping nodal regions of N-body wave functions: Ground-state Hartree–Fock wave functions for Li–C,” *J. Chem. Phys.* **97**(12), 9200–9215 (1992).
- ⁵⁰K. Nakano, S. Sorella, D. Alfè, and A. Zen, “Beyond single-reference fixed-node approximation in ab initio diffusion Monte Carlo using antisymmetrized geminal power applied to systems with hundreds of electrons,” *J. Chem. Theory Comput.* **20**(11), 4591–4604 (2024).
- ⁵¹I. N. Levine, *Quantum Chemistry, Pearson Advanced Chemistry Series* (Pearson, 2014).
- ⁵²D. Bressanini and P. J. Reynolds, “Generalized variational principle for excited states using nodes of trial functions,” *Phys. Rev. E* **84**, 046705 (2011).
- ⁵³J. Feldt and C. Filippi, “Excited-state calculations with quantum Monte Carlo,” in *Quantum Chemistry and Dynamics of Excited States* (John Wiley & Sons, Ltd., 2020), pp. 247–275, Chap. 8.
- ⁵⁴R. M. Grimes, B. L. Hammond, P. J. Reynolds, and W. A. Lester, “Quantum Monte Carlo approach to electronically excited molecules,” *J. Chem. Phys.* **85**(8), 4749–4750 (1986).
- ⁵⁵As stated in the Foulkes *et al.* work,⁵⁶ the selected trial wave function must transform according to a one-dimensional irreducible representation of that symmetry group.
- ⁵⁶W. M. C. Foulkes, R. Q. Hood, and R. J. Needs, “Symmetry constraints and variational principles in diffusion quantum Monte Carlo calculations of excited-state energies,” *Phys. Rev. B* **60**, 4558–4570 (1999).
- ⁵⁷P. J. Reynolds, R. N. Barnett, B. L. Hammond, and W. A. Lester, “Molecular physics and chemistry applications of Quantum Monte Carlo,” *J. Stat. Phys.* **43**, 1017–1026 (1986).
- ⁵⁸M. Caffarel and P. Claverie, “Development of a pure diffusion quantum Monte Carlo method using a full generalized Feynman–Kac formula. II. Applications to simple systems,” *J. Chem. Phys.* **88**(2), 1100–1109 (1988).
- ⁵⁹M. Mella, G. Morosi, and D. Bressanini, “Time step bias improvement in diffusion Monte Carlo simulations,” *Phys. Rev. E* **61**, 2050–2057 (2000).
- ⁶⁰S. A. Chin, “Quadratic diffusion Monte Carlo algorithms for solving atomic many-body problems,” *Phys. Rev. A* **42**, 6991–7005 (1990).
- ⁶¹F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, 2017).
- ⁶²J. B. Anderson, “Quantum chemistry by random walk: Higher accuracy,” *J. Chem. Phys.* **73**(8), 3897–3899 (1980).
- ⁶³F. Mentch and J. B. Anderson, “Quantum chemistry by random walk: Importance sampling for H_3^+ ,” *J. Chem. Phys.* **74**(11), 6307–6311 (1981).
- ⁶⁴D. M. Ceperley, “The stochastic solution of the many-body Schrödinger equation for fermions,” in *Recent Progress in Many-Body Theories*, edited by J. G. Zabolitzky, M. de Llano, M. Fortes, and J. W. Clark (Springer, Berlin, Heidelberg, 1981), pp. 262–269.
- ⁶⁵M. A. Suhm and R. O. Watts, “Quantum Monte Carlo studies of vibrational states in molecules and clusters,” *Phys. Rep.* **204**(4), 293–329 (1991).
- ⁶⁶T. Kato, “On the eigenfunctions of many-particle systems in quantum mechanics,” *Commun. Pure Appl. Math.* **10**(2), 151–177 (1957).
- ⁶⁷R. T. Pack and W. B. Brown, “Cusp conditions for molecular wavefunctions,” *J. Chem. Phys.* **45**(2), 556–559 (1966).
- ⁶⁸N. Cerf and O. C. Martin, “Finite population-size effects in projection Monte Carlo methods,” *Phys. Rev. E* **51**, 3679–3693 (1995).
- ⁶⁹D. M. Ceperley and M. H. Kalos, *Quantum Many-Body Problems* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1986), pp. 145–194.
- ⁷⁰F. R. Petruzielo, J. Toulouse, and C. J. Umrigar, “Approaching chemical accuracy with quantum Monte Carlo,” *J. Chem. Phys.* **136**(12), 124116 (2012).
- ⁷¹R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos, CASINO User’s Guide, Cambridge University, Version 2.13, https://casinoqmc.net/casino_manual_dir/casino_manual.pdf.
- ⁷²R. J. Needs, M. D. Towler, N. D. Drummond, P. López Ríos, and J. R. Trail, “Variational and diffusion quantum Monte Carlo calculations with the CASINO code,” *J. Chem. Phys.* **152**(15), 154106 (2020).
- ⁷³K. Nakano, C. Attaccalite, M. Barborini, L. Capriotti, M. Casula, E. Coccia, M. Dagrada, C. Genovese, Y. Luo, G. Mazzola, A. Zen, and S. Sorella, “TurboRVB: A many-body toolkit for ab initio electronic simulations by quantum Monte Carlo,” *J. Chem. Phys.* **152**(20), 204121 (2020).
- ⁷⁴J. Kim, A. D. Baczewski, T. D. Beaudet, A. Benali *et al.*, “QMCPACK: An open source *ab initio* quantum Monte Carlo package for the electronic structure of atoms, molecules and solids,” *J. Phys.: Condens. Matter* **30**(19), 195901 (2018).
- ⁷⁵R. Shinde, E. J. Landinez Borda, S. Shepard, E. Sloatman, A. Cuzzocrea, V. Azizi, P. Lopez-Tarifa, N. Renaud, C. Umrigar, S. Moroni, and C. Filippi (2024). “Cornell-Holland ab-initio materials package (CHAMP-EU),” Github. <https://github.com/filippi-claudia/champ>
- ⁷⁶M. Dubecký, L. Mitas, and P. Jurečka, “Noncovalent interactions by quantum Monte Carlo,” *Chem. Rev.* **116**(9), 5188–5215 (2016).
- ⁷⁷Y. S. Al-Hamdani and A. Tkatchenko, “Understanding non-covalent interactions in larger molecular complexes from first principles,” *J. Chem. Phys.* **150**(1), 010901 (2019).
- ⁷⁸A. Zen, Y. Luo, G. Mazzola, L. Guidoni, and S. Sorella, “*Ab initio* molecular dynamics simulation of liquid water by quantum Monte Carlo,” *J. Chem. Phys.* **142**(14), 144111 (2015).
- ⁷⁹P. Seth, P. L. Ríos, and R. J. Needs, “Quantum Monte Carlo study of the first-row atoms and ions,” *J. Chem. Phys.* **134**(8), 084105 (2011).
- ⁸⁰M. J. Gillan, F. R. Manby, M. D. Towler, and D. Alfè, “Assessing the accuracy of quantum Monte Carlo and density functional theory for energetics of small water clusters,” *J. Chem. Phys.* **136**(24), 244105 (2012).
- ⁸¹J. Kolorenc and L. Mitas, “Applications of quantum Monte Carlo methods in condensed systems,” *Rep. Prog. Phys.* **74**(2), 026502 (2011).
- ⁸²R. K. Pathria, *Statistical Mechanics, International Series of Monographs in Natural Philosophy* (Elsevier Science & Technology Books, 1972).
- ⁸³K. Huang, *Statistical Mechanics* (Wiley, 1987).
- ⁸⁴F. Schwabl, *Statistical Mechanics, Advanced Texts in Physics* (Springer, Berlin, Heidelberg, 2006).