Calculation of Intermolecular Interactions via Diffusion Monte Carlo

by

Kevin E. Gasperich

B.S. Chemical Engineering, Pennsylvania State University, 2010

Submitted to the Graduate Faculty of

the Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Kevin E. Gasperich

It was defended on

October 25, 2019

and approved by

Kenneth D. Jordan, Richard King Mellon Professor and Distinguished Professor of Computational Chemistry, Department of Chemistry

Daniel S. Lambrecht, Assistant Professor, Department of Chemistry and Physics, Florida Gulf Coast University

Sean Garrett-Roe, Associate Professor, Department of Chemistry

David Yaron, Professor, Department of Chemistry, Carnegie Mellon University

Dissertation Director: Kenneth D. Jordan, Richard King Mellon Professor and Distinguished Professor of Computational Chemistry, Department of Chemistry

Calculation of Intermolecular Interactions via Diffusion Monte Carlo

Kevin E. Gasperich, PhD

University of Pittsburgh, 2019

Quantum chemistry is a useful tool that provides insight into the properties and behavior of chemical systems. Modern software packages have made quantum chemistry methods more easily accessible, and the continued increase in available computational resources has allowed them to be applied to larger systems at higher levels of theory. Two significant problems that the field faces are the high computational complexity of high-level methods and the shift toward parallelism in high performance computing architectures. This work examines the treatment of weakly interacting molecular systems with the fixed-node diffusion Monte Carlo (DMC) method. DMC and other quantum Monte Carlo (QMC) methods offer a possible solution to both of the aforementioned problems: they can produce near-exact results with a lower scaling (with respect to problem size) than other similarly-accurate methods, and they are inherently parallel, so there is little additional cost associated with distributing the work of a single QMC calculation across a large number of processing units.

The only error in DMC that is not systematically improvable is the constraint of a fixed nodal surface of the many-particle wave function of the system being studied. There are many cases in which a single Slater determinant trial function is sufficient to obtain accurate results, but there are others in which more sophisticated multi-determinant trial functions are necessary. Furthermore, it is non-trivial to generate nodal surfaces of similar quality for isolated and interacting molecules, so cancellation of errors is not guaranteed. We examine the use of different single- and multi-determinant trial functions in DMC calculations on small chemical systems with the goal of further understanding how to construct appropriate trial functions for molecules and clusters.

Table of Contents

List of Tables

List of Figures

Preface

I'd like to thank my advisor, Ken Jordan, for his guidance and patience over the past several years. I'd like to thank Anouar Benali for giving me the opportunity to spend several months working at Argonne National Laboratory.

I'm grateful for the people who were willing to put up with what at times may have felt like an endless stream of questions, especially Mike Deible, Albert DeFusco, Eric Berquist, and Thomas Applencourt.

I'd like to thank Bob Wienand, whose infectious energy and enthusiasm made a lasting impression that helped me get to where I am today, and I'd like to thank Julie d'Itri for helping me realize that I might enjoy theoretical chemistry and for encouraging me to apply to Pitt's graduate chemistry program.

This would've been much more difficult without the support of my family and friends, especially Shiv Upadhyay, Amanda Dumi, Amy Carlson, and Micaelan Valesky.

1.0 Introduction

Electronic structure theory encompasses a variety of methods that are used in solid-state physics, materials science, and chemistry. These methods allow one to calculate properties of molecules and solids from first principles by solving the Schrödinger equation. With a few exceptions (involving small numbers of particles and/or a high degree of symmetry), this equation cannot be solved exactly; approximations must be made, and different approximations give rise to the many different methods in electronic structure theory. These approximations can be of different types, including the form of the wave function, the type and size of the basis set, the terms included in the Hamiltonian, and the methods used for evaluating certain terms.

In general, the more accurate ab initio methods are more computationally demanding. One of the most basic ab initio methods, Hartree-Fock theory, formally scales as $O(N^4)$, where N is the size of the basis set. It is computationally inexpensive, but due to its failure to describe electron correlation, it is not accurate enough to reliably predict chemical properties. In spite of its deficiencies, Hartree-Fock calculations are useful for defining an initial set of orbitals to be used in correlated methods.

Second-order Møller-Plesset perturbation theory[1] is slightly more computationally expensive, with $O(N^5)$ scaling, but it recovers the most important correlation effects missed by Hartree-Fock, and it is a valuable approach for the study of weakly-correlated systems. Coupled cluster with singles, doubles, and perturbative triples[2] is reliably accurate for systems that are not strongly correlated, but its $O(N^7)$ scaling (where N is now a general measure of system size rather than strictly the number of orbitals) limits its applicability to larger systems. Full configuration interaction, which gives the exact nonrelativistic energy in a particular basis, has a cost which scales exponentially with system size, so it cannot be used for systems with more than a few dozen correlated electrons.

Quantum Monte Carlo methods are becoming more popular as the world's largest computers increasingly rely on parallelism for performance. One of these methods, diffusion Monte Carlo (DMC),[3–5] will be described in this chapter, and results of DMC calculations

will be presented in later chapters. DMC scales roughly as $O(N^3M)$, where N is the number of electrons and M is the number of Slater determinants in the trial function. In addition to this low-scaling computational complexity, DMC also benefits from nearly ideal parallel scaling —there is negligible overhead required to distribute the computational work of a single DMC calculation across many processing units. This makes it well-suited for modern high-performance computing.

With its $O(N^3)$ scaling, density functional theory is an inexpensive alternative to wave function-based ab initio methods. Ref. [6] offers a recent review of the performance of 200 functionals applied to several classes of problems: there are several functionals which perform well (RMS errors below $1\,\mathrm{kcal\,mol^{-1}}$) across many of the test sets, but there are some sets for which no functional has an RMS error below a few $\mathrm{kcal\,mol^{-1}}$. While DFT is useful in studies of systems for which appropriate functionals have been developed and shown to work, there is (in practice) no universal functional that will reliably give accurate results for any system, and the error due to the use of an inexact functional is not systematically improvable.

1.1 The electronic Schrödinger equation

The work described in the following chapters concerns the calculation of properties (primarily energies) of electronic systems. Such systems are assumed to be well-described by the non-relativistic time-independent electronic Schrödinger equation,

$$\hat{H}\Psi(\vec{r}_1,\ldots,\vec{r}_N) = E\Psi(\vec{r}_1,\ldots,\vec{r}_N) \tag{1.1}$$

where $\{\vec{r}\}$ represents the positions of the electrons, $\hat{H}$ is a Hamiltonian, and E is the energy associated with the system described by the wave function Ψ. For the purposes of this work, we will consider the nonrelativistic electronic Hamiltonian for a system of N electrons and M ions:

$$\hat{H} = -\sum_{i=1}^{N}\frac{\hbar^2}{2m_e}{\nabla_i}^2 + \frac{e^2}{4\pi\varepsilon_0}\left(\sum_{i<j}\frac{1}{|\vec{r}_i-\vec{r}_j|} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{|\vec{r}_i-\vec{R}_A|} + \sum_{A<B}\frac{Z_AZ_B}{|\vec{R}_A-\vec{R}_B|}\right) \tag{1.2}$$

where the first term is the operator corresponding to the kinetic energy of the electrons, and the remaining three terms are the operators corresponding to electron-electron, electron-ion, and ion-ion potential energies due to Coulomb interactions. In everything that follows, Hartree atomic units will be used (i.e. the mass of an electron m_e, reduced Planck's constant $\hbar = \frac{h}{2\pi}$, the elementary charge e, and the Coulomb constant $k_e = \frac{1}{4\pi\varepsilon_0}$ will all be equal to one.

Wave function-based electronic structure theory methods rely on applying a set of constraints to (or assuming a certain form of) the many-electron wave function for a given system. The wave function is then optimized within these constraints to give an approximate solution of the electronic Schrödinger equation. Equation (1.1) describes the effect of the Hamiltonian operating on its eigenfunctions, but it is also useful to be able to evaluate an energy associated with a wave function which is not an eigenfunction of $\hat{H}$. The expectation value of the energy associated with any wave function Ψ can be found by left-multiplying the equation by Ψ^* and integrating over the electronic degrees of freedom:

$$E = \langle\Psi|\hat{H}|\Psi\rangle = \int \mathrm{d}\vec{x}_1 \int \mathrm{d}\vec{x}_2 \ldots \int \mathrm{d}\vec{x}_N \Psi^*(\vec{x}_1, \ldots, \vec{x}_N)\hat{H}\Psi(\vec{x}_1, \ldots, \vec{x}_N), \tag{1.3}$$

where Ψ is assumed to be normalized, and where we have now also included spin as well as spatial coordinates (i.e., $\vec{x}_i = (\vec{r}_i, \omega_i)$ where ω_i represents the spin coordinate of electron i).

In practice, it is impossible to find an exact solution to (1.1) for any system of more than a few particles, so approximations must be made. One of these, the Born-Oppenheimer approximation, is already implied by the omission of a term for the nuclear kinetic energy operator in Eq. (1.2). In this approximation, the electronic part of the Schrödinger equation is solved while the positions of the nuclei remain fixed. This makes the nuclear repulsion term a constant; for simplicity, it will also be omitted from the following discussion in this chapter. In atomic units, after dropping the nuclear repulsion term, the Hamiltonian simplifies to

$$\hat{H} = \sum_i \left[-\frac{1}{2}\nabla_i^2 - \sum_{iA} \frac{Z_A}{|\vec{r}_i - \vec{R}_A|} \right] + \sum_{i<j} \frac{1}{|\vec{r}_i - \vec{r}_j|} \tag{1.4}$$

$$= \sum_i \hat{h}(i) + \sum_{i<j} \hat{v}(i,j) \tag{1.5}$$

where $\hat{h}$ is a one-electron operator representing kinetic energy and electron-ion potential energy, and $\hat{v}$ is a two-electron operator representing the electron-electron repulsion energy.

1.1.1 Slater determinants

Another useful approximation is the assumption that a many-electron wave function Ψ can be represented as a product of single-electron functions, or orbitals ψ. A simple product $\Psi(\vec{x}_1, \ldots, \vec{x}_N) = \psi_i(\vec{x}_1)\psi_j(\vec{x}_2)\ldots\psi_k(\vec{x}_N)$ is not antisymmetric (i.e., it does not change sign when the coordinates (spatial and spin) of two electrons are exchanged). A Slater determinant is a linear combination of these products that *is* antisymmetric with respect to exchange:

$$\Psi(\vec{x}_1, \vec{x}_2, ..., \vec{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_i(\vec{x}_1) & \psi_j(\vec{x}_1) & \ldots & \psi_k(\vec{x}_1) \\ \psi_i(\vec{x}_2) & \psi_j(\vec{x}_2) & \ldots & \psi_k(\vec{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_i(\vec{x}_N) & \psi_j(\vec{x}_N) & \ldots & \psi_k(\vec{x}_N) \end{vmatrix} \tag{1.6}$$

$$= |ij...k\rangle \tag{1.7}$$

The orbitals in the determinant can be taken to be parametrized functions, and the variation of these parameters gives flexibility to the orbitals and the wave function. For the rest of this work, we assume that the set of orbitals is orthonormal (i.e. $\int \mathrm{d}\vec{x}\psi_i^*(\vec{x})\psi_j(\vec{x}) = \delta_{ij}$, where δ_{ij} is the Kronecker delta); this simplifies some equations without any loss of generality.

1.1.2 One-electron basis sets

In practice, the orbitals are represented in a basis set. In periodic systems, plane wave basis sets are often used. These have the advantages of periodicity and orthogonality, and they are eigenfunctions of the momentum operator (and the kinetic energy operator). An infinite number of plane waves would provide a complete description of space (i.e., an orbital expressed as a linear combination of such functions would be infinitely flexible within the unit cell), so increasing the number of functions used will generally allow results to converge (or at least be extrapolated) to the limit of this complete basis. For a central potential, spherical harmonics are well-suited as angular basis functions, and there are several ways to represent the radial behavior. For a hydrogen-like system (a potential proportional to $\frac{1}{r}$ and a single electron), Laguerre polynomials in r along with functions of the form $e^{-\zeta r}$ (where $\zeta > 0$) are eigenfunctions of the electronic Hamiltonian. Along with the spherical harmonics, these

radial functions also form an orthogonal basis that is complete when an infinite number of them are used (and as with the plane waves, they will converge to the limit of a complete basis in a well-behaved way). In systems with more electrons and more nuclei, these are no longer exact eigenfunctions, but when expansions are placed on each nucleus they still form a useful basis that can be converged in a well-behaved way. Gaussian radial functions of the form $e^{-\alpha r^2}$ (where $\alpha > 0$) are much easier to work with than Slater-type functions ($e^{-\zeta r}$). They lack certain desirable properties (e.g., nuclear cusps, correct long-range behavior), but their ease of use makes them preferable to Slater-type functions.[a] Atom-centered Gaussian basis functions, or atomic orbitals (AOs), are used in most software packages that perform ab initio calculations on finite systems.

Once a basis set has been defined, the spatial part ϕ_i of each orbital can be represented as a linear combination of basis functions $\{\chi_\mu\}$:

$$\phi_i(\vec{r}) = \sum_\mu c_{\mu i}\, \chi_\mu(\vec{r}) \tag{1.8}$$

Orbitals must also account for the spin coordinate of electrons. This can be accomplished by multiplying the spatial part of the orbital by a spin basis function. The one-electron spin basis is usually taken to be the eigenfunctions of $\hat{s}_z$ (the z-component of the spin operator), which form a complete orthonormal basis in this space:

$$\hat{s}_z\,|\alpha\rangle = \frac{1}{2}\,|\alpha\rangle \tag{1.9}$$

$$\hat{s}_z\,|\beta\rangle = -\frac{1}{2}\,|\beta\rangle \tag{1.10}$$

$$\langle\alpha|\beta\rangle = 0 \tag{1.11}$$

$$\langle\alpha|\alpha\rangle = \langle\beta|\beta\rangle = 1 \tag{1.12}$$

[a] Efficient evaluation schemes exist for integrals over Slater-type orbitals involving only one or two centers, but not for the three- or four-center integrals that are required for ab initio calculations on polyatomic molecules.

1.2 Hartree-Fock

One of the simplest ab initio electronic structure theory methods, Hartree-Fock, aims to find the single-determinant wave function with the lowest energy in a given basis set. For a Slater determinant Ψ composed of a set of orthonormal orbitals, the expression for the energy simplifies to:

$$E = \langle\Psi|\hat{H}|\Psi\rangle = \sum_i \langle i|\hat{h}|i\rangle + \frac{1}{2}\sum_{i,j}\Big(\langle ij|ij\rangle - \langle ij|ji\rangle\Big) \tag{1.13}$$

where the sums are over occupied orbitals, and

$$\langle i|\hat{h}|j\rangle = \int \mathrm{d}\vec{x}_1\, \psi_i^*(\vec{x}_1)\hat{h}(1)\psi_j(\vec{x}_1) \tag{1.14}$$

$$\langle ij|kl\rangle = \iint \mathrm{d}\vec{x}_1\, \mathrm{d}\vec{x}_2\, \psi_i^*(\vec{x}_1)\psi_j^*(\vec{x}_2)\hat{v}(1,2)\psi_k(\vec{x}_1)\psi_l(\vec{x}_2) \tag{1.15}$$

As described above, the molecular orbitals (MOs) are represented as linear combinations of basis functions (for this discussion, these will be described as AOs, but they can be any type of basis function).

The method of Lagrange multipliers can be used to minimize the single-determinant energy in the parameter space of the coefficients $c_{\mu i}$ of Eq. (1.8) while enforcing the constraint that the MOs remain normalized. This leads to an eigenvalue problem where the optimal MOs are eigenfunctions of the Fock operator $\hat{f}$:

$$\hat{f}(1)\psi_i(\vec{x}_1) = \varepsilon_i\psi_i(\vec{x}_1) \tag{1.16}$$

$$\hat{f}(1) = \hat{h}(1) + \hat{v}_{HF}(1), \tag{1.17}$$

where ε_i is the orbital energy associated with the MO ψ_i, and the argument of $\hat{f}$ refers to the fact that it is acting on electron 1 The first term on the right-hand side of Eq. (1.17) is the one-electron operator in Eq. (1.5). The second term, $\hat{v}_{HF}$, is an approximation of the two-electron operator in Eq. (1.5). It represents the potential due to the average distribution of the other $N-1$ electrons in the system, and it is formed by summing over all occupied orbitals and integrating over the coordinates of one electron.

$$\hat{v}_{HF}(1) = \sum_{\substack{k \\ k\neq i}}^{\text{occ.}} \int \mathrm{d}\vec{x}_2 \psi_k^*(\vec{x}_2)\frac{1}{r_{12}}(1-\mathcal{P}_{12})\psi_k(\vec{x}_2) \tag{1.18}$$

where the sum runs over all occupied MOs, and the $\mathcal{P}_{12}$ operator permutes the coordinates of electrons 1 and 2. The restriction on the sum in Eq. (1.18) would necessitate a different $\hat{v}_{HF}$ (and thus a different Fock operator) for each orbital that it acts on; however, the $k \neq i$ restriction can be removed with no effect on the sum. The sum over all N occupied orbitals would appear to imply that $v\hat{_{HF}}$ acting on an electron in ψ_i will describe repulsion between that electron and itself (because it is included in the sum), but this term is exactly cancelled due to presence of the exchange term $\mathcal{P}_{12}$. If we take $v\hat{_{HF}}(1)$ to act on $\psi_i(\vec{x}_1)$, then in the $k = i$ term of the sum in Eq. (1.18), the integrand will be zero:

$$(1 - \mathcal{P}_{12})\psi_i(\vec{x}_2)\psi_i(\vec{x}_1) = \psi_i(\vec{x}_2)\psi_i(\vec{x}_1) - \psi_i(\vec{x}_1)\psi_i(\vec{x}_2) = 0 \tag{1.19}$$

Through $\hat{v}_{HF}$, the Fock operator depends on the occupied MOs, so this equation is solved iteratively in a self-consistent manner: using an initial set of MOs, the Fock operator is formed; solution of Eq. (1.16) gives a new set of MOs and associated MO energies, which can be used to form a new Fock operator. This is repeated until convergence[b]. The determinant obtained in this way is the Hartree-Fock wave function, and its energy is the Hartree-Fock energy.

1.3 Electron correlation

Electron correlation is a general term describing the correlated motion of electrons. It is useful to separate this correlation into different types.

1.3.1 Exchange correlation

Exchange (or Fermi) correlation arises from the fact that electrons are Fermions, so they must be described by an antisymmetric wave function. This means that the wave function must have a node (i.e., it must be equal to zero) wherever the coordinates (spatial and spin) of two electrons coincide. This has the effect of decreasing the likelihood that two electrons

[b] Typically, this is done until the change in total energy or some measure of the electron density difference between two successive iterations falls below a small threshold

with the same spin will be found near each other in space. This behavior is solely due to the antisymmetry of the wave function, so even a simple single Slater determinant wave function will describe this type of correlation.

1.3.2 Coulomb correlation

Coulomb correlation (or just correlation) describes the correlated motion due to electrons repelling each other through Coulomb interactions. A complete description of these interactions includes instantaneous (rather than average) positions of electrons. Hartree-Fock, a mean-field theory[c], does not account for this, so it is said to be uncorrelated. The electronic correlation energy is defined as the difference between the Hartree-Fock limit[d] and the exact nonrelativistic energy of a system. For the remainder of this work, the term "correlation" will refer only to this Coulomb correlation.

1.4 Determinants as a basis

Just as a set of AOs can form a basis in which to represent a set of one-electron orbitals, a set of Slater determinants forms a basis in which many-electron wave functions can be represented. The Hartree-Fock energy only depends on the occupied orbitals, but if the basis set has more functions than electrons (which is nearly always the case), the Fock operator will have additional orbitals (virtual orbitals) which are not occupied in the Hartree-Fock determinant. If one or more of the occupied MOs in the HF determinant is replaced by these virtual orbitals, a new Slater determinant is formed. A singly-excited determinant is formed when one occupied MO is replaced by a virtual MO, a double-excited determinant is formed when two occupied MOs are replaced by virtual MOs, and so on up to N-tuple excitations. While the functional form of an individual determinant cannot describe any

[c]Here, the term "mean-field" describes the fact that electron repulsion is accounted for by averaging over the positions of all electrons to form a field which each individual electron interacts with. (Note that because this includes an exchange term, each electron in an N-electron system only "sees" the other $N-1$ electrons in this field.)

[d]The Hartree-Fock limit is the energy that would be obtained from a Hartree-Fock calculation in the limit of a complete basis set.

correlated behavior of electrons, a linear combination of determinants can describe this correlation. A simple example of this can be seen with a minimum basis set description of H_2, where the basis consists of one s-type AO on each nucleus. The MOs in this case are determined by symmetry:

$$\sigma_g(\vec{r}) = \chi_A(\vec{r}) + \chi_B(\vec{r}) \tag{1.20}$$

$$\sigma_u(\vec{r}) = \chi_A(\vec{r}) - \chi_B(\vec{r}) \tag{1.21}$$

$$\tag{1.22}$$

where $\chi_A(\chi_B)$ is the AO centered on nucleus $A(B)$, and normalization constants have been omitted for simplicity. The lowest-energy determinant is

$$\Psi_1 = |\sigma_g\overline{\sigma_g}\rangle \tag{1.23}$$

$$\Psi_1(\vec{x}_1, \vec{x}_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \sigma_g(\vec{x}_1) & \overline{\sigma_g}(\vec{x}_1) \\ \sigma_g(\vec{x}_2) & \overline{\sigma_g}(\vec{x}_2) \end{vmatrix} \tag{1.24}$$

$$= \frac{1}{\sqrt{2}} \begin{vmatrix} \sigma_g(\vec{r}_1)\alpha(\omega_1) & \sigma_g(\vec{r}_1)\beta(\omega_1) \\ \sigma_g(\vec{r}_2)\alpha(\omega_2) & \sigma_g(\vec{r}_2)\beta(\omega_2) \end{vmatrix} \tag{1.25}$$

$$= \frac{1}{\sqrt{2}}\sigma_g(\vec{r}_1)\sigma_g(\vec{r}_2)\left[\alpha(\omega_1)\beta(\omega_2) - \beta(\omega_1)\alpha(\omega_2)\right] \tag{1.26}$$

where the overbar represents multiplication by the β spin function, and lack of an overbar represents multiplication by the α spin function. After integrating over spin coordinates, the probability density $P(\vec{r}_1, \vec{r}_2)$ of finding electron 1 at $\vec{r}_1$ and electron 2 at $\vec{r}_2$ can be factored into a product form $p(\vec{r}_1)p(\vec{r}_2)$, where it is clear that the spatial distributions of the two electrons are completely independent of each other:

$$P(\vec{r}_1, \vec{r}_2) = \int \mathrm{d}\omega_1 \int \mathrm{d}\omega_2 \left|\Psi_1(\vec{x}_1, \vec{x}_2)\right|^2 \tag{1.27}$$

$$= \left|\sigma_g(\vec{r}_1)\right|^2 \left|\sigma_g(\vec{r}_2)\right|^2 \tag{1.28}$$

This can also be thought of as a linear combination of half and half ionic (with both electrons near the same nucleus) and covalent (electrons on opposite nuclei) wave functions:

$$\sigma_g(\vec{r}_1)\sigma_g(\vec{r}_2) = \left[\chi_A(\vec{r}_1)\chi_A(\vec{r}_2) + \chi_B(\vec{r}_1)\chi_B(\vec{r}_2)\right] + \left[\chi_A(\vec{r}_1)\chi_B(\vec{r}_2) + \chi_B(\vec{r}_1)\chi_A(\vec{r}_2)\right] \tag{1.29}$$

$$= \Phi_{\text{ion}} + \Phi_{\text{cov}} \tag{1.30}$$

Another determinant with the same symmetry is $\Psi_2 = |\sigma_u\overline{\sigma_u}\rangle$. This determinant has a higher energy than Ψ_1, and it also lacks any description of spatial correlation between electrons. As with Ψ_1, this also contains half ionic and half covalent character, but these terms now have opposite signs:

$$\sigma_u(\vec{r}_1)\sigma_u(\vec{r}_2) = [\chi_A(\vec{r}_1)\chi_A(\vec{r}_2) + \chi_B(\vec{r}_1)\chi_B(\vec{r}_2)] - [\chi_A(\vec{r}_1)\chi_B(\vec{r}_2) + \chi_B(\vec{r}_1)\chi_A(\vec{r}_2)] \tag{1.31}$$

$$= \Phi_{\text{ion}} - \Phi_{\text{cov}} \tag{1.32}$$

If we use these two determinants as a basis to form a new wavefunction $\Phi = c_1\Psi_1 + c_2\Psi_2$, the resulting wave function is now correlated. Because of the difference in sign between Φ_{ion} and Φ_{cov} in Ψ_2, these two terms will no longer have the same weight in Φ. If the energy of Φ is minimized by varying c_1 and c_2, the resulting wave function will have less ionic character (i.e. the electrons will be less likely to be located near the same nucleus).

1.5 Correlated methods

Correlated methods go beyond the mean-field approximation to account for the correlated motion of electrons; in general, they are more accurate and more expensive than Hartree-Fock. As described above, a wave function comprised of a linear combination of Slater determinants has the flexibility to describe correlated behavior of electrons. This can be done in any set of MOs, although we will primarily focus on the case where the MOs are obtained from a Hartree-Fock calculation. Post-Hartree-Fock methods operate in this space of determinants formed from the converged Hartree-Fock orbitals. The Slater-Condon rules describe the matrix elements of the Hamiltonian in this basis[e]:

$$\langle\Psi|\hat{H}|\Psi\rangle = \sum_i \langle i|\hat{h}|i\rangle + \frac{1}{2}\sum_{i,j}\langle ij||ij\rangle\,, \tag{1.33a}$$

$$\langle\Psi|\hat{H}|\Psi_i^a\rangle = \langle i|\hat{h}|a\rangle + \sum_j \langle ij||aj\rangle\,, \tag{1.33b}$$

$$\left\langle\Psi\middle|\hat{H}\middle|\Psi_{ij}^{ab}\right\rangle = \langle ij||ab\rangle\,, \tag{1.33c}$$

[e] These are applicable to any basis of determinants created from a set of orthonormal MOs.

where Ψ_i^a represents the determinant created by removing orbital ψ_i from Ψ and replacing it with ψ_a (ψ_i is occupied in Ψ, and ψ_a is not), and the sums are over all occupied orbitals. If all of the MOs (occupied and virtual) are orthonormal (as we have been assuming), then the set of all Slater determinants that can be formed from these MOs is also orthonormal.

1.5.1 Configuration interaction

Configuration interaction (CI) is conceptually one of the most simple post-HF methods. In its simplest form, a CI wave function is represented as a linear combination of Slater determinants constructed from a single set of orthonormal molecular orbitals. This is usually expressed as a linear expansion of excited determinants,

$$|\Psi_{\mathrm{CI}}\rangle = (1 + \hat{C})\,|\Psi_0\rangle\,, \tag{1.34}$$

$$\hat{C} = \sum_{m=1}^{N} \hat{C}_m \tag{1.35}$$

$$\hat{C}_1 = \sum_{i,a} c_i^a \hat{\tau}_i^a \tag{1.36}$$

where $\hat{\tau}_i^a$ is a single-excitation operator, and $\hat{C}_m$ is a linear combination of m-tuple excitation operators. Using these determinants as a basis, one can use the Slater-Condon rules (Eq. (1.33a)) to construct the matrix representation of the Hamiltonian, which will be sparse (matrix elements between determinants that differ by more than two orbitals will be zero). Each eigenvector of this Hamiltonian matrix corresponds to a CI wave function with an energy given by the corresponding eigenvalue. Often, one is only interested in at most a few of these states, so it is possible to use efficient algorithms which return only a few eigenvalues (rather than diagonalizing the entire matrix).

1.5.1.1 Full configuration interaction Full CI (FCI) uses as a basis all possible determinants (possibly with some symmetry restrictions) with a given number of electrons. This yields the exact nonrelativistic energy in a given basis, but it is prohibitively expensive for systems above a moderate size.

The number of determinants grows as $\binom{2M}{N}$, where N and M are the number of electrons and spatial orbitals, respectively. If one restricts S_z to be a certain value by fixing the

number of spin-up and spin-down electrons ($N_\uparrow$ and $N_\downarrow$), then this is reduced to $\binom{M}{N_\uparrow} \times \binom{M}{N_\downarrow}$, which, although smaller, still scales exponentially (for a fixed ratio of N/M, this will scale exponentially with N; in certain limiting cases it is not exactly exponential).

1.5.1.2 Truncated configuration interaction One can also truncate $\hat{C}$ at a certain order of excitation in order to decrease the size of the wave function: $\hat{C} = \hat{C}_1$ gives CI singles (CIS); $\hat{C} = \hat{C}_1 + \hat{C}_2$ gives CI singles and doubles (CISD), and likewise for triples (CISDT), quadruples (CISDTQ), etc. The additional number of determinants at excitation level m grows as $\sim \binom{N}{m} \times \binom{2M-N}{m}$, so even though these truncated CI approaches are less expensive than FCI, they are still very costly ($\sim O(M^{m+2}) - O(M^{2m+2})$).

One significant disadvantage of truncated CI is that it is not size consistent. If E^A_{CISD} and E^B_{CISD} are the energies of molecule A and molecule B evaluated with CISD, and E^{A+B}_{CISD} is the energy of the non-interacting system comprised of molecules A and B separated by a long distance, then $E^{A+B}_{\text{CISD}} \neq E^A_{\text{CISD}} + E^B_{\text{CISD}}$. This is due to the fact that certain excitations in the direct product of the two individual molecule CISD wave functions are not present in the combined system (e.g. a double excitation on A and a double excitation on B is a quadruple excitation in A+B, which is not present in CISD).

1.5.1.3 Occupation restrictions Another way to reduce the cost of CI calculations is to disallow certain excitations from the reference configuration. The most common type of this restriction is to use a frozen core (i.e., to only include configurations in which the core orbitals are all fully occupied). It is also often useful to disallow excitations into some of the higher-lying virtual orbitals. Perhaps the simplest way to do this is with a complete active space (CAS) CI: A certain set of orbitals is defined to be the active space; the occupancies of all other orbitals are fixed, and the determinant basis is defined to include all possible rearrangements of electrons within the active orbitals (while keeping the total number of electrons constant).

1.5.2 Coupled cluster

Coupled cluster is similar to CI in that it represents a correlated wave function in terms of excitations from a reference determinant, but it differs from CI in the way that the excitation operators are applied. Whereas CI uses a linear expansion over excitation operators, coupled cluster uses an exponential expansion:[f]

$$|\Psi_{\mathrm{CC}}\rangle = e^{\hat{T}} |\Psi_0\rangle , \tag{1.37}$$

$$\hat{T} = \sum_{m=1}^{N} \hat{T}_m \tag{1.38}$$

$$\hat{T}_1 = \sum_{i,a} t_i^a \hat{\tau}_i^a \tag{1.39}$$

This exponential form means that even when $\hat{T}$ is truncated, the wave function will still contain excitations at all orders:

$$e^{\hat{T}} = 1 + \hat{T} + \frac{\hat{T}^2}{2!} + \frac{\hat{T}^3}{3!} + \dots \tag{1.40}$$

$$= 1 + \hat{T}_1 + \left(\hat{T}_2 + \frac{\hat{T}_1^2}{2!}\right) + \left(\hat{T}_3 + \hat{T}_1\hat{T}_2 + \hat{T}_2\hat{T}_1 + \frac{\hat{T}_1^3}{3!}\right) + \dots \tag{1.41}$$

Unlike with truncated CI, truncated CC is size consistent. The most frequently used variant of CC is CCSD(T), in which the singles and doubles are treated fully, and a perturbative correction is applied to approximate the effect of triples. In cases where HF is a good reference, CCSD(T) gives accurate results, but it scales as $\sim O(n^7)$[g] with system size n, so it is expensive for large systems.

[f] This can also be expressed as a product of excitation operators, but the exponential form is usually used.

[g] the leading term in the complexity scales as $O(o^3v^4)$ in the number of occupied orbitals o and virtual orbitals v

1.5.3 Selected configuration interaction

The aforementioned CI methods all suffer from scaling issues due to their equivalent treatment of all determinants within a particular subset (e.g. in CISD, all singles and doubles are included in the space that is diagonalized, regardless of how strongly they contribute to the wave function). In one study of relative importance of different determinants in CI calculations on H_2O, Ne, CO, and C_2, it was found that ~99.99% of the correlation energy could be recovered using only a small fraction of the full determinant space.[7] In light of this fact, it would be useful to have a method that selects determinants to include in the wave function based on some measure of their importance.

One class of methods that has been developed for this purpose is selected CI (SCI). In general, SCI methods are iterative, and at each iteration they evaluate some or all of the determinants that are connected (by single or double excitations) to the current wave function. From these connected determinants, the most important ones are added to the wave function for the next iteration. There are different metrics by which these determinants can be evaluated and selected, and this choice is part of what distinguishes the many different SCI methods.

The basic ideas of SCI have been around for decades,[8–10] but the field has recently seen a resurgence in popularity with the development of several new SCI methods[11–16] as well as improvements to old methods.[17–20] These methods allow near-FCI level results at a fraction of the cost. They can also be used in a black-box way, without relying heavily on any "chemical intuition" from the user. This separates them from traditional restricted CI methods (e.g. CAS, RAS, ORMAS), which are most useful when the user is able to specify chemically-relevant active spaces which include the most important determinants but are not so large as to make the calculations prohibitively expensive.

1.6 Monte Carlo methods

1.6.1 Monte Carlo integration

Eq. (1.3) shows that the energy of a square integrable wave function can be evaluated by performing a multidimensional integral. In the HF and post-HF methods described above, one often chooses to represent orbitals in terms of simple functions (usually Gaussians) in order to keep the integrals easy to analytically evaluate; however, if one is willing to evaluate the integrals numerically it becomes possible to evaluate the energy corresponding to an arbitrary wave function. This allows the use of wave functions with some of the constraints that describe the exact wave function (e.g., nuclear cusps, electron-electron cusps, explicit r_{12} dependence, long range behavior).

Because the integral over all spatial coordinates of an N-electron wave function in 3-dimensional space spans 3N dimensions, standard quadrature methods (i.e., those using fixed abscissas and weights) are not efficient for systems of more than a few electrons. In general, for a method where the error converges as n^{-p} in one dimension, the error in D dimensions will converge as $n_{tot}^{-p/D}$, where n is the number of points in each dimension and $n_{tot} = n^D$ is the total number of points.

$$\langle f \rangle = \frac{1}{b-a} \int_a^b f(x) \mathrm{d}x \approx \frac{1}{n} \sum_{i=1}^{n} w_i f(x_i) \tag{1.42}$$

Instead of integrating by taking a weighted sum of values of the integrand at fixed points within the integration domain, it is possible to approximate the same integral by a sum of values of the function evaluated at randomly selected points; this is known as Monte Carlo integration. The simplest way to implement this would be to take a set of points x_i sampled randomly from a uniform distribution over the domain of integration. The mean value $\langle f \rangle$ of the integrand $f(x)$ can be approximated by taking the mean $\overline{f}$ of the integrand evaluated at

these points:

$$\langle f \rangle = \frac{1}{b-a} \int_a^b f(x) \mathrm{d}x \tag{1.43}$$

$$\overline{f}_n = \frac{1}{n} \sum_{i=1}^{n} f(x_i), \qquad x_i \sim U(a,b) \tag{1.44}$$

$$\langle f \rangle \approx \overline{f}_n \pm \frac{\sigma_f}{\sqrt{n}} \tag{1.45}$$

where $U(a,b)$ is the uniform distribution on the interval (a,b), and the notation $x_i \sim p$ means that the random variates x_i are sampled from distribution p. The error in the approximation given in Eq. (1.44) depends on the standard deviation σ_f of $f(x)$, and it decreases as $n^{-1/2}$, so that in the limit of infinite n, this sum converges to the exact value of the integral. This convergence is statistical, not spatial, so it does not depend on the dimensionality of the integral.

1.6.2 Importance sampling

The simple Monte Carlo method described in the previous section has two significant limitations: it is not possible to sample uniformly over an infinite domain (e.g., when integrating over an N-electron wave function which spans all of $\mathbb{R}^{3N}$), and there is no way to decrease the σ_f factor in the error estimate. One way to solve both of these problems is to transform the integrand and integration variable. Consider some function $p(x)$ which is defined on $x \in [a,b]$ such that it is positive everywhere and normalized $\int_a^b p(x)\mathrm{d}x = 1$. The integral in Equation (1.43) can be rewritten as

$$I = \int_a^b \left(\frac{f(x)}{p(x)} \right) p(x) \mathrm{d}x \tag{1.46}$$

$$= \int_a^b g(x) p(x) \mathrm{d}x \tag{1.47}$$

where $g(x) = \frac{f(x)}{p(x)}$. The cumulative distribution function (CDF) $y(x)$ of the distribution $p(x)$ is

$$y(x) = \int_a^x p(x') \mathrm{d}x', \tag{1.48}$$

and this allows integration over y instead of x:

$$I = \int_0^1 g(x(y))\mathrm{d}y \tag{1.49}$$

$$\approx \frac{1}{n}\sum_{i=1}^{n} g(x(y_i)), \qquad y_i \sim U(0,1) \tag{1.50}$$

$$\approx \frac{1}{n}\sum_{i=1}^{n} g(x_i), \qquad x_i \sim p \tag{1.51}$$

where $x(y)$ is the inverse of $y(x)$. This transformation solves the problem of integrating over an infinite domain, and with a judicious choice of $p(x)$ (so that σ_g is small) it can also reduce the prefactor in the error estimate.

Because $x(y)$ is the inverse CDF of $p(x)$, taking $x(y_i)$ where y_i are sampled uniformly from $[0, 1)$ is equivalent to sampling x_i directly from p. If this is to be done directly (i.e., first sample y_i and then transform according to the inverse CDF), it limits the choices of p to distributions for which the inverse CDF can be evaluated. This greatly limits the ability to choose a distribution that will give a low σ_g, so it is useful to find another method to sample from p. One such method is the Metropolis-Hastings algorithm.[21,22] This can be used to sample from any probability distribution $p(\mathbf{x})$ as long as it can be evaluated for any $\mathbf{x}$ in the integration domain.

Beginning from some initial position $\mathbf{x}$, a new position $\mathbf{x}'$ is chosen from a transition distribution $T(\mathbf{x}' \leftarrow \mathbf{x})$; the move to $\mathbf{x}'$ is accepted with probability $A(\mathbf{x}' \leftarrow \mathbf{x})$. This acceptance probability is not uniquely defined, but it must satisfy the following condition (detailed balance):

$$\frac{A(\mathbf{x}' \leftarrow \mathbf{x})}{A(\mathbf{x} \leftarrow \mathbf{x}')} = \frac{T(\mathbf{x} \leftarrow \mathbf{x}')p(\mathbf{x}')}{T(\mathbf{x}' \leftarrow \mathbf{x})p(\mathbf{x})} \tag{1.52}$$

A common choice for A is:

$$A(\mathbf{x}' \leftarrow \mathbf{x}) = \min\left(1, \frac{T(\mathbf{x} \leftarrow \mathbf{x}')p(\mathbf{x}')}{T(\mathbf{x}' \leftarrow \mathbf{x})p(\mathbf{x})}\right) \tag{1.53}$$

If a set of points is generated according to this procedure, they will be distributed according to $p(\mathbf{x})$.

1.6.3 Variational Monte Carlo

The use of Monte Carlo integration to evaluate or optimize expectation values of quantum mechanical operators is known as variational Monte Carlo (VMC). Consider a slight modification to Eq. (1.3) in which the wave function is not normalized; this can be rearranged to have the same form as Eq. (1.50):

$$E = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{1.54}$$

$$= \frac{\int \Psi^*(\mathbf{R})\hat{H}\Psi(\mathbf{R})\mathrm{d}\mathbf{R}}{\int \Psi^*(\mathbf{R}')\Psi(\mathbf{R}')\mathrm{d}\mathbf{R}'} \tag{1.55}$$

$$= \int \left(\frac{|\Psi(\mathbf{R})|^2}{\int |\Psi(\mathbf{R}')|^2 \mathrm{d}\mathbf{R}'}\right) \frac{\hat{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})}\mathrm{d}\mathbf{R} \tag{1.56}$$

$$= \int p(\mathbf{R}) E_L(\mathbf{R})\mathrm{d}\mathbf{R} \tag{1.57}$$

$$E \approx \frac{1}{n}\sum_{i=1} n E_L(\mathbf{R}_i), \qquad \mathbf{R}_i \sim p \tag{1.58}$$

where $\mathbf{R}$ represents the positions of all electrons in the system, and $E_L(\mathbf{R}) = \frac{\hat{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$ is the local energy evaluated at $\mathbf{R}$. The Metropolis-Hastings algorithm can be used to sample points $\mathbf{R}_i$ from $|\Psi|^2$, so this integral can be performed with any wave function as long as one can evaluate the value and laplacian (for the kinetic energy contribution to E_L) at any point in space. Because consecutive points generated by this algorithm are not independent, the statistical error of the approximation given by Eq. (1.57) will be greater than $\frac{\sigma_{E_L}}{\sqrt{n}}$. If one defines the correlation time t_{corr} to be the average number of steps between independent samples, then the error is given by $\frac{\sigma_{E_L}\sqrt{t_{corr}}}{\sqrt{n}} = \frac{\sigma_{E_L}}{\sqrt{m}}$ where $m = n/t_{corr}$ is the number of independent samples.

1.6.4 Projector Monte Carlo

Projector Monte Carlo techniques[23,24] are a class of methods that take advantage of the fact that imaginary time evolution of a wave function projects out all but the lowest-energy stationary state. The time-dependent Schrödinger equation describes the time evolution of a wave function $\Phi(\mathbf{R}, t)$ for a given Hamiltonian $\hat{H}$.

$$i\frac{\partial}{\partial t}\Phi(\mathbf{R}, t) = \hat{H}\Phi(\mathbf{R}, t) \tag{1.59}$$

When the imaginary unit is incorporated into the differential, the equation describes evolution in imaginary time ($\tau = it$).

$$\frac{\partial}{\partial \tau}\Phi(\mathbf{R},\tau) = -\hat{H}\Phi(\mathbf{R},\tau) \tag{1.60a}$$

$$= \frac{1}{2}\nabla^2\Phi(\mathbf{R},\tau) - V(\mathbf{R})\Phi(\mathbf{R},\tau) \tag{1.60b}$$

When the initial wave function $\Phi(\mathbf{R},0)$ is expressed as a linear combination of eigenfunctions $\phi_i(\mathbf{R})$ of $\hat{H}$,

$$\hat{H}\varphi_i(\mathbf{R}) = E_i\varphi_i(\mathbf{R}) \tag{1.61}$$

$$\Phi(\mathbf{R},0) = \sum_j c_j\varphi_j(\mathbf{R})) \tag{1.62}$$

it becomes apparent that propagation in imaginary time results in convergence to the wave function of the stationary state with the lowest energy. Shifting the Hamiltonian by E_0 will give a stationary solution as $\tau \to \infty$.

$$\Phi(\mathbf{R},\tau) = e^{-(\hat{H}-E_0)\tau}\Phi(\mathbf{R},0) = \sum_j c_j\varphi_j(\mathbf{R})e^{-(E_j-E_0)\tau} \tag{1.63}$$

$$\lim_{\tau\to\infty}\Phi(\mathbf{R},\tau) = c_0\varphi_0(\mathbf{R}) \tag{1.64}$$

In general, these projector methods use repeated application of a short-time propagator to project to the ground state. The methods differ in their representation of the electronic wave function and in the form of the short-time propagator. Much of the rest of this work will focus on one such method, fixed-node diffusion Monte Carlo (DMC).[3–5,25–28]

1.6.5 Diffusion Monte Carlo

DMC represents the electronic distribution by a set of walkers moving through the space of electronic coordinates. When considered separately, each of the two terms in the Hamiltonian describes a time evolution process that is simple to simulate. If the potential energy term were not present, the time evolution over an interval $\Delta\tau$ could be simulated by having each walker undergo a random walk with steps from $\mathbf{R}$ to $\mathbf{R}'$ (diffusive steps) sampled according to Eq. (1.65)

$$p_{\text{diff}}(\mathbf{R}' \leftarrow \mathbf{R}) = (2\pi\Delta\tau)^{-3N/2} \exp\left(-\frac{|\mathbf{R}-\mathbf{R}'|^2}{2\Delta\tau}\right) \tag{1.65}$$

If the kinetic energy term were not present, the time evolution could be simulated by scaling the weight of each walker in the distribution by $\exp[(V(\mathbf{R}) - E_0)\Delta\tau]$ (branching steps). To simulate time evolution due to both of these terms, one must take both branching and diffusive steps. Because the kinetic and potential energy operators do not commute, there is an error due to this method of modeling the time evolution.

$$e^{\hat{H}\Delta\tau} = e^{(\hat{T}+\hat{V})\Delta\tau} = e^{\hat{T}\Delta\tau}e^{\hat{V}\Delta\tau} + \frac{[\hat{V},\hat{T}]}{2}\Delta\tau^2 + O(\Delta\tau^3) \tag{1.66}$$

By taking small time steps and alternating between diffusive and branching steps, this error can be reduced. One can also eliminate the quadratic term in Eq. (1.66) with the second order Trotter-Suzuki formula:

$$e^{\hat{H}\Delta\tau} = e^{(\hat{T}+\hat{V})\Delta\tau} = e^{\hat{V}\Delta\tau/2}e^{\hat{T}\Delta\tau}e^{\hat{V}\Delta\tau/2} + O(\Delta\tau^3) \tag{1.67}$$

By this process, the distribution of walkers converges to that of the wave function of the ground state for the given potential. To correct for the error due to the use of a finite time step, one typically performs several calculations at different time steps, and the resulting energies are extrapolated to a time step of zero.

The representation of the wavefunction as an ensemble of unsigned walkers means that it will correspond to a nodeless (i.e. bosonic) state. In order to apply DMC to fermionic systems, antisymmetry of the wave function must be enforced. In practice, this is done by forming a mixed density $\rho(\mathbf{R},\tau)$, which is the product of a trial wave function $\Psi_T(\mathbf{R})$ and the "exact" wave function $\Phi(\mathbf{R},\tau)$ (exact within the constraint of the given nodal surface). This

density will have the same nodes as the trial wave function. Because these nodes are fixed, the density will not converge to that of the (nodeless) bosonic ground state, and the energy obtained by sampling this density will be that of the lowest energy wave function with the given nodal surface. Formation of the mixed density not only solves the problem of enforcing antisymmetry, but it also reduces statistical error by improving sampling, and it changes the physical interpretation of the walker distribution to a density rather than a wave function.

The time propagation of the mixed density differs from that of the exact wave function: multiplication by Ψ_T changes the branching term and introduces a new drift term in the equation. This process can be modeled in a similar manner as described above.

$$\frac{\partial}{\partial\tau}\rho(\mathbf{R},\tau) = \frac{1}{2}\nabla^2\rho(\mathbf{R},\tau) - \nabla\cdot(\mathbf{V}_d(\mathbf{R})\rho(\mathbf{R},\tau)) - (E_L(\mathbf{R}) - E_T)\,\rho(\mathbf{R},\tau) \tag{1.68}$$

$$\mathbf{V}_d(\mathbf{R}) = \frac{\nabla\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \tag{1.69}$$

The drift velocity $\mathbf{V}_d(\mathbf{R})$ acts to steer walkers away from nodes and toward regions where the magnitude of the wave function is larger. In the stochastic simulation of the time evolution of the distribution of walkers, the drift term introduces a deterministic step to the algorithm.

1.6.6 Improved trial functions

Because of the fixed-node error and the effect of the wave function variance on DMC convergence, it is important to use an accurate trial function (i.e., one with a low variance and with nodes that closely approximate those of the exact system). DMC can be useful for calculating interaction energies between molecules/clusters: one can calculate the energy of the individual monomers and of the interacting system and then subtract one from the other. If the trial functions for all calculations are of similar quality, then the errors in each will roughly cancel and an accurate interaction energy can be obtained; however, evaluation of the quality of a trial function is not straightforward. In cases of weakly interacting molecules, it is often sufficient to use a single-determinant (SD) trial function.[29,30] In these systems, the nodal surface near each monomer (i.e., near each region of high electron density) is relatively

unchanged when the other parts of the system are added; however, there are counterexamples of weakly interacting systems for which SD trial functions are not sufficient, some of which will be discussed in later chapters.

1.6.6.1 Jastrow correlation factor In DMC calculations, a determinant (or linear combination of determinants) $D(\mathbf{R})$ of single-particle orbitals is usually multiplied by a Jastrow factor[31–33] $\exp[J(\{r_{ij}\},\{r_{iI}\})]$ which is a positive function that depends explicitly on electron-electron and electron-nucleus distances r_{ij} and r_{iI}, respectively. Because the Jastrow factor is strictly positive, it does not alter the nodal surface of the trial function,

$$\Psi_T(\mathbf{R}) = D(\mathbf{R}) \exp[J(\{r_{ij}\},\{r_{iI}\})] \tag{1.70}$$

so it has no effect on the fixed-node error; however, it does decrease the variance of the trial function, which leads to improved importance sampling and faster convergence of the calculated energy. The Jastrow functions J used in this work consist of one-body (electron-nucleus), two-body (electron-electron) and three-body (electron-electron-nucleus) terms, J_{1I}, J_2, and J_{3I}, respectively, where the subscript I refers to a unique nucleus.

$$J_{1I} = \sum_{i}^{N_{elec}} \chi_I(r_{iI}) \tag{1.71}$$

$$J_2 = \sum_{i<j}^{N_{elec}} u(r_{ij}) \tag{1.72}$$

$$J_{3I} = \sum_{i<j}^{N_{elec}} f_I(r_{ij}, r_{iI}, r_{jI}) \tag{1.73}$$

In this work, the functions χ_I, u, and f_I are either polynomials or cubic splines, and separate u and f_I terms are used depending on whether pairs of electrons have the same or opposite spins.

1.7 DMC trial functions: several case studies

As mentioned above, single determinant trial functions are often assumed to be sufficiently accurate to obtain chemically accurate results when used in DMC calculations. The following chapters present three case studies in which we explore this assumption.

1.7.1 Benzene dimer

The parallel displaced benzene dimer is a model for π-stacking, the attractive interaction between aromatic rings. Chapter 2 describes a study of the use of SD trial functions to calculate the strength of this interaction. We find that the binding energy is significantly underestimated, and that more sophisticated trial functions are necessary.

1.7.2 H_4

By placing four hydrogen atoms at the corners of a square and then stretching it to a rectangular geometry, one obtains a simple model system that is easily tunable by varying a single parameter. At the square geometry, the frontier orbitals are degenerate; as the system is stretched, this degeneracy breaks, and at long distances it resembles two isolated H_2 molecules (although not at equilibrium bond length). In Chapter 3, we calculate the energy of this system with several DMC trial functions and with traditional quantum chemistry methods. We show that a SD trial function fails to give accurate results at or near the square geometry, but that it becomes more accurate as the system is stretched. We show how the fixed-node error associated with the use of a SD trial function compares to several measures of the near-degeneracy and multiconfigurational nature of the system, including the ROHF triplet orbital gap, the size of the leading determinant coefficient in a CAS calculation, and the correlation energy recovered by a CAS calculation.

1.7.3 Be_2

In Chapter 4, we calculate the binding energy of the beryllium dimer with DMC. For this weakly-interacting system, SD trial functions do not give a quantitatively correct binding energy. In small determinantal expansions, the fixed-node error is also not balanced between the atom and the dimer. We describe a novel method for extrapolating to the full configurational space and find that this gives an accurate binding energy (within 0.1 kcal mol^{-1} of the best experimental estimate).

1.8 Future work: CIPSI

It is clear that there are many cases in which it is non-trivial to generate DMC trial functions for interacting systems that will result in a cancellation of fixed-node error between the isolated monomers and the full system. A procedure that could reliably generate multideterminant trial functions of similar accuracy for isolated and interacting monomers would allow much more accurate DMC energies to be obtained over a wider variety of systems. The SCI methods discussed above are promising candidates for this type of trial function generation. Chapter 5, which is adapted from Ref. [20], describes new features implemented in Quantum Package, a software package that can perform SCI calculations. Quantum Package offers several ways to estimate and control the quality of a multideterminant wave function. If certain measures of wave function quality are determined to be strongly coupled to fixed-node error, this could allow more reliable cancellation of error in DMC calculations.

2.0 Diffusion Monte Carlo Study of the Parallel Displaced Form of the Benzene Dimer

The text and figures in this chapter have been reprinted with permission from Gasperich, K.; D. Jordan, K., Diffusion Monte Carlo Study of the Parallel Displaced Form of the Benzene Dimer. In *Recent Prog. Quantum Monte Carlo*, Tanaka, S., Roy, P.-N., Mitas, L., Eds.; ACS Symposium Series, Vol. 1234; American Chemical Society: 2016, pp 107–117, DOI: 10.1021/bk-2016-1234.ch007. The author's contribution to the work included performing all calculations, generating all figures, and editing/revising the manuscript.

2.1 Summary

The diffusion Monte Carlo (DMC) method is used to calculate the interaction energy of the parallel displaced form of the benzene dimer. The calculations were performed with single-determinant Slater-Jastrow trial functions at time steps between 0.0025 and 0.04 a.u., allowing for extrapolation to zero time step. Our calculated interaction energy is considerably smaller in magnitude than the best coupled cluster singles and doubles with perturbative triples [CCSD(T)] estimate, leading us to conclude that there is a sizable fixed-node error due to the use of a single-determinant trial function.

2.2 Introduction

The π-stacked parallel displaced (PD) form of the benzene dimer is an important test system for examining the performance of electronic structure methods at describing dispersion interactions.[35–49] In addition to serving as a prototype for π stacking, it is also a prototypical system for which the Møller-Plesset second-order perturbation theory (MP2) method[1]

significantly overbinds the dimer due to its overestimation of the dispersion contribution to the interaction energy.[50] Insight into the origin of the overbinding at the MP2 level can be gained by considering the definition of the dispersion energy in terms of the Casimir-Polder integral[51] over complex frequencies of the polarizabilities of the two monomers. The MP2 dispersion energy is equivalent to that obtained from the Casimir-Polder integral employing uncoupled Hartree-Fock (UCHF) polarizabilities of the monomers.[52] In the case of the PD form of the benzene dimer, the magnitude of the dispersion contribution is significantly reduced if the coupled perturbed Hartree-Fock (CPHF) polarizability function is used instead of the UCHF polarizability function in evaluating the integral.[48,53,54] The coupled cluster singles and doubles with perturbative triples [CCSD(T)] method,[2] when used with sufficiently flexible basis sets, provides a quantitatively accurate description of the interaction between the benzene rings;[55] however, due to its $O(N^7)$ scaling with respect to system size, this approach is computationally prohibitive for much larger systems. This has led naturally to an interest in lower-scaling methods that can achieve accuracies comparable to that of CCSD(T). Diffusion Monte Carlo (DMC)[3–5,25,27,28] is one such method that is being increasingly employed to characterize weak interactions in dimers and larger clusters as well as in molecular crystals and layered materials. DMC has several advantages over CCSD(T), including lower scaling ($\sim O(N^3)$) with system size, weaker sensitivity to the basis sets employed, and better scalability over large numbers of CPU cores.

Most of the error in DMC energies in the zero time step limit is due to the fixed-node approximation, which is made to ensure that a Fermionic wave function results. The vast majority of DMC calculations employ a single Slater determinant of Hartree-Fock or density functional theory (DFT) orbitals to impose the fixed nodes. In the calculation of interaction energies for weakly interacting systems, it is generally assumed that the errors due to the fixed-node approximation cancel when twice the energy of the monomer is subtracted from the energy of the dimer at its equilibrium geometry.[29,30] However, there are weakly interacting dimers, e.g., Be_2, for which the use of a single Slater determinant to impose the nodal surface is known to be inadequate.[56] In practice, in describing weakly interacting systems with DMC, there are also the challenges of reducing the statistical and finite time step errors to a small fraction of the interaction energy of interest. In this work, we apply the DMC method to the

PD form of the benzene dimer. The motivation of the present study is to test the suitability of single-determinant trial functions to fix the nodal surfaces in describing the interaction energy of the PD benzene dimer.

2.3 Methodology

The quantum Monte Carlo calculations were carried out using the standard procedure of first generating a single-determinant Slater-Jastrow trial function (i.e., a trial function consisting of a single Slater determinant multiplied by a Jastrow factor[31–33]). The Jastrow factor does not affect the nodal surface of the trial function. Trail-Needs AREP pseudopotentials[57,58] were employed together with contracted Gaussian-type-orbital (GTO) basis sets that were designed for use with these pseudopotentials.[59] $5s5p2d1f$ and $5s2p1d$ basis sets were used for C and H, respectively. The Becke3LYP[60–63] density functional method was used to generate the orbitals employed in the trial functions. The Jastrow factors included electron-electron (e-e), electron-nuclear (e-n), and three-body electron-electron-nuclear (e-e-n) terms, with 17, 22, and 34 parameters in the e-e, e-n, and e-e-n factors, respectively; these parameters were optimized by means of the variational Monte Carlo (VMC) method. The resulting trial functions were then used in carrying out the DMC calculations. The geometries of the benzene monomer and PD dimer are taken from Miliordos et al.[55] who optimized the structures at the CCSD(T)/aug-cc-pVTZ[64,65] level of theory, with the dimer geometry being optimized with the constraint of rigid monomers. This structure has one benzene monomer displaced with respect to the other as shown in Figure 1. The center to center displacement in the plane of the rings (R_1) is 1.6835 Å, and the distance between the planes of the two rings (R_2) is 3.4507 Å. Two different reference energies were used for calculating the interaction energy: in one, the reference energy was taken to be twice the energy calculated for the monomer, and, in the other, the reference energy was obtained from calculations on the dimer with the two monomers separated by 10 Å.

The DMC calculations were carried out with 64 000 walkers at time steps of 0.0025, 0.005, 0.0075, 0.01, and 0.04 a.u. The size-consistent version of the T-move method[66] was used to

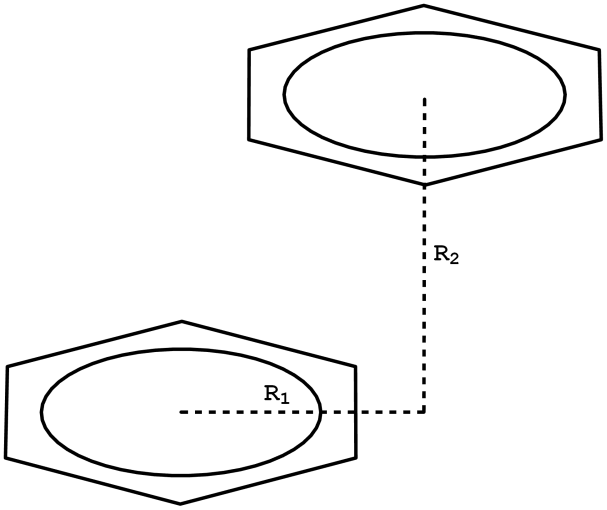

Figure 1: PD form of the benzene dimer used in this study ($R_1 = 1.6835\,\text{Å}$; $R_2 = 3.4507\,\text{Å}$).

deal with the nonlocality of the pseudopotential. The orbitals for the trial functions were generated using the Gaussian 09 program.[67] The QMC calculations were carried out using the CASINO code.[27]

2.4 Results

Figure 2 reports the energies of the PD benzene dimer at its equilibrium geometry, the dimer at a separation of 10 Å, and twice the energy of the monomer, all obtained from DMC calculations with time steps of 0.0025, 0.005, 0.0075, and 0.01 a.u.. Figure 3 reports the associated interaction energies. Both linear and quadratic ($a + b\tau + c\tau^2$) fits of the data are presented. Linear extrapolation to zero time step yields interaction energies of $-1.86(12)$ and $-2.02(14)\,\text{kcal mol}^{-1}$ when using as the reference twice the energy of the monomer and the long distance dimer calculation, respectively. The corresponding results with the quadratic fits are $-1.80(30)$ and $-1.90(34)$ kcal/mol. We place more confidence in the latter values because the DMC moves were governed by the algorithm of Umrigar

and coworkers[4] which was designed to be quadratically convergent in the time step. Also, as seen from the results in Figures 2 and 3, the $\tau \to 0$ limits of the energies obtained using the two different reference energies agree much better when using the quadratic than when using the linear fits. Comparison of the DMC values of the interaction energy with the large basis set CCSD(T) result of $-2.65(2)\,\mathrm{kcal\,mol^{-1}}$[55] suggests that there is a $\sim 0.8\,\mathrm{kcal\,mol^{-1}}$ error in the DMC value of the interaction energy of the PD benzene dimer resulting from the use of a single-determinant Slater-Jastrow trial function with B3LYP orbitals. Azadi and Cohen[68] recently reported DMC values of the interaction energy of the PD benzene dimer obtained using a single-determinant Slater-Jastrow trial function both with and without a backflow transformation[69] which allows for a relaxation of the nodal surface. Their structure for the dimer is similar to that employed in our calculations; for R_1 fixed at 1.6 Å, they calculated DMC energies at several values of R_2 to obtain a minimum in the interaction energy at $R_2 = 3.8$ Å. In the absence of backflow, Azadi and Cohen obtained an interaction energy of $-1.8(2)\,\mathrm{kcal\,mol^{-1}}$, which is consistent with our result. With backflow, Azadi and Cohen obtained an interaction energy of $-2.7(3)\,\mathrm{kcal\,mol^{-1}}$, which is in agreement with the CCSD(T) result. However, these results were obtained from linear extrapolation of DMC energies obtained at 0.01 and 0.04 a.u. time steps, and it is not clear how large an error may have resulted from a linear extrapolation of the interaction energies calculated at these relatively large time steps. We return to this issue below. As the present paper was being prepared, we learned of unpublished results of Dubecký and Mitas[70] who obtained, using the protocol described by Dubecký et al.[71] (employing a single-determinant Slater-Jastrow trial function with B3LYP/aug-TZV orbitals and a time step of 0.005 a.u.), an interaction energy of $-2.13(13)\,\mathrm{kcal\,mol^{-1}}$ for the PD benzene dimer, in excellent agreement with our results. The calculations of Dubecký and Mitas used the geometry from the S22 test set,[50] which is very close to that employed here, but the in-plane displacement of one ring relative to the other is in a direction perpendicular to that considered here. Our calculations differ from those of Dubecký and Mitas by our use of larger atomic basis sets to represent the orbitals in the trial functions and by our use of multiple time steps and extrapolation to zero time step.

As noted above, Azadi and Cohen used time steps of 0.01 and 0.04 a.u. in their DMC calculations, and one might expect that the 0.04 a.u. time step, in particular, is outside

the regime that a linear fit is valid. This motivated us to also carry out DMC calculations (without backflow) on the PD benzene dimer and benzene monomer at a time step of 0.04 a.u. time step. Figure 4 displays the results of our DMC calculations including the 0.04 a.u. time step. It is clear from this figure that the energies at the 0.04 a.u. time step are indeed far outside the linear regime and that linear extrapolation of the results at 0.04 and 0.01 a.u. time steps results in much higher total energies than obtained from extrapolation of the energies calculated at the 0.0025–0.01 a.u. time steps. Nonetheless, the values of the interaction energy ($-1.72(10)$ and $-2.05(12)$ kcal mol^{-1} using as the reference twice the energy of the monomer and the energy of the dimer with 10 Å separation of the rings, respectively) obtained from the linear extrapolation of total energies at 0.04 and 0.01 a.u. time steps are close to those obtained from extrapolation of the small time step results (Figure 3 and Table 1) ($-1.86(12)$ and $-2.02(14)$ kcal mol^{-1} using as the reference twice the energy of the monomer and the energy of the dimer with 10 Å separation of the rings, respectively) (or $-1.83(12)$ and $-1.98(14)$ kcal mol^{-1} from extrapolation of the binding energy). However, this check was done in the absence of backflow correlation, and there is still the possibility that a sizable error could result from linear extrapolation of the interaction energies at 0.04 and 0.01 a.u. time steps in the calculations including backflow.

In summary, we conclude that DMC calculations using a single-determinant trial function significantly underestimate the magnitude of the interaction energy of the PD form of the benzene dimer. It is expected that single-determinant trial functions are likely to prove inadequate for accurate calculations of the interaction energies of π-stacked systems in general. Although a recent study of Azadi and Cohen concluded that backflow correlation can remedy this problem, we believe that it is necessary to extend such calculations to smaller time steps than used in the Azadi/Cohen study to establish definitively that this is indeed the case. In principle, the limitation of the use of single Slater determinants to define the nodal surfaces can be overcome by the adoption of multi-configurational trial functions; the challenge with such an approach is the selection of configuration spaces that introduce negligible size consistency errors. In this context we note that Sorella et al.[72] have carried out lattice regularized DMC (LRDMC)[73] calculations of the PD benzene dimer using Jastrow correlated antisymmetrized geminal power wavefunctions[74,75] to establish the nodal surfaces. These

calculations gave an interaction energy of $-2.2(3)\,\mathrm{kcal\,mol^{-1}}$, which is only slightly larger in magnitude than that obtained in this study using a single-determinant trial function. They also use a geometry in which the rings are displaced similarly to those of the PD benzene dimer from the S22 test set.

2.5 Acknowledgments

This research was supported by grant CHE1362334 from the National Science Foundation and by computational resources at the University of Pittsburgh's Center for Simulation and Modeling and at ALCF, funded by DOE INCITE award CPH005. We acknowledge valuable discussions with Dr. M. Deible

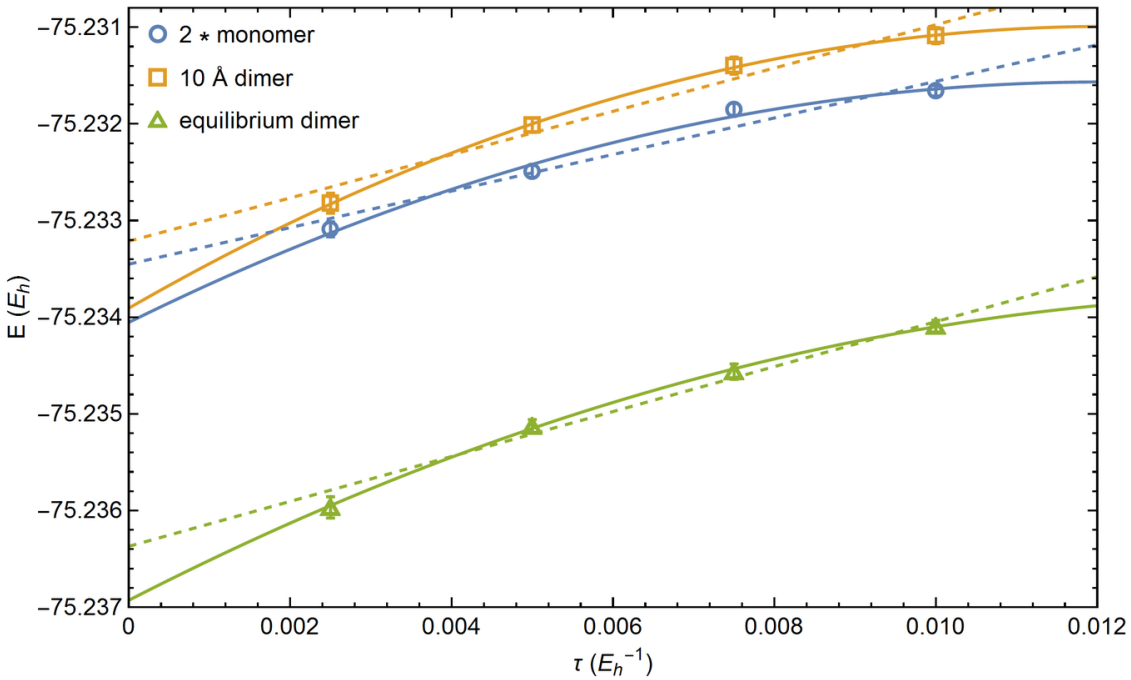

Figure 2: Total energies from DMC calculations on the PD form of the benzene dimer at its equilibrium geometry as well as on the dimer with 10 Å separation between the rings and twice the energy of the benzene monomer. Both linear and quadratic fits (dotted and solid lines, respectively) of the energies at time steps ranging from 0.0025 to 0.010 a.u. are presented.

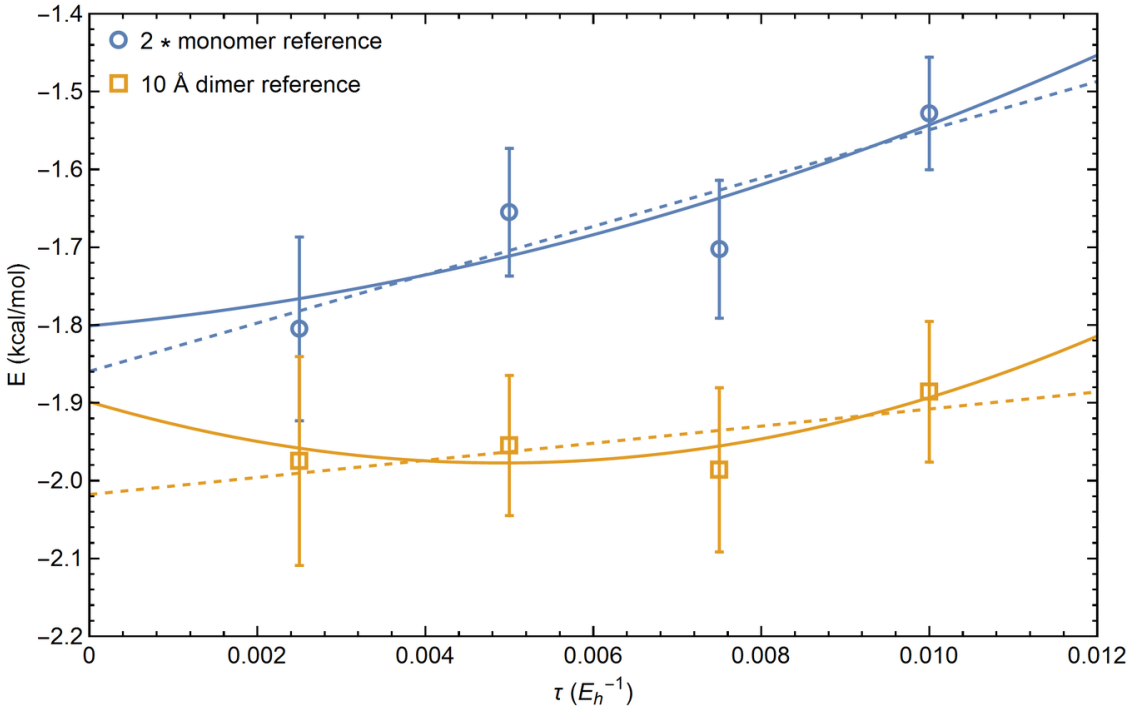

Figure 3: Interaction energies from DMC calculations on the PD form of the benzene dimer calculated as the difference between the energy at equilibrium geometry and either twice the energy of the monomer or the energy of the dimer with 10 Å separation between the rings. Both linear and quadratic fits of the energies at time steps ranging from 0.0025 to 0.010 a.u. are presented (dashed and solid lines, respectively).

Table 1: DMC values of the interaction energy of the PD form of the benzene dimer.

Time step (a.u.)	Interaction energy (kcal/mol)	
	2 × monomer ref.	dimer at 10 Å ref.
0.01	-1.53(7)	-1.89(9)
0.0075	-1.70(9)	-1.99(11)
0.0050	-1.65(8)	-1.95(9)
0.0025	-1.80(12)	-1.97(13)
0[a]	-1.86(12)	-2.02(14)
0[b]	-1.80(30)	-1.90(34)

[a] extrapolated using a linear fit
[b] extrapolated using a quadratic fit

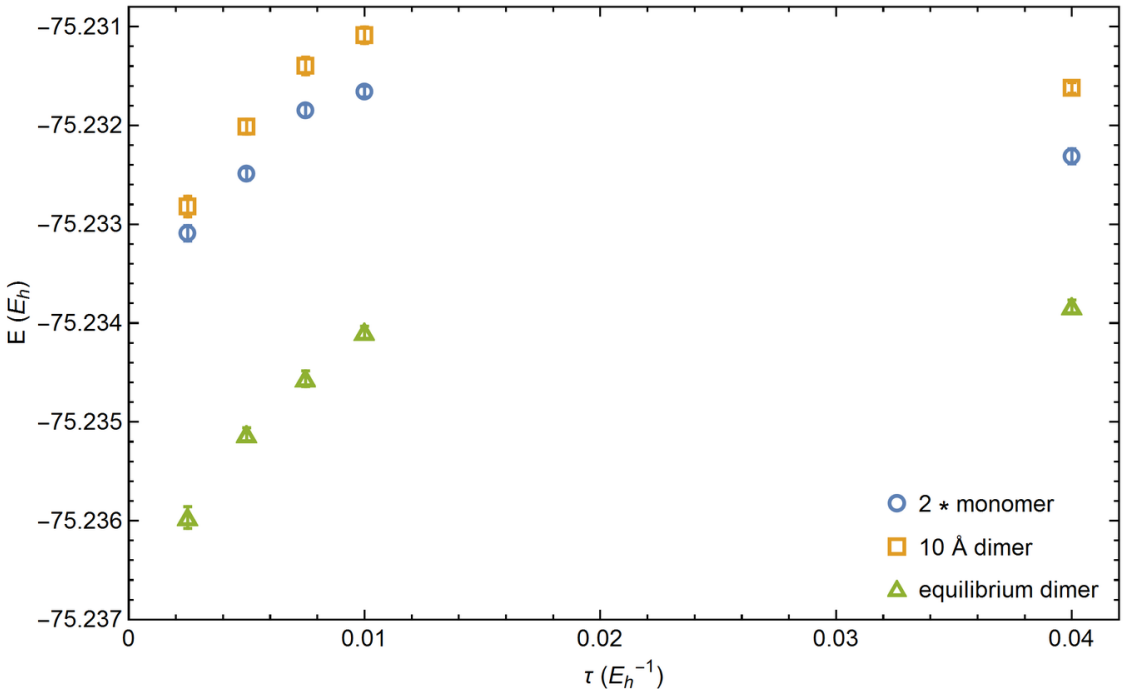

Figure 4: DMC energies of the PD benzene dimer at its equilibrium geometry, the dimer with the monomers separated by 10 Å, and twice the energy of the monomer. In addition to the data presented in Figure 2, DMC energies for the time step of 0.04 a.u. are reported.

3.0 H_4: A Model System for Assessing the Performance of Diffusion Monte Carlo Calculations Using a Single Slater Determinant Trial Function

The text and figures in this chapter have been adapted from Gasperich, K.; Deible, M.; Jordan, K. D. H_4 : A model system for assessing the performance of diffusion Monte Carlo calculations using a single Slater determinant trial function. *J. Chem. Phys.* **2017**, *147*, 074106, DOI: 10.1063/1.4986216, with the permission of AIP Publishing. The author's contribution to the work included performing all calculations, generating all figures, and editing/revising the manuscript.

3.1 Summary

A model H_4 system is used to investigate the accuracy of diffusion Monte Carlo (DMC) calculations employing a single Slater determinant to fix the nodal surface. The lowest energy singlet state of square H_4 is a diradical which is poorly described by DMC calculations using a single determinant (SD) trial function. Here we consider distortions to rectangular structures, which decrease the amount of diradical character. The falloff of the error in the SD-DMC energy with increasing separation between the two H_2 molecules is found to be much more rapid for small distortions away from square than for large distortions. This behavior is shown to be correlated with the extent of mixing between the two configurations needed to properly describe the diradical character. The error in the SD-DMC energy is found to be sizeable ($\sim$0.1 eV) even for separations at which the coefficient of the dominant configuration in a four-electron, four-orbital complete active space self-consistent-field wave function is as large as 0.9.

3.2 Introduction

The diffusion Monte Carlo (DMC) method[5,25] is increasingly being used to address problems in electronic structure theory. In general, to ensure fermionic behavior, it is necessary to employ a trial wave function to impose a nodal surface for exchange of electrons.[5] In principle, fixed-node DMC calculations, if run for a sufficiently long time and extrapolated to the zero time step, can yield the exact energy for the imposed nodal surface. The vast majority of DMC calculations have employed a single Slater determinant of Hartree-Fock or density functional theory orbitals to fix the nodal surface. While this approach has proven successful for a wide range of problems, it is known to be inadequate for systems that have a strong static correlation.[56,77–80] However, the relationship between the error in SD-DMC energies and the extent of configuration mixing, e.g., as found in multiconfigurational methods such as complete active space self-consistent field (CASSCF)[81] or configuration interaction (CI), is not well understood. Except for a few model systems, knowledge of how the inclusion of configuration mixing in the trial function impacts the nodal surface for exchange of electrons is lacking.[82–91] We are especially interested in quantifying how the error in the SD-DMC energy depends on the energy gap between the frontier orbitals of a molecule or cluster. In this work, we explore this issue for a model planar $(H_2)_2$ system.

The motivation for choosing the $(H_2)_2$ system for exploring the near degeneracy issue can be seen from Fig. 5 which plots as a function of the separation between the two H_2 molecules the energies of the four valence molecular orbitals from restricted open-shell Hartree-Fock (ROHF) calculations of the lowest energy triplet state. (Details of these calculations are given below.) The nodal patterns of the valence orbitals are also depicted in Fig. 5. For the square structure, the two frontier orbitals (of b_{2u} and b_{3u} symmetry as labeled in the D_{2h} point group) are degenerate; hence, the lowest energy singlet state of H_4 is diradical in nature, requiring a minimum of two Slater determinants for a proper description of its wave function. This state belongs to the B_{1g} representation in the D_{4h} point group. By distorting the structure from square to rectangular, with increasing (or decreasing) distance between the two H_2 molecules, one can tune the extent of configuration mixing in the wave function. In this study, SD-DMC energies are calculated and compared with the results of full configuration interaction (FCI),

coupled-cluster singles and doubles plus perturbative triples [CCSD(T)][2] calculations, as well as of multi-determinant DMC calculations employing CAS(2,2) and CAS(4,4) trial wave functions, where CAS(n,m) denotes CASSCF with n active electrons in m orbitals.

DMC calculations on square H_4 using multi-configurational trial functions have been reported by Anderson.[78] The present study extends that of Anderson in that it considers rectangular structures which remove the orbital degeneracy, employs a larger basis set for the trial functions, includes a comparison with the results of FCI and CCSD(T) calculations, and correlates the error in the SD-DMC energy with various measures of the near degeneracy in the system.

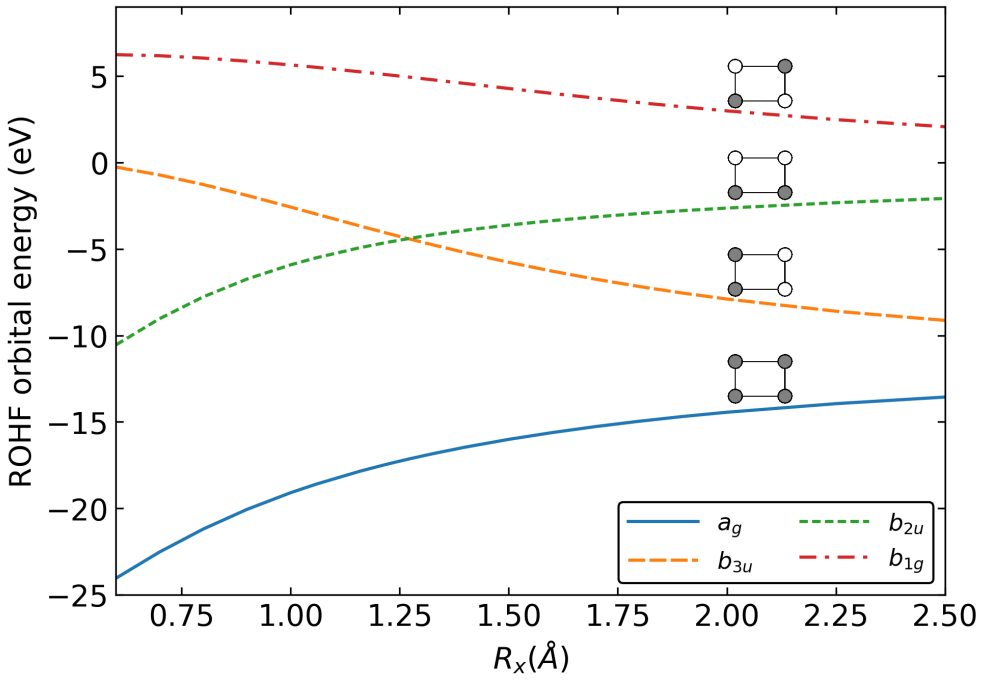

Figure 5: Energies of the four valence orbitals of $(H_2)_2$ from ROHF (Guest and Saunders)/cc-pVQZ calculations on the lowest energy triplet state as a function of the distance between the two H_2 molecules, with a bond length of 1.27 Å).

3.3 Computational details

The structure of square H_4 with sides of 1.27 Å was taken from a theoretical study of Silver and Stevens.[92] In generating rectangular structures, two parallel sides of the rectangle were fixed at 1.27 Å, and the lengths, R, of the other two sides were varied. Thus, one can view the system as two interacting stretched H_2 molecules. The Hartree-Fock (HF), CASSCF, FCI, and CCSD(T) energies were calculated using the correlation-consistent cc-pVQZ basis set.[64] The HF and CASSCF trial wave functions used in the DMC calculations were expanded in a basis set consisting of the s functions from the cc-pV5Z basis set[64] and the p and d functions from the cc-pVTZ basis set.[64] The ROHF calculations on the triplet state used the parametrization of Guest and Saunders[93] to define the diagonal blocks of the Fock matrix.

The DMC calculations employed Slater-Jastrow trial functions, which are the products of a Jastrow factor[31] and a determinantal wave function, consisting of either a single Slater determinant (from a HF calculation) or a linear combination of determinants (from a CASSCF calculation). The one- and two-body Jastrow factors in this work are represented using cubic B-splines as described by Esler et al.[94] The form of the three-body Jastrow proposed by Drummond et al.[32] was also used. Because the Jastrow factors are positive everywhere, they do not affect the nodal surface of an all-electron trial function; thus, the nodes (and the fixed-node error) depend only on the Slater determinant part of the trial function. The parameters in the Jastrowfactors were optimized via variational Monte Carlo (VMC) by minimizing a cost function consisting of 95% energy and 5% variance.[95] In the case of CASSCF wave functions, the coefficients of the Slater determinants and the Jastrow parameters were optimized simultaneously in the VMC step. The DMC calculations were carried out with time steps of 0.005, 0.0025, and 0.001 a.u., with the resulting energies being extrapolated to the zero time step using linear fits. (Due to the small values of the employed time steps, similar results are obtained when using quadratic fits for the extrapolation.) The HF, CCSD(T), CASSCF, and CI calculations were performed using the GAMESS program,[96] and the quantum Monte Carlo calculations were carried out using the QMCPACK program.[97]

3.4 Results and discussion

Before reporting the energies from the various calculations, we first consider how configuration mixing impacts the nodal surface of H_4 in the square structure. Figure 6 reports cuts through the nodal surfaces of the $(1a_g)^2(1b_{2u})^2$ and $(1a_g)^2(1b_{3u})^2$ configurations as well as of the CAS(2,2) wave function involving these two configurations. The cuts are analogous to those considered by Anderson in his study of square H_4; two electrons of the same spin are held stationary,and changes in the node are shown as the position of a third electron moves in the plane of the molecule, i.e., the xy-plane. As expected, the nodal surface of the CAS(2,2) wave function is fundamentally different from that of either of the single determinant wave functions.

Figure 7 reports as a function of separation between the two H_2 molecules the energies obtained using various theoretical methods as well as the square of the coefficient of the dominant Slater determinant in the CAS(4,4) wave function. As seen from this figure, the FCI and CAS(4,4)-DMC potential energy curves are nearly identical, with the latter being slightly lower in energy as a result of finite basis set errors in the FCI calculations. For the square structure, the CAS(4,4)-DMC calculations give an energy of $-2.119\,31(4)$ a.u., compared to the $-2.091(26)$ a.u. DMC energy reported in Ref. [78]. Although not included in the figure, we note that the potential energy curve from the CAS(2,2)-DMC calculations is very close to that from the CAS(4,4)-DMC calculations. In contrast, the potential energy curve from the SD-DMC calculations differs markedly from those from the FCI and CAS(4,4)-DMC calculations, with the maximum error ($\sim$0.9 eV) occurring for the square structure, at which the CAS(4,4) wave function has two dominant configurations each entering with a coefficient of 0.67 in magnitude. Moreover, the error in the SD-DMC energy is still $\sim$0.1 eV (relative to the CAS(4,4)-DMC energy) when the dominant configuration in the CAS(4,4) wave function has a coefficient of 0.9. Interestingly, the CCSD(T) potential energy curve lies close to the FCI potential energy curve except for geometries very close to the square transition state structure.

Figure 8 reports the error in the SD-DMC energy for R values between 1.27 and 3.0 Å. In each of the two regimes, $R \lesssim 1.4$ Å and $R \gtrsim 1.8$ Å, the data points are well fit by exponentials,

as shown in the figure. Since the DMC calculations using CAS(2,2) and CAS(4,4) trial wave functions give nearly identical energies, which, in turn, are close to the full CI energies, it follows that the true nodal surface is well represented by the two-configuration wave function. This has motivated us to examine how the error in the SD-DMC energy correlates with the extent of mixing between the two configurations as described in a minimum basis set treatment. For this model, the orbitals are fully determined by symmetry, so that the CAS(2,2) calculation is equivalent to the two-configuration CI given by

$$\begin{pmatrix} 0 & K_{bc} \\ K_{bc} & 2\Delta \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \Delta E \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{3.1}$$

where the energy of the $(1a_{\mathrm{g}})^2(1b_{2\mathrm{u}})^2$ configuration is set to zero, 2Δ is the energy difference between the $(1a_{\mathrm{g}})^2(1b_{2\mathrm{u}})^2$ and $(1a_{\mathrm{g}})^2(1b_{3\mathrm{u}})^2$ configurations, K_{bc} is the coupling of the two configurations, and b and c are used to denote the $1b_{2\mathrm{u}}$ and $1b_{3\mathrm{u}}$ orbitals, respectively. The solution of the eigenvalue problem gives the energy lowering due to the configuration mixing,

$$\Delta E = \Delta - \sqrt{\Delta^2 + {K_{bc}}^2} \tag{3.2}$$

where

$$\Delta = \varepsilon_b - \varepsilon_c + \frac{1}{2}\left(J_{cc} - J_{bb}\right) - 2J_{bc} + K_{bc} \tag{3.3}$$

The expression for Δ in Eq. 3.3 is in terms of the restricted HF orbitals of the $(1a_{\mathrm{g}})^2(1b_{2\mathrm{u}})^2$ configuration. If this is expressed instead in terms of the ROHF orbitals of the triplet state, using the Guest-Saunders definition, Δ becomes equal to the orbital energy difference. There are two limiting regimes for ΔE,

$$\Delta E \approx \begin{cases} -K_{bc} + \Delta & \Delta \ll K_{bc} \quad \text{(3.4a)} \\ -\dfrac{{K_{bc}}^2}{2\Delta} & K_{bc} \ll \Delta \quad \text{(3.4b)} \end{cases}$$

The square of the normalized c_2 coefficient is

$${c_2}^2 = \frac{{K_{bc}}^2}{2\left(\Delta^2 + \Delta\sqrt{{K_{bc}}^2 + \Delta^2} + {K_{bc}}^2\right)} \tag{3.5}$$

We note also that ΔE and c_2 are related as

$$\Delta E = \frac{K_{bc} c_2}{\sqrt{1 - {c_2}^2}} \tag{3.6}$$

Obviously, c_2 also displays limiting regimes for $\Delta \ll K_{bc}$ and $K_{bc} \ll \Delta$.

In light of these results, we now consider how the error in the SD-DMC energy correlates with the energy lowering, ΔE, due to the mixing of the two configurations for R values between 1.27 and 3.0 Å. As seen from Fig. 9, for $\Delta E \gtrsim 0.5$ eV, the error in the SD-DMC energy varies linearly with ΔE. For small (0.2–0.5 eV) ΔE values, the error in the SD-DMC energy falls off with decreasing ΔE more slowly than would be expected from extrapolation of the linear fit to the large ΔE data. (ΔE remains finite at large separations between the two H_2 molecules due to electron correlation in the isolated monomers.)

Additional insight into the origin of the trends reported in Figs. 8 and 9 is provided by the examination of how K_{bc}, Δ, and ΔE depend on R, with these results being reported in Fig. 10, from which it is seen that both K_{bc} and Δ display an exponential dependence on R, with that in Δ being much steeper than that in K_{bc}. It is also seen that $\Delta \approx K_{bc}$ near $R = 1.42$ Å at which $\Delta E \approx 0.5$ eV. Thus the crossover between the linear and the nonlinear regimes in the plot of the SD-DMC error vs. ΔE (Fig. 9)corresponds to the region where $\Delta \approx K_{bc}$. Figure 11 reports the error in the SD-DMC energy vs. the c_2 coefficient from the 2×2 CI calculation. Again, two limiting regimes are apparent, which is consistent with the relationship between ΔE and c_2 given above.

We now return to the R dependence of the error in the SD-DMC energy reported in Fig. 8, where it was seen that there are two limiting regimes, one for $1.27\,\text{Å} < R \lesssim 1.31\,\text{Å}$ (near the square structure) and the other for large R, each showing an exponential dependence on R but with very different exponents. In these two regimes, ΔE is closely approximated by Eqs. 3.4b and 3.4a, respectively. Given the dependence of the error in the SD-DMC energy on ΔE and the exponential dependence of both K_{bc} and Δ on R, it follows that each of the two limiting regimes in the error in the SD-DMC vs. R plot (Fig. 8) displays an exponential dependence on R.

Orbital energy gaps provide, perhaps, the simplest measure of the configuration mixing, leading us to also examine how the error in the SD-DMC energy depends on Δ, the energy

gap of the two singly occupied orbitals from the calculations on the triplet state. From Fig. 12 it is seen that the error in the SD-DMC energy when reported vs. Δ also displays two limiting regimes, each displaying approximately exponential behavior. The switch over in the behavior occurs near $R = 1.5\,\text{Å}$, which differs slightly from the R value at which $\Delta = K_{bc}$ as reported in Fig. 10. This difference is likely caused by the use of the cc-pVQZ basis set for the calculation of the orbital energies used in Fig. 12 as compared to the minimal basis set used in calculating the quantities shown in Fig. 10.

The errors in the SD-DMC energies for the H_4 model system are due to the inadequacy of a single Slater determinant for describing the nodal surface. Based on earlier studies, we know that each of the $(1a_g)^2(1b_{2u})^2$ and $(1a_g)^2(1b_{3u})^2$ configurations for rectangular H_4 has four nodal domains, while the exact wave function is expected to have two nodal domains.[82–85] Interestingly, exploratory calculations show that the error in the SD-DMC energy for the lowest energy $^2B_{1g}$ state of square ${H_4}^+$, which, like the $^1B_{1g}$ state of H_4, requires two configurations —in this case, $(1a_g)^2(1b_{2u})^2$ and $(1a_g)^2(1b_{3u})^2$ (where the orbital symmetries are specified in the D_{2h} point group) —for a proper description of its wave function, is slightly larger than that found in the SD-DMC calculations on the $^1B_{1g}$ state of H_4.

Although the detailed nature of how the nodal surface of the H_4 model is impacted by the mixing of the second configuration in the CAS(2,2) wave function is not known, it is instructive to focus on the cut that goes through $(x, y) = (0, 0)$. This cut occurs at $\theta = 0°$ and $90°$ for the two single determinant configurations but occurs at an intermediate angle θ for the two-configuration wave function, with θ being $45°$ for the square structure and between $0°$ and $45°$ for structures with $R > 1.27\,\text{Å}$. For the two-configuration wave function $\tan(\theta) = c_2/c_1$, which, based on Eq. 3.6, can also be expressed as $\Delta E/K_{bc}$. Hence, ΔE can be viewed as a measure of the change in the nodal surface brought about by configuration mixing.

3.5 Conclusions

In this work, we have applied the DMC method to a $(H_2)_2$ model system in square and rectangular geometries. It is demonstrated that the error in the energy calculated using the DMC method with a single determinant trial function is sizable even for geometries in which the magnitude of the coefficient of the leading configuration in a CAS(4,4) calculation is as large as 0.9. Most significantly, for geometries where the orbital gap (determined from ROHF calculations) is smaller than 4 eV, the CCSD(T) method outperforms the SD-DMC method. The error in the SD-DMC energy is found to correlate with ΔE, the amount of correlation energy recovered by a CAS(2,2) calculation. The dependence of the error in the SD-DMC energy displays two limiting regimes when plotted against ΔE or R. From analysis of a two-configuration wave function treatment of the system, it is seen that the two limiting regimes correspond to the situations in which off-diagonal coupling is much larger or much smaller than the energy gap between the two non-interacting configurations. We also find that the error in the SD-DMC energy for the $^2B_{1g}$ state of square ${H_4}^+$ is comparable to that for the $^1B_{1g}$ state of the neutral molecule at the same geometry even though in the cation there is one rather than two electrons in the a1g orbital engaged in exchange with electrons in the frontier orbitals.

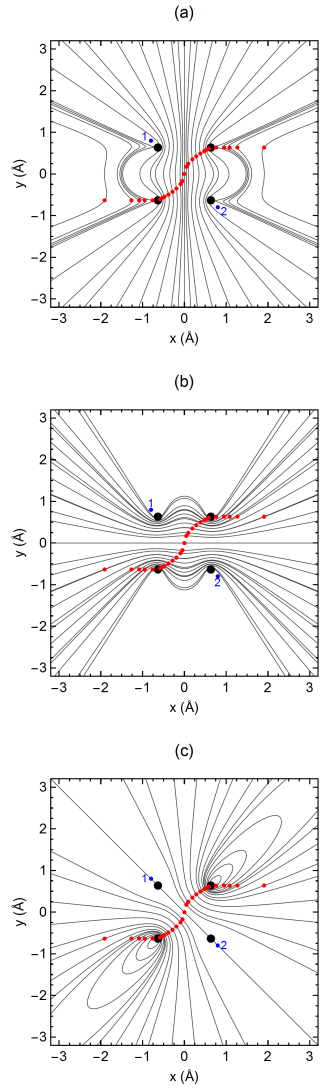

Figure 6: Cuts of the nodal surface of the (a)$(1a_{\rm g})^2(1b_{3\rm u})^2$, (b)$(1a_{\rm g})^2(1b_{2\rm u})^2$, (c)CAS(2,2) wave functions for square H_4. The positions of the nuclei are indicated by the black circles, and the positions of electrons 1 and 2 (both up spin) are indicated by blue dots. Several Choices for the position of electron 3 (red dot) are considered, and the positions for electron 4 for which the wave function is zero are indicated by the curves.

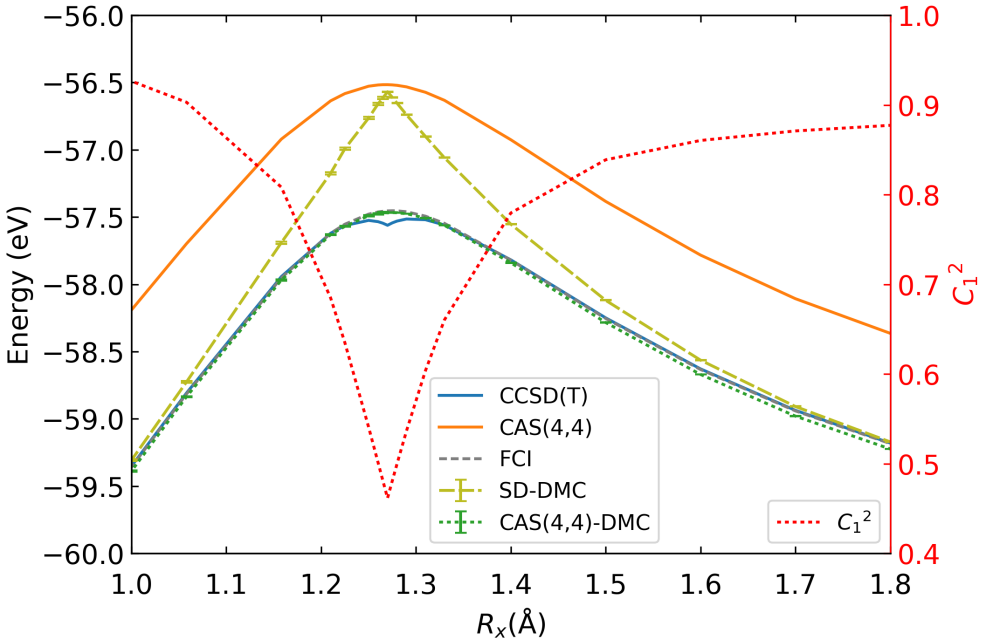

Figure 7: Energy from various theoretical methods and the square of the largest CAS(4,4) vector coefficient as a function of separation between the H_2 molecules, keeping the monomer bond lengths fixed at the value of the square transition state structure. For $R > 1.27$ Å, the SD-DNC calculations used the $(1a_g)^2(1b_{3u})^2$ configuration in the trial wave function, and for $R < 1.27$ Å, the $(1a_g)^2(1b_{2u})^2$ configuration was employed.

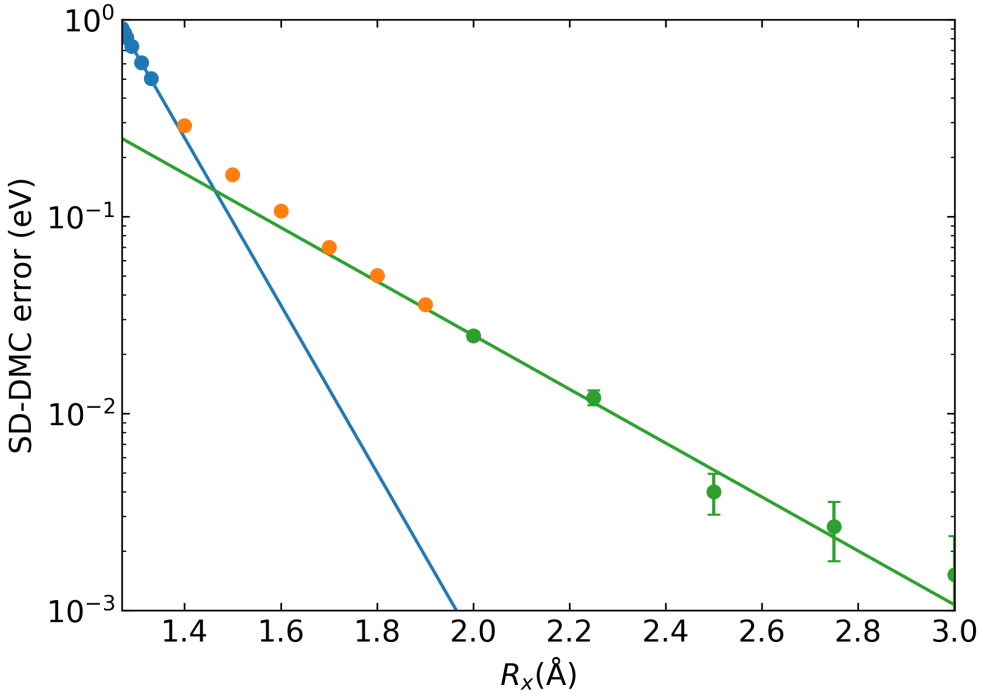

Figure 8: Error in the SD-DMC energy of H_4 (relative to the CAS(4,4)-DMC energy) as a function of the distance between the two H_2 molecules.

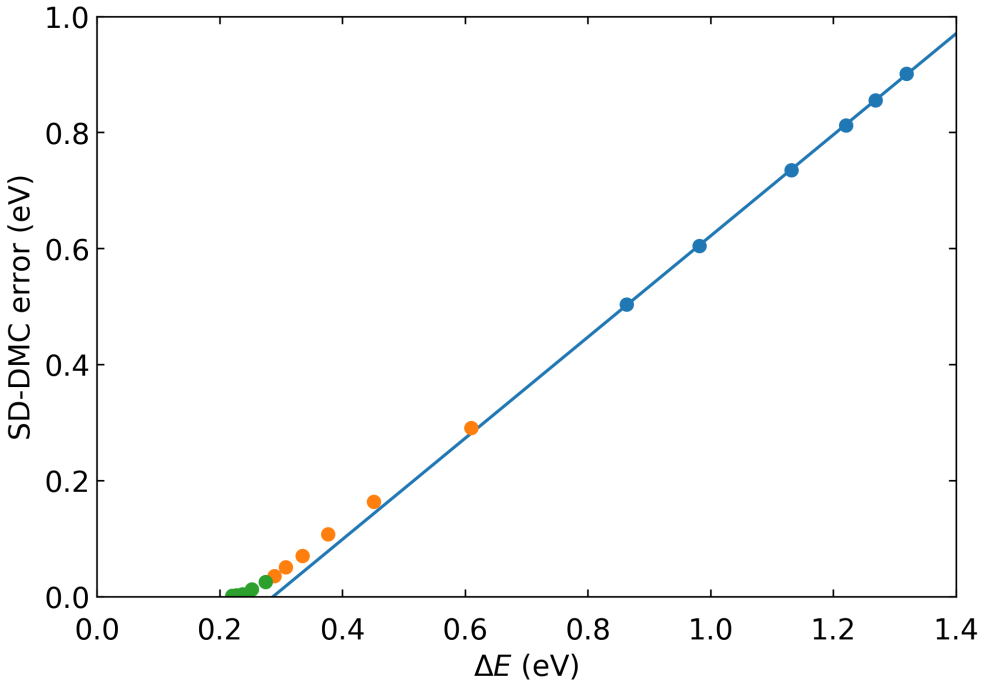

Figure 9: Error in the SD-DMC energy of H_4 (relative to the CAS(4,4)-DMC energy) as a function of ΔE, the HF-CAS(2,2) energy difference, for distances between the two H_2 molecules ranging from 1.27 to 3.0 Å.

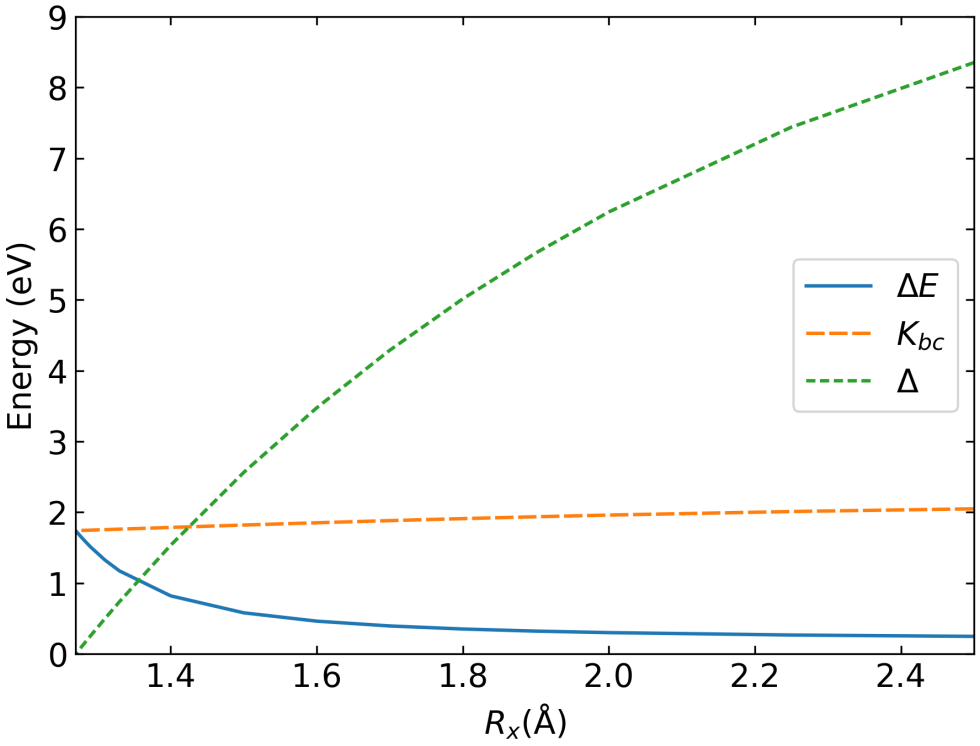

Figure 10: Values of K_{bc}, Δ, and ΔE for H_4 as a function of the distance between H_2 molecules. Results obtained using a STO-6G basis set.

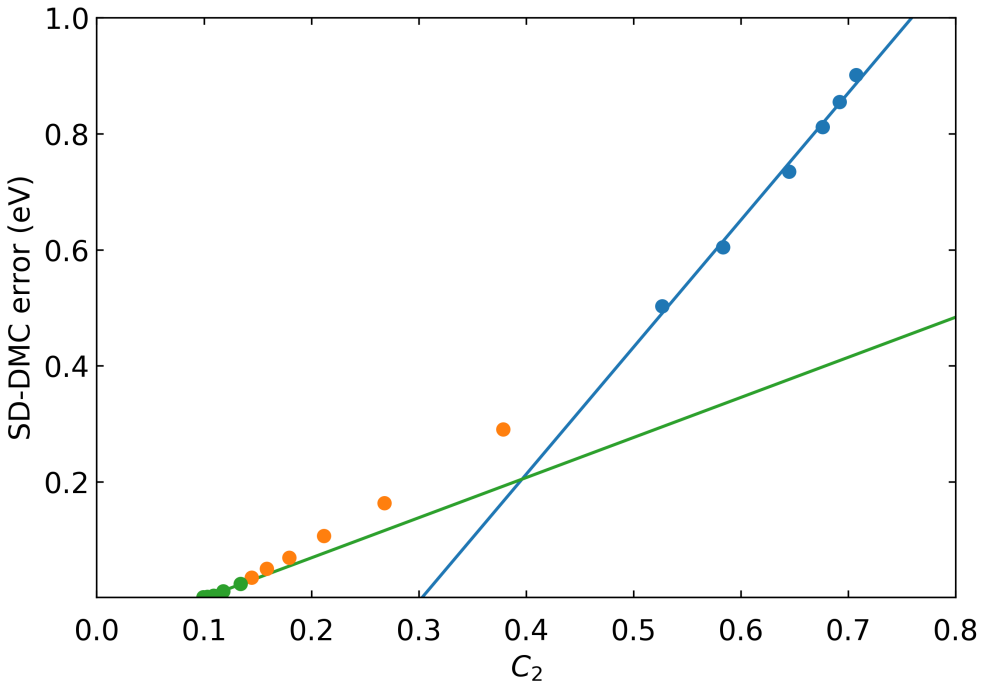

Figure 11: Error in the SD-DMC energy as a function of c_2, the coefficient of the $(1a_g)^2(1b_{2u})^2$ configuration from CAS(2,2) calculations for distances between the two H_2 molecules ranging from 1.27 to 3.0 Å. In each of the two limiting regimes (large and small Δ/K), the error in the SD-DMC energy depends approximately exponentially on c_2.

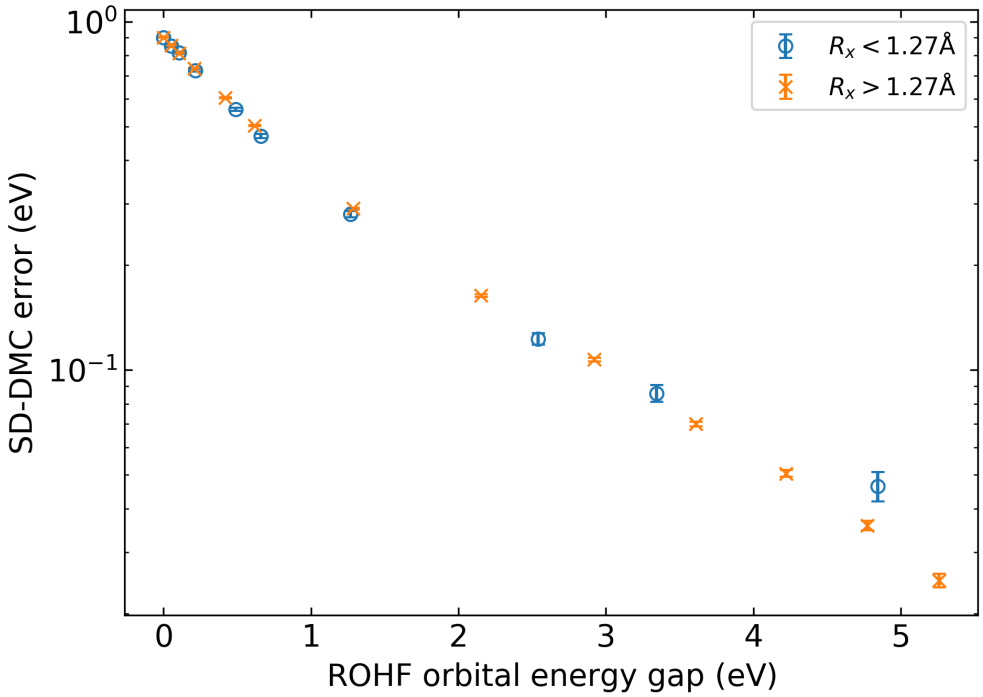

Figure 12: Error in the SD-DMC energy of H_4 vs. the orbital energy gap from ROHF/cc-pVQZ calculations on the lowest energy triplet state.

4.0 Quantum Monte Carlo Calculation of the Binding Energy of the Beryllium Dimer

The text and figures in this chapter have been reprinted from Deible, M. J.; Kessler, M.; Gasperich, K. E.; Jordan, K. D. Quantum Monte Carlo calculation of the binding energy of the beryllium dimer. *J. Chem. Phys.* **2015**, *143*, 084116, DOI: 10.1063/1.4929351, with the permission of AIP Publishing. The author's contribution to the work included performing single-determinant DMC calculations on the Be atom, creating/editing figures, and revising the manuscript.

4.1 Summary

The accurate calculation of the binding energy of the beryllium dimer is a challenging theoretical problem. In this study, the binding energy of Be_2 is calculated using the diffusion Monte Carlo (DMC) method, using single Slater determinant and multiconfigurational trial functions. DMC calculations using single-determinant trial wave functions of orbitals obtained from density functional theory calculations overestimate the binding energy, while DMC calculations using Hartree-Fock or CAS(4,8), complete active space trial functions significantly underestimate the binding energy. In order to obtain an accurate value of the binding energy of Be_2 from DMC calculations, it is necessary to employ trial functions that include excitations outside the valence space. Our best estimate DMC result for the binding energy of Be_2, obtained by using configuration interaction trial functions and extrapolating in the threshold for the configurations retained in the trial function, is $908\,cm^{-1}$, only slightly below the $935\,cm^{-1}$ value derived from experiment.

4.2 Introduction

The beryllium dimer has been the subject of numerous experimental and theoretical studies.[74,98–129] In 1984, Bondybey and English, using ro-vibrational data from near the bottom of the ground state $^1\Sigma_g^+$ potential energy curve of Be_2, deduced a value of $790 \pm 30\,cm^{-1}$ for the binding energy (here defined from the potential energy minimum, i.e., neglecting vibrational zero-point energy).[98–100] Based on rotational structure in the $v = 0$ level, Bondybey and English determined a bond length of 2.45 Å. More recently, Merritt and coworkers experimentally observed eleven vibrational levels of Be_2, allowing them to obtain a more refined estimate of $929.74\,cm^{-1}$ for the well depth.[101] This was subsequently revised to $934.9\,cm^{-1}$ upon further analysis of the experimental data.[102] Over the past few years, several electronic structure calculations have been reported that obtained well depths close to the recent experimental value.[116–125] The keys to the successful calculations are the use of large, flexible basis sets and the recovery of a large portion of the correlation energy including contributions from the 1*s* core orbitals. To illustrate the difficulty of calculating an accurate binding energy of Be_2, we note that a complete basis set limit coupled cluster singles plus doubles with perturbative triples [CCSD(T)] calculation including correlation of the 1s core electrons underestimates the binding energy by $224\,cm^{-1}$.[124] Moreover, basis functions beyond those included in the aug-cc-pCVQZ basis set[130,131] contribute $79\,cm^{-1}$ to the CCSD(T) value of the binding energy.[124]

In this study, we apply the diffusion Monte Carlo (DMC) method[5,25,26,28] to the Be dimer. The DMC method is capable of giving the exact ground state energy under the constraint of the fixed-node approximation,[29,30,82,90,91] which is required to maintain the fermionic nature of the wave function. The constraint is imposed by the use of a trial function often taken to be a single Slater determinant of Hartree-Fock (HF) or density functional theory (DFT) orbitals. If the nodal surface of the trial wave function were exact, then the DMC method, if run for a sufficient number of steps and extrapolated to zero time step, would give the exact ground state energy. It is generally assumed[29] that for weakly interacting dimers, the errors introduced by the use of single determinant trial functions to impose the fixed nodes largely cancel when the interaction energy is calculated by subtracting the sum of the energies of

the two monomers from that of the dimer, and this has been confirmed for systems such as the water dimer and the methane dimer.[127] However, it is not clear that this will be the case for weakly interacting species for which static correlation effects are important. The Be dimer is thus a particularly interesting test system, as the ground state wavefunction of Be has considerable $2s^2 \rightarrow 2p^2$ character. Indeed, all-electron DMC calculations on Be using a CAS(2,4), complete active space, trial function allowing for $2s^2 \rightarrow 2p^2$ mixing give a significantly lower total energy than do DMC calculations using a single Slater determinant trial function.[74,90,120] However, DMC calculations using a CAS(4,8) trial function for the dimer and a CAS(2,4) trial function for the atom considerably underestimate the binding of the dimer.[118] Harkless and Irikura[120] used a truncated CAS(4,8) space and Anderson and Goddard[128] used a generalized valence bond (GVB) trial function and each reported DMC values of the binding energy of Be_2 in good agreement with experiment. As will be discussed later in the manuscript, this is likely to be fortuitous. In the present study, we calculate the binding energy of Be_2 using the DMC method in conjunction with more flexible multiconfigurational trial functions than were employed in earlier studies.

4.3 Computational details

The experimental value of the equilibrium bond length, 2.453 603 Å,[101] was used for all calculations on the beryllium dimer. In the first set of calculations, single determinant trial functions were considered, with the orbitals being obtained from the HF approximation and from several DFT methods including the local density approximation (LDA), the Perdew-Burke-Ernzerhof (PBE)[132] and Becke-Lee-Yang-Parr (BLYP)[62,133] generalized gradient approximation (GGA) functionals, and the Becke3LYP,[60–63] PBE0,[134] and Becke half and half exchange plus LYP correlation (BH&HLYP)[135] hybrid functionals, which contain 20 %, 25 %, and 50 % exact exchange, respectively. In addition, a trial function comprised of a single Slater determinant of Brueckner orbitals[136,137] was considered. The cc-pVQZ-g $5s4p3d2f$ contracted Gaussian-type orbital basis set[130,131] was used to represent the orbitals in the single Slater determinant trial functions. Both cc-pVQZ-fg and cc-pVQZ-g basis sets

were used in generating the multiconfigurational trial functions. Here, -*fg* indicates that both the *f* and *g* functions were omitted from the basis set, while -*g* indicates that only the *g* functions were omitted.

DMC calculations were also carried out using multiconfigurational trial functions generated from CAS and configuration interaction (CI) calculations. For the beryllium dimer, both CAS(4,8) and CAS(4,16) trial functions were considered. The CAS(4,8) wave function allows all arrangements of the four valence electrons in the space of the molecular orbitals (MOs) derived predominantly from the 2*s* and 2*p* atomic orbitals (AOs). The CAS(4,16) wavefunction expands the active space to include the π_g, π_u, σ_g, σ_u molecular orbitals derived from the 3*s* and 3*p* atomic orbitals and has 816 configuration state functions (CSFs). The DMC calculations were carried out retaining all CSFs with coefficients greater than 0.001, 0.0025, 0.005, and 0.01 in magnitude, and these results were used to extrapolate the energies to the value for the full configuration space. The extrapolation is shown in Figure 13. With the 0.001 coefficient threshold, 341 CSFs are retained from the CAS(4,16) space. Truncations of the configuration space were carried out after the CAS (or CI) calculations using GAMESS but prior to the optimizations in the variational Monte Carlo (VMC) calculations.

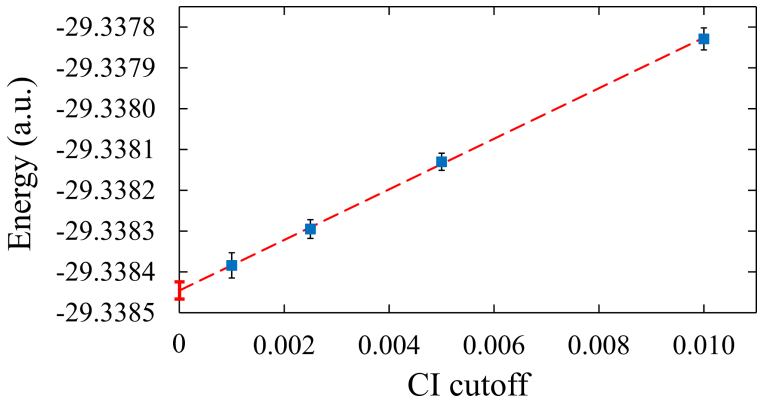

Figure 13: Extrapolation of the DMC energies of the beryllium dimer in the calculations using the CAS(4,16)/cc-pVQZ-*g* trial function, as described in the text. The dashed red line is a linear fit to the DMC energies (blue squares).

CI trial functions were generated by carrying out configuration interaction calculations, allowing for up to four electron excitations from the valence space into the full virtual space

and employing CAS(4,8) orbitals. Natural orbitals were then generated and used to carry out subsequent CI calculations allowing up to quadruple excitations in the space of all natural orbitals with occupations greater than 0.0001 in the first CI calculation (again keeping the $1\,\sigma_g$ and $1\,\sigma_u$ orbitals frozen). Trial functions with reduced configuration spaces were then generated by discarding configurations with coefficients above particular thresholds (0.01, 0.005, 0.0025, and 0.001). For the smallest threshold (0.001) 484 of 4500 CSFs were retained. For calculating the binding energy, a single plus double excitation CI (SDCI) calculation was carried out on the atom using CAS(2,8) orbitals and followed by a subsequent SDCI calculation using natural orbitals with occupations greater than 0.0001.

Each of the trial functions was combined with a Jastrow factor[32] with electron-electron, electron-nucleus, and electron-electron- nucleus terms. VMC calculations were used to optimize the Jastrow factors via energy minimization. For the multiconfigurational trial functions, the coefficients of the CSFs were optimized simultaneously with the parameters in the Jastrow function. The resulting trial functions, including the Jastrow factors, were then used to carry out DMC simulations using 40 000 to 50 000 walkers at a single time step of 0.001 a.u.. The correction scheme of Ma et al.[138] was used to account for the electron-nuclear cusps. For estimating statistical errors, the blocking procedure of Flyvbjerg and Petersen was used.[139] For one set of DMC calculations using the CAS(4,16) trial function, time steps of 0.0005, 0.003, and 0.005 a.u. were also used, allowing extrapolation of the energies to the zero time step limit. This extrapolation is shown in Figure 14.

The single determinant trial functions were generated using Gaussian09[67] and the multiconfigurational trial functions were generated using GAMESS.[96] The quantum Monte Carlo calculations were carried out using the CASINO[27] and QMCPACK[97] codes for the single determinant and multideterminant trial functions, respectively. QMCPACK was used for the latter calculations due to its implementation of an efficient algorithm for handling multideterminant trial functions.[140]

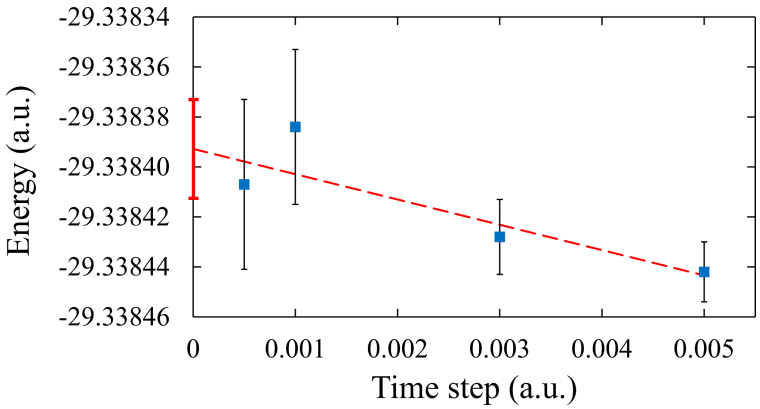

Figure 14: Extrapolation to zero time step of the DMC energy of the Be dimer at the equilibrium bond length of 2.453 603 Å. The calculations were based on the CAS(4,16)/cc-pVQZ-g trial function and used a 0.001 threshold on the CI coefficients.

4.4 Results

The results of the DMC calculations at the 0.001 a.u. time step are reported in Table 2. With the HF trial function, the DMC calculations give a binding energy of 724 cm^{-1}, significantly smaller than the experimental value of 935 cm^{-1}. On the other hand, the DMC calculations using trial functions employing LDA or GGA orbitals considerably overestimate the binding energy of Be_2. Significantly, improved agreement with experiment is obtained when using orbitals from hybrid functionals containing a component of exact exchange or from Brueckner CCSD calculations. Specifically, the DMC calculations using PBE0, BH&HLYP, and Brueckner orbitals give binding energies of 992, 966, and 955 cm^{-1}, respectively. Toulouse and Umrigar[126] obtained a binding energy of 1008 cm^{-1} from DMC calculations using single determinant trial functions but optimizing the orbitals and basis functions in the VMC step. For both Be and Be_2, regardless of the orbitals used, the DMC calculations using single determinant trial functions give energies considerably above the exact energies of these species, suggesting that the good agreement with experiment of these calculated binding energies is fortuitous. Support for this conjecture is provided by Figure 15, from which it

Table 2: Total energies of Be and Be_2 dissociation energy computed with DMC using various trial functions.

Trial function[a]	Total energy (a.u.) Be	Be_2	$D_e(cm^{-1})$
HF/QZ-*g*	-14.65730(4)	-29.31789(6)	724(21)
LDA/QZ-*g*	-14.65721(4)	-29.31977(7)	1174(25)
PBE/QZ-*g*	-14.65731(5)	-29.31960(8)	1094(26)
BLYP/QZ-*g*	-14.65725(4)	-29.31956(8)	1113(26)
B3LYP/QZ-*g*	-14.65727(3)	-29.31946(8)	1079(23)
PBE0/QZ-*g*	-14.65728(3)	-29.31907(8)	992(21)
BH&HLYP/QZ-*g*	-14.65726(5)	-29.31891(7)	966(26)
BD/QZ-*g*	-14.65718(4)	-29.31872(7)	955(24)
CAS(4,8)/QZ-*fg*[c]	-14.66723(1)	-29.33707(3)	573(8)
CAS(4,16)/QZ-*fg*[c]	-14.66730(1)	-29.33832(3)	819(8)
Ext. CAS(4,16)/QZ-*fg*	-14.66730(1)	-29.33841(2)	838(7)
CAS(4,16)/QZ-*g*[c]	-14.66727(2)	-29.33838(3)	845(8)
Ext. CAS(4,16)/QZ-*g*	-14.66727(2)	-29.33845(2)	857(9)
CI/QZ-*g*[c]	-14.66725(1)	-29.33848(2)	873(6)
Ext. CI/QZ-*g*	-14.66725(1)	-29.33864(2)	908(6)
Experimental[d]	-14.667356	-29.33897	934.9(4)

[a] QZ refers to the cc-pVQZ basis set. The "-*g*" and "-*fg*" indicate, respectively, that the *g* functions and *f* and *g* functions were omitted from the basis sets. Ext. refers to CAS and CI results extrapolated to the full configuration space for the active orbital list as described in the text.

[b] The DMC energies of the Be atom calculated using various single determinant trial functions should agree. The spread of the energies in the table is the result of statistical errors and the use of a finite (0.001 a.u.) time step.

[c] 0.001 threshold on CI coefficients for retained configurations.

[d] The experimental D_e value for Be_2 is from Ref. [101]. The non-relativistic energy of the Be atom is from Ref. [141]

is seen that the calculations that give binding energies closest to experiment do so because they give a higher energy for the dimer. It should be noted that all single determinant trial functions should give the same DMC energy of the Be atom.[142] The spread in the DMC energies of the Be atom calculated using different single determinant trial functions is only about 27 cm^{-1} with part of that being statistical and part being due to finite time step errors (i.e., using a time step of 0.001 a.u.).

As expected, based on earlier studies,[118,120] DMC calculations using valence-space CAS

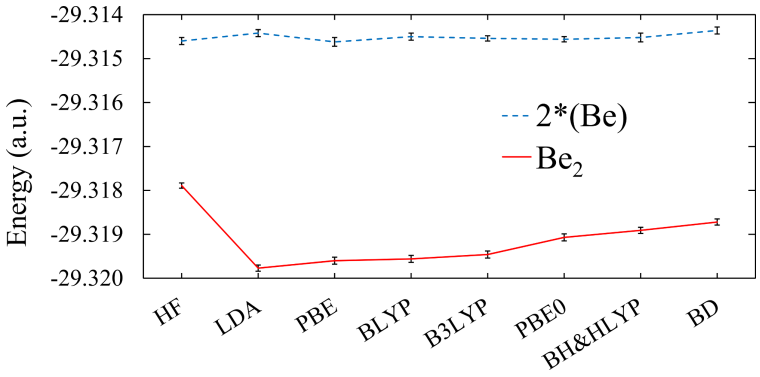

Figure 15: DMC energies of twice the beryllium atom and the dimer for several single-determinant trial wave functions.

trial functions give significantly lower energies for the Be atom and dimer than do the DMC calculations using the trial functions based on single Slater determinants. However, the DMC calculations using the CAS(4,8) trial function for the dimer and CAS(2,4) for the atom give a binding energy of only 573 cm^{-1}, which is even smaller than that obtained using HF trial functions. This indicates that use of valence space CAS trial functions does not result in a balanced treatment of the nodal surfaces of the atom and molecule. Most of the error is due to the inadequacy of the CAS(4,8) space in describing the nodal surfaces of the dimer since the DMC calculations on the atom using the CAS(2,4) trial function give an energy close to the current best estimate[141] (−14.667 228 vs. −14.667 356 a.u.). Expanding the CAS space to include also the MOs derived from the 3*s* and 3*p* AOs, giving CAS(2,8) for the atom and CAS(4,16) for the dimer, lowers the DMC energies of the atom and dimer, by 10 and 300 cm^{-1}, respectively, and results in a dimer binding energy of 845 cm^{-1}, at the 0.001 coefficient threshold and using the cc-pVQZ-*g* basis set. The corresponding binding energy obtained using the cc-pVQZ-*fg* basis set is 819 cm^{-1}, indicating that the nodal surface of Be_2 is slightly improved by including *f* functions in the basis set. Extrapolating these results along the sequence of coefficient cutoffs gives binding energies of 838 and 857 cm^{-1} for trial functions expanded in terms of the cc-pVQZ-*fg* and cc-pVQZ-*g* basis sets, respectively. The extrapolation to zero time step of the DMC/CAS(4,16) energies obtained with the 0.001

coefficient threshold when using the cc-pVQZ-g basis set gives a DMC binding energy of $849\,\mathrm{cm}^{-1}$ (see Figure 15) vs. the $845\,\mathrm{cm}^{-1}$ value obtained with the 0.001 a.u. time step. Thus, we conclude that the error due to the use of the finite time step is inconsequential for the calculation of the binding energy of the dimer.

The DMC calculations using CI trial functions with 0.001 coefficient cutoffs and the cc-pVQZ-g basis set yielded a dimer binding energy of $873\,\mathrm{cm}^{-1}$, while the corresponding result obtained by extrapolation to the full configuration space is $908\,\mathrm{cm}^{-1}$, which is only $27\,\mathrm{cm}^{-1}$ smaller than the experimental value of the binding energy. These results demonstrate that correlation effects involving configurations outside the CAS(8,16) space are important for describing the nodal surface of Be_2.

It should be noted that the SDTQ CI calculations using the cc-pVQZ-g basis set and freezing the $1s$ orbitals give a binding energy of only $601\,\mathrm{cm}^{-1}$, which is $334\,\mathrm{cm}^{-1}$ lower than the experimental value. About $70\,\mathrm{cm}^{-1}$ of the error in this result is due to the neglect of the correlation effects involving the core $1s$ orbitals,[125] while the remaining error is due to correlation effects that are not captured due to the basis set truncation. This underscores one of the major advantages of the DMC method, namely, that it achieves convergence with much smaller basis sets (for the trial functions) than required for traditional quantum chemistry methods.

4.5 Conclusions

In conclusion, the binding energy of the beryllium dimer has been calculated using the diffusion Monte Carlo method in conjunction with a wide variety of trial wave functions. Even DMC calculations with a trial wave function as flexible as CAS(4,16) considerably underestimate the binding energy of the beryllium dimer. CI trial functions allowing excitations from the valence space into the entire virtual space give a binding energy within $27\,\mathrm{cm}^{-1}$ of the experimental value. It is possible that this small remaining discrepancy from experiment is due to the neglect of excitations from the $1s$ orbitals in the trial functions used for the DMC calculations. Although DMC calculations using small configurational spaces

that give binding energies close to experiment have been reported for Be_2, they also give energies for the atom and dimer that are appreciably higher than those obtained using the CI trial functions employed here. Thus, the good agreement of the binding energy of Be_2 with the experimental value obtained with such small multiconfigurational trial function spaces is likely fortuitous. We believe that our findings are relevant for a wide range of other dimers, e.g., the benzene dimer, where there is appreciable configuration mixing in the wave functions of the monomers. In particular, achieving well converged binding energies for such systems is likely to require the use of multiconfigurational trial functions allowing for high-order excitations as well as excitations outside the valence space.

4.6 Acknowledgements

This research was supported by Grant No. CHE136234 from the National Science Foundation. The calculations were done on computers in the University of Pittsburgh's Center for Simulation and Modeling. We acknowledge valuable discussions with M. Morales.

5.0 Quantum Package 2.0: An Open-Source Determinant-Driven Suite of Programs

The text and figures in this chapter have been reprinted with permission from Garniron, Y.; Applencourt, T.; Gasperich, K.; Benali, A.; Ferté, A.; Paquier, J.; Pradines, B.; Assaraf, R.; Reinhardt, P.; Toulouse, J.; Barbaresco, P.; Renon, N.; David, G.; Malrieu, J.-P.; Véril, M.; Caffarel, M.; Loos, P.-F.; Giner, E.; Scemama, A. Quantum Package 2.0: An Open-Source Determinant-Driven Suite of Programs. *J. Chem. Theory Comput.* **2019**, *15*, 3591–3609, DOI: 10.1021/acs.jctc.9b00176. The author contributed to the implementation of the algorithm for generating spin-adapted sets of determinants in addition to making minor revisions of the manuscript.

5.1 Summary

Quantum chemistry is a discipline which relies heavily on very expensive numerical computations. The scaling of correlated wave function methods lies, in their standard implementation, between $\mathcal{O}(N^5)$ and $\mathcal{O}(e^N)$, where N is proportional to the system size. Therefore, performing accurate calculations on chemically meaningful systems requires i) approximations that can lower the computational scaling, and ii) efficient implementations that take advantage of modern massively parallel architectures. Quantum Package is an open-source programming environment for quantum chemistry specially designed for wave function methods. Its main goal is the development of determinant-driven selected configuration interaction (sCI) methods and multi-reference second-order perturbation theory (PT2). The determinant-driven framework allows the programmer to include any arbitrary set of determinants in the reference space, hence providing greater methodological freedom. The sCI method implemented in Quantum Package is based on the CIPSI (Configuration Interaction using a Perturbative Selection made Iteratively) algorithm which complements the variational sCI energy with a PT2 correction. Additional external plugins have been

recently added to perform calculations with multireference coupled cluster theory and range-separated density-functional theory. All the programs are developed with the IRPF90 code generator, which simplifies collaborative work and the development of new features. Quantum Package strives to allow easy implementation and experimentation of new methods, while making parallel computation as simple and efficient as possible on modern supercomputer architectures. Currently, the code enables, routinely, to realize runs on roughly 2000 CPU cores, with tens of millions of determinants in the reference space. Moreover, we have been able to push up to 12 288 cores in order to test its parallel efficiency. In the present manuscript, we also introduce some key new developments: i) a renormalized second-order perturbative correction for efficient extrapolation to the full CI limit, and ii) a stochastic version of the CIPSI selection performed simultaneously to the PT2 calculation at no extra cost.

5.2 Introduction

In 1965, Gordon Moore predicted that the number of transistors in an integrated circuit would double about every two years (the so-called Moore's law).[143] Rapidly, this "law" was interpreted as an expected two-fold increase in performance every 18 months. This became an industrial goal. The development of today's most popular electronic structure codes was initiated in the 1990's (or even before). At that time, the increase of computational power from one supercomputer generation to the next was mostly driven by an increase of processors' frequency. Indeed, the amount of random access memory was small, the time to access data from disk was slow, and the energy consumption of the most powerful computer was 236 kW, hence far from being an economical concern.[144] At the very beginning of the 21st century, having increased continuously, both the number of processors and their frequency raised the supercomputer power consumption by two orders of magnitude, inflating accordingly the electricity bill. The only way to slow down this frenetic growth of power consumption while keeping alive Moore's dream was to freeze the processor's frequency (between 1 and 4 GHz), and increase the number of CPU cores. The consequence of such a choice was that *"free*

lunch" was over: the programmers now had to parallelize their programs to make them run faster.[145] At the same time, computer scientists realized that the increase of performance in memory access was slower than the increase in computational power,[146] and that the floating-point operation (or flop) count would soon stop being the bottleneck. From now on, data movement would be the main concern. This paradigm shift was named the *memory wall*. Moore's law is definitely near the end of its life.[147]

The traditional sequential algorithms of quantum chemistry are currently being redesigned and replaced by parallel equivalents by multiple groups around the world.[11–14,148–156] This has obviously a significant influence on methodological developments. The most iconic example of this move towards parallel-friendly methods is the recently developed *full configuration interaction quantum Monte Carlo* (FCIQMC) method by Alavi and coworkers.[11] FCIQMC can be interpreted as a Monte Carlo equivalent of older selected configuration interaction (sCI) algorithms[8–10,14–16,18,19,156–185] such as CIPSI (Configuration Interaction using a Perturbative Selection made Iteratively),[10] that are iterative and thus *a priori* not well adapted to massively parallel architecture. As we shall see here, things turn out differently, and the focus of the present article is to show that sCI methods can be made efficient on modern massively parallel supercomputers.

Quantum Package[186] is an open-source suite of wave function quantum chemistry methods mainly developed at the *Laboratoire de Chimie et Physique Quantiques* (LCPQ) in Toulouse (France), and the *Laboratoire de Chimie Théorique* (LCT) in Paris. Its source code is freely available on GitHub at the following address: https://github.com/QuantumPackage/qp2. Quantum Package strives to allow easy implementation and experimentation of new methods, while making parallel computation as simple and efficient as possible. Accordingly, the initial choice of Quantum Package was to go towards *determinant-driven* algorithms. Assuming a wave function expressed as a linear combination of determinants, a determinant-driven algorithm essentially implies that the outermost loop runs over determinants. On the other hand, more traditional *integral-driven* algorithms have their outermost loop running on the two-electron integrals appearing in the expression of the matrix elements in the determinant basis (see Sec. 5.3.2). Determinant-driven algorithms allow more flexibility than their integral-driven counterparts,[187] but they have been known for years to be less efficient

than their integral-driven variant for solving electronic structure problems. In high-precision calculations, the number of determinants is larger than the number of integrals, justifying the integral-driven choice. However, today's programming standards impose parallelism, and if determinant-driven calculations prove to be better adapted to parallelism, such methods could regain popularity. More conventional approaches have also been very successfully parallelized: CCSD(T),[188,189] DMRG,[190] GW,[191] QMC,[192–194] and many others.

Quantum Package was used in numerous applications, in particular to obtain reference ground-state energies[165–169,195] as well as excitation energies[19,181,196] for atomic and molecular systems. For example, in Ref. [19], Quantum Package has been used to compute more than hundred very accurate transition energies for states of various characters (valence, Rydberg, $n \to \pi^*$, $\pi \to \pi^*$, singlet, triplet, ...) in 18 small molecules. The high quality and compactness of the CIPSI wave function was also used for quantum Monte Carlo calculations to characterize the ground state of the water and the FeS molecules,[169,173] and obtained highly accurate excitation energies.[174,197,198] Of course, the technical considerations were not the main concern of the different articles that were produced. Because the present work focused on the actual implementation of the methods at least as much as on the theory behind them, this article is a perfect opportunity to discuss in depth their implementation.

This manuscript is organized as follows. In Sec. 5.3, we briefly describe the main computational methods implemented in Quantum Package as well as newly developed methods and extrapolation techniques. Section 5.4 deals with their implementation. In particular, Sec. 5.4.1 discusses the computation of the Hamiltonian matrix elements using determinant-driven algorithms, while Sec. 5.4.3 focuses on the acceleration of the Davidson diagonalization, a pivotal point of sCI methods. In Sec. 5.4.4, we focus on the determinant selection step used to build compact wave functions. In a nutshell, the principle is to incrementally build a reference wave function by scavenging its external space for determinants that interact with it. To make this step more affordable, we designed a new stochastic scheme which selects *on the fly* the more important determinants while the second-order perturbative (PT2) energy is computed using a hybrid stochastic-deterministic scheme.[156] Therefore, the selection part of this new stochastic CIPSI selection is virtually free as long as one is interested in the second-order perturbative correction, which is crucial in many cases in order

to obtain near *full configuration interaction* (FCI) results. Section 5.5 briefly explains how we produce spin-adapted wave functions, and Sec. 5.6 describes parallelism within QUANTUM PACKAGE. The efficiency of the present algorithms is demonstrated in Sec. 5.7.3 where illustrative calculations and parallel speedups are reported. Finally, Sec. 5.8 discusses the development philosophy of QUANTUM PACKAGE as well as other relevant technical details. Unless otherwise stated, atomic units are used throughout.

5.3 Methods

5.3.1 Generalities

The correlation energy is defined as[199]

$$E_{\mathrm{c}} = E_{\mathrm{exact}} - E_{\mathrm{HF}}, \tag{5.1}$$

where E_{exact} and E_{HF} are, respectively, the exact (non-relativistic) energy and the Hartree-Fock (HF) energy in a complete (one-electron) basis set.

To include electron correlation effects, the wave function associated with the kth electronic state, $|\Psi_k\rangle$, may be expanded in the set of all possible N-electron Slater determinants, $|I\rangle$, built by placing $N_\uparrow$ spin-up electrons in N_{orb} orbitals and $N_\downarrow$ spin-down electrons in N_{orb} orbitals (where $N = N_\uparrow + N_\downarrow$). These so-called molecular orbitals (MOs) are defined as linear combinations of atomic orbitals (AOs)

$$\phi_p(\mathbf{r}) = \sum_{\mu}^{N_{\mathrm{orb}}} C_{\mu p} \chi_\mu(\mathbf{r}). \tag{5.2}$$

Note that the MOs are assumed to be real valued in the context of this work. The eigenvectors of the Hamiltonian $\hat{H}$ are consequently expressed as linear combinations of Slater determinants, i.e.,

$$|\Psi_k\rangle = \sum_{I}^{N_{\mathrm{det}}} c_{Ik} |I\rangle , \tag{5.3}$$

where N_{det} is the number of determinants. For sake of conciseness, we will restrict the discussion to the ground state (i.e. $k = 0$) and drop the subscript k accordingly. Solving the

eigenvalue problem in this basis is referred to as FCI and yields, for a given basis set, the exact solution of the Schrödinger equation. Unfortunately, FCI is usually computationally intractable because of its exponential scaling with the size of the system.

5.3.2 Matrix elements of the Hamiltonian

In the N-electron basis of Slater determinants, one expects the matrix elements of $\hat{H}$ to be integrals over $3N$ dimensions. However, given the two-electron nature of the Hamiltonian, and because the MOs are orthonormal, Slater determinants that differ by more than two spinorbitals yield a zero matrix element. The remaining elements can be expressed as sums of integrals over one- or two-electron coordinates, which can be computed at a reasonable cost. These simplifications are known as Slater-Condon's rules, and reads

$$\langle I|\hat{H}|I\rangle = \sum_{i\in|I\rangle}(i|\hat{h}|i) + \frac{1}{2}\sum_{(i,j)\in|I\rangle}(ii||jj), \tag{5.4a}$$

$$\langle I|\hat{H}|I_p^r\rangle = (p|\hat{h}|r) + \sum_{i\in|I\rangle}(pr||ii), \tag{5.4b}$$

$$\langle I|\hat{H}|I_{pq}^{rs}\rangle = (pr||qs), \tag{5.4c}$$

where $\hat{h}$ is the one-electron part of the Hamiltonian (including kinetic energy and electron-nucleus attraction operators),

$$(p|\hat{h}|q) = \int \phi_p(\mathbf{r})\hat{h}(\mathbf{r})\phi_q(\mathbf{r})d\mathbf{r} \tag{5.5}$$

are one-electron integrals, $i \in |I\rangle$ means that ϕ_i belongs to the Slater determinant $|I\rangle$, $|I_p^r\rangle$ and $|I_{pq}^{rs}\rangle$ are determinants obtained from $|I\rangle$ by substituting orbitals ϕ_p by ϕ_r, and ϕ_p and ϕ_q by ϕ_r and ϕ_s, respectively,

$$(pq|rs) = \iint \phi_p(\mathbf{r}_1)\phi_q(\mathbf{r}_1)r_{12}^{-1}\phi_r(\mathbf{r}_2)\phi_s(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \tag{5.6}$$

are two-electron electron repulsion integrals (ERIs), $r_{12}^{-1} = |\mathbf{r}_1 - \mathbf{r}_2|^{-1}$ is the Coulomb operator, and $(pq||rs) = (pq|rs) - (ps|rq)$ are the usual antisymmetrized two-electron integrals.

Within the HF method, Roothaan's equations allow to solve the problem in the AO basis.[200] In this context, one needs to compute the $\mathcal{O}(N_{\text{orb}}^4)$ two-electron integrals $(\mu\nu|\lambda\sigma)$ over

the AO basis. Thanks to a large effort in algorithmic development and implementation,[201–208] these integrals can now be computed very fast on modern computers. However, with post-HF methods, the computation of the two-electron integrals is a potential bottleneck. Indeed, when computing matrix elements of the Hamiltonian in the basis of Slater determinants, ERIs over MOs are required. Using Eq. (5.2), the cost of computing a single integral $(pq|rs)$ scales as $\mathcal{O}(N_{\text{orb}}^4)$. A naive computation of all integrals in the MO basis would cost $\mathcal{O}(N_{\text{orb}}^8)$. Fortunately, computing all of them can be scaled down to $\mathcal{O}(N_{\text{orb}}^5)$ by transforming the indices one by one.[209] This step is known as the four-index integral transformation. In addition to being very costly, this step is hard to parallelize in a distributed way, because it requires multiple collective communications.[210–213] However, techniques such as density fitting (also called the resolution of the identity),[214–216] low-rank approximations,[217–220] or the combination of both[221] are now routinely employed to overcome the computational and storage bottlenecks.

5.3.3 Selected CI methods

The sCI methods rely on the same principle as the usual configuration interaction (CI) approaches, except that determinants are not chosen *a priori* based on occupation or excitation criteria, but selected among the entire set of determinants based on their estimated contribution to the FCI wave function. Indeed, it has been noticed long ago that, even inside a predefined subspace of determinants, only a small number of them significantly contributes.[164,222] Therefore, an *on-the-fly* selection of determinants is a rather natural idea that has been proposed in the late 1960's by Bender and Davidson[8] as well as Whitten and Hackmeyer.[9] sCI methods are still very much under active development. The main advantage of sCI methods is that no *a priori* assumption is made on the type of electronic correlation. Therefore, at the price of a brute force calculation, a sCI calculation is less biased by the user's appreciation of the problem's complexity.

The approach that we have implemented in Quantum Package is based on the CIPSI algorithm developed by Huron, Rancurel and Malrieu in 1973,[10] that iteratively selects external determinants $|\alpha\rangle$ — determinants which are not present in the (reference or variational)

zeroth-order wave function

$$|\Psi^{(0)}\rangle = \sum_I c_I \, |I\rangle \tag{5.7}$$

at a given iteration — using a perturbative criterion

$$e_\alpha^{(2)} = \frac{\langle \Psi^{(0)}|\hat{H}|\alpha\rangle^2}{E^{(0)} - \langle \alpha|\hat{H}|\alpha\rangle}, \tag{5.8}$$

where

$$E^{(0)} = \frac{\langle \Psi^{(0)}|\hat{H}|\Psi^{(0)}\rangle}{\langle \Psi^{(0)}|\Psi^{(0)}\rangle} \geq E_{\text{FCI}} \tag{5.9}$$

is the zeroth-order (variational) energy, and $e_\alpha^{(2)}$ the (second-order) estimated gain in correlation energy that would be brought by the inclusion of $|\alpha\rangle$. The second-order perturbative correction

$$E^{(2)} = \sum_\alpha e_\alpha^{(2)} = \sum_\alpha \frac{\langle \alpha|\hat{H}|\Psi^{(0)}\rangle^2}{E^{(0)} - \langle \alpha|\hat{H}|\alpha\rangle} \tag{5.10}$$

is an estimate of the total missing correlation energy, i.e., $E^{(2)} \approx E_{\text{FCI}} - E^{(0)}$, for large enough expansions.

Let us emphasize that sCI methods can be applied to any determinant space. Although presented here for the FCI space, it can be trivially generalized to a complete active space (CAS), but also to standard CI spaces such as CIS, CISD or MR-CISD. The only required modification is to set to zero the contributions associated with the determinants which do not belong to the target space.

There is, however, a computational downside to sCI methods. In conventional CI methods, the rule by which determinants are selected is known *a priori*, and therefore, one can map a particular determinant to some row or column indices.[223] As a consequence, it can be systematically determined to which matrix element of $\hat{H}$ a two-electron integral contributes. This allows for the implementation of so-called *integral-driven* methods that work essentially by iterating over integrals. On the contrary, in (most) sCI methods, the determinants are selected *a posteriori*, and an explicit list has to be maintained as there is no immediate way to know whether or not a determinant has been selected. Consequently, we must rely on the so-called *determinant-driven* approach in which iterations are performed over determinants rather than integrals. This can be a lot more expensive, since the number of determinants N_{det} is typically much larger than the number of integrals. The number of determinants

scales as $\mathcal{O}(N_{\text{orb}}!)$ while the number of integrals scales (formally) as $\mathcal{O}(N_{\text{orb}}^4)$. What makes sCI calculations possible in practice is that sCI methods generate relatively compact wave functions, i.e. wave functions where N_{det} is much smaller (by orders of magnitude) than the size of the FCI space. Furthermore, determinant-driven methods require an effective way to compare determinants in order to extract the corresponding excitation operators, and a way to rapidly fetch the associated integrals involved, as described in Sec. 5.4.1.

Because of this high computational cost, approximations have been proposed.[157] Recently, the semi-stochastic heat-bath configuration interaction (SHCI) algorithm has taken further the idea of a more approximate but extremely cheap selection.[14,170,180] Compared to CIPSI, the selection criterion is simplified to

$$e_\alpha^{\text{SHCI}} = \max_I \left(\left| c_I \langle I|\hat{H}|\alpha\rangle \right| \right). \tag{5.11}$$

This algorithmically allows for an extremely fast selection of doubly-excited determinants by an integral-driven approach. Nonetheless, the bottlenecks of the SHCI are the diagonalization step and the computation of $E^{(2)}$, which remain determinant driven.

As mentioned above, FCIQMC is an alternative approach of stochastic nature recently developed in Alavi's group,[11–13] where signed walkers spawn from one determinant to connected ones, with a probability that is a function of the associated matrix element. The average proportion of walkers on a determinant converges to its coefficient in the FCI wave function. A more "brute force" approach is the purely stochastic selection of Monte Carlo CI (MCCI),[224,225] where determinants are randomly added to the zeroth-order wave function. After diagonalization, the determinants of smaller coefficient are removed, and new random determinants are added.

A number of other variants exist but are not detailed here.[8–10,14–16,18,19,156–181,226] Although these various approaches appear under diverse acronyms, most of them rely on the very same idea of selecting determinants iteratively according to their contribution to the wave function or energy.

5.3.4 Extrapolation techniques

5.3.4.1 Usual extrapolation procedure In order to extrapolate the sCI results to the FCI limit, we have adopted the method recently proposed by Holmes, Umrigar and Sharma[171] in the context of the SHCI method.[14,170,171] It consists of extrapolating the sCI energy, $E^{(0)}$, as a function of the second-order Epstein-Nesbet energy, $E^{(2)}$, which is an estimate of the truncation error in the sCI algorithm, i.e $E^{(2)} \approx E_{\text{FCI}} - E^{(0)}$.[10] When $E^{(2)} = 0$, the FCI limit has effectively been reached. This extrapolation procedure has been shown to be robust, even for challenging chemical situations.[14,18,19,171–174,181] Below, we propose an improved extrapolation scheme which *renormalizes* the second-order perturbative correction.

5.3.4.2 Renormalized PT2 At a given sCI iteration, the sCI+PT2 energy is given by

$$E = E^{(0)} + E^{(2)}, \tag{5.12}$$

where $E^{(0)}$ and $E^{(2)}$ are given by Eqs. (5.9) and (5.10), respectively. Let us introduce the following energy-dependent second-order self-energy

$$\Sigma^{(2)}[E] = \sum_{\alpha} \frac{\langle \alpha | \hat{H} | \Psi^{(0)} \rangle^2}{E - \langle \alpha | \hat{H} | \alpha \rangle}. \tag{5.13}$$

Obviously, we have $\Sigma^{(2)}[E^{(0)}] = E^{(2)}$. Now, let us consider the more general problem, which is somewhat related to Brillouin-Wigner perturbation theory, where we have

$$E = E^{(0)} + \Sigma^{(2)}[E], \tag{5.14}$$

and assume that $\Sigma^{(2)}[E]$ behaves linearly for $E \approx E^{(0)}$, i.e.,

$$\Sigma^{(2)}[E] \approx \Sigma^{(2)}[E^{(0)}] + (E - E^{(0)}) \left. \frac{\partial \Sigma^{(2)}[E]}{\partial E} \right|_{E=E^{(0)}}. \tag{5.15}$$

This linear behavior is corroborated by the findings of Nitzsche and Davidson.[227] Substituting Eq. (5.15) into (5.14) yields

$$\begin{aligned} E &= E^{(0)} + \Sigma^{(2)}[E^{(0)}] + (E - E^{(0)}) \left. \frac{\partial \Sigma^{(2)}[E]}{\partial E} \right|_{E=E^{(0)}} \\ &= E^{(0)} + Z\, E^{(2)}, \end{aligned} \tag{5.16}$$

where the renormalization factor is

$$Z = \left[1 - \left.\frac{\partial \Sigma^{(2)}[E]}{\partial E}\right|_{E=E^{(0)}}\right]^{-1}, \tag{5.17}$$

and

$$\left.\frac{\partial \Sigma^{(2)}[E]}{\partial E}\right|_{E=E^{(0)}} = -\sum_{\alpha} \frac{\langle \alpha|\hat{H}|\Psi^{(0)}\rangle^2}{(E^{(0)} - \langle \alpha|\hat{H}|\alpha\rangle)^2} < 0. \tag{5.18}$$

Therefore, the renormalization factor fulfills the condition $0 \leq Z \leq 1$, and its actual computation does not involve any additional cost when computed alongside $E^{(2)}$ as they involve the same quantities. This renormalization procedure of the second-order correction, that we have named rPT2, bears obvious similarities with the computation of quasiparticle energies within the G_0W_0 method.[228–231] Practically, the effect of rPT2 is to damp the value of $E^{(2)}$ for small wave functions. Indeed, when N_{det} is small, the sum $E^{(0)} + E^{(2)}$ usually overestimates (in magnitude) the FCI energy, yielding a pronounced non-linear behavior of the sCI+PT2 energy. Consequently, by computing instead the (renormalized) energy $E^{(0)} + Z\,E^{(2)}$, one observes a much more linear behavior of the energy, hence an easier extrapolation to the FCI limit. Its practical usefulness is illustrated in Sec. 5.7.2.

5.4 Implementation

In this section, we give an overview of the implementation of the various methods present in Quantum Package. The implementation of the crucial algorithms is explained in detail in the PhD thesis of Dr Y. Garniron[232] as well as in the Appendix of the present manuscript.

5.4.1 Determinant-driven computation of the matrix elements

For performance sake, it is vital that some basic operations are done efficiently and, notably, the computation of the Hamiltonian matrix elements. This raises some questions about the data structures chosen to represent the two-electron integrals and determinants, as well as their consequences from an algorithmic point of view. This section is going to address these questions by going through the basic concepts of our determinant-driven approach.

5.4.1.1 Storage of the two-electron integrals In QUANTUM PACKAGE, the two-electron integrals are kept in memory because they require a fast random access. Considering the large number of two-electron integrals, a hash table is the natural choice allowing the storage of only non-zero values with a data retrieval in near constant time.[233] However, standard hashing algorithms tend to shuffle data to limit the probability of collisions. Here, we favor data locality using the hash function given in Algorithm 1. This hash function i) returns the same value for all keys related by permutation symmetry, ii) keeps some locality in the storage of data, and iii) can be evaluated in 10 CPU cycles (estimated with MAQAO[234]) if the integer divisions by two are replaced by right bit shift instructions.

Function `HASH`(i,j,k,l)**:** `/* Hash function for two-electron integrals */`

Data: i,j,k,l are the orbital indices

Result: The corresponding hash

$p \leftarrow \min(i,k)$;

$r \leftarrow \max(i,k)$;

$t \leftarrow p + r(r-1)/2$;

$q \leftarrow \min(j,l)$;

$s \leftarrow \max(j,l)$;

$u \leftarrow q + s(s-1)/2$;

$v \leftarrow \min(t,u)$;

$w \leftarrow \max(t,u)$;

return $v + w(w-1)/2$;

Algorithm 1: Hash function that maps any orbital quartet (i,j,k,l) related by permutation symmetry to a unique integer.

The hash table is such that each bucket can potentially store 2^{15} consecutive key-value pairs. The 15 least significant bits of the hash value are removed to give the bucket index $[i_{\text{bucket}} = \lfloor \text{hash}(i,j,k,l)/2^{15} \rfloor]$, and only those 15 bits need to be stored in the bucket for the key storage $[\text{hash}(i,j,k,l) \mod 2^{16}]$. Hence, the key storage only requires two bytes per key, and they are sorted in increasing order, enabling a binary search within the bucket. The key search is always fast since the binary search is bounded by 15 misses and the maximum size of the key array is 64 kiB, the typical size of the L1 cache. The efficiency of the integral

storage is illustrated in Section 5.12.1.

5.4.2 Internal representation of determinants

Determinants can be conveniently written as a string of creation operators applied to the vacuum state $|\rangle$, e.g., $a_i^\dagger a_j^\dagger a_k^\dagger|\rangle = |I\rangle$. Because of the fermionic nature of electrons, a permutation of two contiguous creation operators results in a sign change $a_j^\dagger a_i^\dagger = -a_i^\dagger a_j^\dagger$, which makes their ordering relevant, e.g., $a_j^\dagger a_i^\dagger a_k^\dagger|\rangle = -|I\rangle$. A determinant can be broken down into two pieces of information: i) a set of creation operators corresponding to the set of occupied spinorbitals in the determinant, and ii) an ordering of the creation operators responsible for the sign of the determinant, known as *phase factor*. Once an ordering operator $\hat{\mathcal{O}}$ is chosen and applied to all determinants, the phase factor may simply be included in the CI coefficient.

The determinants are built using the following order: i) spin-up ($\uparrow$) spinorbitals are placed before spin-down ($\downarrow$) spinorbitals, as in the Waller-Hartree double determinant representation[235] $\hat{\mathcal{O}}\,|I\rangle = \hat{I}|\rangle = \hat{I}_\uparrow \hat{I}_\downarrow|\rangle$, and ii) within each operator $\hat{I}_\uparrow$ and $\hat{I}_\downarrow$, the creation operators are sorted by increasing indices. For instance, let us consider the determinant $|J\rangle = a_j^\dagger a_k^\dagger a_{\bar{i}}^\dagger a_i^\dagger|\rangle$ built from the set of spinorbitals $\{i_\uparrow, j_\uparrow, k_\uparrow, i_\downarrow\}$ with $i < j < k$. If we happen to encounter such a determinant, our choice of representation imposes to consider its re-ordered expression $\hat{\mathcal{O}}\,|J\rangle = -a_i^\dagger a_j^\dagger a_k^\dagger a_{\bar{i}}^\dagger|\rangle = -|J\rangle$, and the phase factor must be handled.

The indices of the creation operators (or equivalently the spinorbital occupations), are stored using the so-called *bitstring* encoding. A bitstring is an array of bits; typically, the 64-bit binary representation of an integer is a bitstring of size 64. Quite simply, the idea is to map each spinorbital to a single bit with value set to its occupation number. In other words, 0 and 1 are associated with the *unoccupied* and *occupied* states, respectively. Additional information about the internal representation of determinants can be found in Appendix 5.12.2.

5.4.3 Davidson diagonalization

Finding the lowest root(s) of the Hamiltonian is a necessary step in CI methods. Standard diagonalization algorithms scale as $\mathcal{O}(N_{\text{det}}^3)$ and $\mathcal{O}(N_{\text{det}}^2)$ in terms of computation and storage, respectively. Hence, their cost is prohibitive as N_{det} is usually, at least, of the order of few millions. Fortunately, not all the spectrum of $\hat{H}$ is required: only the first few lowest eigenstates are of interest. The Davidson diagonalization[236–240] is an iterative algorithm which aims at extracting the first N_{states} lowest eigenstates of a large matrix. This algorithm reduces the cost of both the computation and storage to $\mathcal{O}(N_{\text{states}}N_{\text{det}}^2)$ and $\mathcal{O}(N_{\text{states}}N_{\text{det}})$, respectively. It is presented as Algorithm 2 and further details about the present Davidson algorithm implementation are gathered in Appendix 5.12.3.

5.4.4 CIPSI selection and PT2 energy

5.4.4.1 The basic algorithm The largest amount of work for this second version of Quantum Package has been devoted to the improvement of the CIPSI algorithm implementation.[241] As briefly described in Sec. 5.3, this is an iterative selection algorithm, where determinants are added to the reference wave function according to a perturbative criterion.

The nth CIPSI iteration can be described as follows:

1. The zeroth-order (reference or variational) wave function

$$|\Psi^{(0)}\rangle = \sum_{I \in \mathcal{I}_n} c_I \, |I\rangle \tag{5.19}$$

   is defined over a set of determinants $\mathcal{I}_n$ — characterized as *internal* determinants — from which the lowest eigenvector of $\hat{H}$ are obtained.
2. For all *external* determinants $|\alpha\rangle \notin \mathcal{I}_n$ but connected to $\mathcal{I}_n$, i.e., $\langle\Psi^{(0)}|\hat{H}|\alpha\rangle \neq 0$, we compute the individual perturbative contribution $e_\alpha^{(2)}$ given by Eq. (5.8). This set of external determinants is labeled $\mathcal{A}_n$.
3. Summing the contributions of all the external determinants $\alpha \in \mathcal{A}_n$ gives the second-order perturbative correction provided by Eq. (5.10) and the FCI energy can be estimated as $E_{\text{FCI}} \approx E^{(0)} + E^{(2)}$.

Function `DAVIDSON_DIAG`(N_{states}, $\mathbf{U}$)**:**

Data: N_{states}: Number of requested states

Data: N_{det}: Number of determinants

Data: $\mathbf{U}$: Guess vectors, $N_{\text{det}} \times N_{\text{states}}$

Result: N_{states} lowest eigenvalues eigenvectors of $\mathbf{H}$

`converged` $\leftarrow$ `FALSE` ;

while $\neg$ *converged* **do**

Gram-Schmidt orthonormalization of $\mathbf{U}$;

$\mathbf{W} \leftarrow \mathbf{H}\,\mathbf{U}$;

$\mathbf{h} \leftarrow \mathbf{U}^{\dagger}\,\mathbf{W}$;

Diagonalize $\mathbf{h}$: eigenvalues E and eigenvectors $\mathbf{y}$;

$\mathbf{U}' \leftarrow \mathbf{U}\,\mathbf{y}$;

$\mathbf{W}' \leftarrow \mathbf{W}\,\mathbf{y}$;

for $k \leftarrow 1, N_{states}$ **do**

for $i \leftarrow 1, N_{det}$ **do**

$\mathbf{R}_{ik} \leftarrow \frac{E_k \mathbf{U}'_{ik} - \mathbf{W}'_{ik}}{\mathbf{H}_{ii} - E_k}$;

end

end

`converged` $\leftarrow \|\mathbf{R}\| < \epsilon$;

$\mathbf{U} \leftarrow [\mathbf{U}, \mathbf{R}]$;

end

return $\mathbf{U}$;

Algorithm 2: Davidson diagonalization procedure. Note that $[.,.]$ stands for column-wise matrix concatenation.

4. We extract $|\alpha^\star\rangle \in \mathcal{A}_n^\star$, the subset of determinants $|\alpha\rangle \in \mathcal{A}_n$ with the largest contributions $e_\alpha^{(2)}$, and add them to the variational space $\mathcal{I}_{n+1} = \mathcal{I}_n \cup \mathcal{A}_n^\star$. In practice, in the case of the single-state calculation, we aim at doubling the size of the reference wave function at each iteration.
5. Iterate until the desired convergence has been reached.

All the details of our current implementation are reported in Appendix 5.12.4. In the remaining of this section, we only discuss the algorithm of our new stochastic CIPSI selection.

5.4.4.2 New stochastic selection In the past, CIPSI calculations were only possible in practice thanks to approximations. The first approximation restricts the set $\mathcal{A}_n$ by defining a set of *generators*. Indeed, it is very unlikely that $|\alpha\rangle$ will be selected if it is not connected to any $|I\rangle$ with a large coefficient, so only the determinants with the largest coefficients are generators. A second approximation defines a set of *selectors* in order to reduce the cost of $e_\alpha^{(2)}$ by removing the determinants with the smallest coefficients in the expression of $\Psi^{(0)}$ in $E^{(2)}$. This approximate scheme was introduced in the 80's and is known as *three-class CIPSI*.[157] The downside of these approximations is that the calculation is biased and, consequently, does not strictly converge to the FCI limit. Moreover, similar to the initiator approximation in FCIQMC,[13] this scheme suffers from a size-consistency issue.[242] The stochastic selection that we describe in this section (asymptotically) cures this problem, as there is *no threshold* on the wave function: if the calculation is run long enough, the unbiased FCI solution is obtained.

Recently, some of us developed a hybrid deterministic/stochastic algorithm for the computation of $E^{(2)}$.[17] The main idea is to rewrite the expression of

$$E^{(2)} = \sum_\alpha c_\alpha \langle \Psi^{(0)} | \hat{H} | \alpha \rangle \tag{5.20}$$

into elementary contributions labeled by the determinants of the internal space:

$$E^{(2)} = \sum_I \sum_{\alpha \in \mathcal{A}_I} c_\alpha \langle \Psi^{(0)} | \hat{H} | \alpha \rangle = \sum_I \varepsilon_I, \tag{5.21}$$

where

$$c_\alpha = \frac{\langle \Psi^{(0)} | \hat{H} | \alpha \rangle}{E^{(0)} - \langle \alpha | \hat{H} | \alpha \rangle} \tag{5.22}$$

is the corresponding coefficient estimated via first-order perturbation theory, and $\mathcal{A}_I$ is the subset of determinants $|\alpha\rangle$ connected to $|I\rangle$ by $\hat{H}$ such that $|\alpha\rangle \notin \cup_{K<I}\mathcal{A}_K$. The sum is decomposed into a stochastic and a deterministic contribution

$$E^{(2)} = \sum_{J\in\mathcal{D}} \varepsilon_J + \sum_{K\in\mathcal{S}} \varepsilon_K, \tag{5.23}$$

where $\mathcal{D}$ and $\mathcal{S}$ are the sets of determinants included in the deterministic and stochastic components, respectively.

The $|I\rangle$'s are sorted by decreasing c_I^2, and two processes are used simultaneously to compute the contributions ε_I. The first process is stochastic and $|I\rangle$ is drawn according to c_I^2. When a given ε_I has been computed once, its contribution is stored such that if $|I\rangle$ is drawn again later the contribution does not need to be recomputed. The only update is to increment the number of times it has been drawn for the Monte Carlo statistics. In parallel, a deterministic process is run, forcing to compute the contribution ε_I with the smallest index which has yet to be computed. The deterministic component is chosen as the first contiguous set of ε_I. Hence, the computation of $E^{(2)}$ is unbiased, and the exact deterministic value can be obtained in a finite time if the calculation is run long enough. The stochastic part is only a convergence accelerator providing a reliable error bar. The computation of $E^{(2)}$ is run with a default stopping criterion set to $|\delta E^{(2)}/E^{(2)}| = 0.002$, where $\delta E^{(2)}$ is the statistical error associated with $E^{(2)}$. We would like to stress that, thanks to the present semistochastic algorithm, the complete wave function is considered, and that no threshold is required. Consequently, size-consistency will be preserved if a size-consistent perturbation theory is applied.

While performing production runs, we have noticed that the computation of $E^{(2)}$ was faster than the CIPSI selection. Hence, we have slightly modified the routines computing $E^{(2)}$ such that the selection of determinants is performed alongside the computation of $E^{(2)}$. This new on-the-fly CIPSI selection performed during the stochastic PT2 calculation completely removes the conventional (deterministic) selection step, and the determinants are selected with no additional cost. We have observed that, numerically, the curves of the variational energy as a function of N_{det} obtained with either the deterministic or the stochastic selections are indistinguishable, so that the stochastic algorithm does not harm the selection's quality.

For the selection of multiple states, one PT2 calculation is run for each state and, as proposed by Angeli *et al.*,[243] the selection criterion is modified as

$$\tilde{e}_\alpha^{(2)} = \sum_k^{N_{\text{states}}} \frac{c_{\alpha k}}{\max_I c_{Ik}^2} \langle \Psi_k^{(0)} | \hat{H} | \alpha \rangle \,, \tag{5.24}$$

with

$$c_{\alpha k} = \frac{\langle \Psi_k^{(0)} | \hat{H} | \alpha \rangle}{\langle \Psi_k^{(0)} | \hat{H} | \Psi_k^{(0)} \rangle - \langle \alpha | \hat{H} | \alpha \rangle}. \tag{5.25}$$

This choice gives a balanced selection between states of different multi-configurational nature.

5.5 Spin-adapted wave functions

Determinant-based sCI algorithms generate wave functions expressed in a truncated space of determinants. Obviously, the selection presented in the previous section does not enforce that $\hat{H}$ commutes with $\hat{S}^2$ in the truncated space. Hence, the eigenstates of $\hat{H}$ are usually not eigenvectors of $\hat{S}^2$, although the situation improves when the size of the internal space tends to be complete. A natural way to circumvent this problem is to work in the basis of *configuration state functions* (CSFs), but this representation makes the direct computation of the Hamiltonian less straightforward during the Davidson diagonalization.

Instead, we follow the same path as the MELD and SCIEL codes,[244–246] and identify all the spatial occupation patterns associated with the determinants.[247] We then generate all associated spin-flipped configurations, and add to the internal space all the missing determinants. This procedure ensures that $\hat{H}$ commutes with $\hat{S}^2$ in the truncated space, and spin-adapted states are obtained by the diagonalization of $\hat{H}$. In addition, we apply a penalty method in the diagonalization by modifying the Hamiltonian as[248]

$$\tilde{\mathbf{H}} = \mathbf{H} + \gamma \Big(\mathbf{S}^2 - \mathbf{I} \langle S^2 \rangle_{\text{target}} \Big)^2, \tag{5.26}$$

where $\mathbf{I}$ is the identity matrix and γ is a fixed parameter set to 0.1 by default. This improves the convergence to the desired spin state, but also separates degenerate states with different spins, a situation that can potentially occurs with Rydberg states. In the Davidson algorithm,

this requires the additional computation of $\mathbf{S}^2\,\mathbf{U}$, for which the cost is expected to be the same as the cost of $\mathbf{H}\,\mathbf{U}$ (see Algorithm 2). The cost of computing $\mathbf{H}\,\mathbf{U}$ and $\mathbf{S}^2\,\mathbf{U}$ is mostly due to the search of the connected pairs of determinants, namely the determinants $\langle I|$ and $|J\rangle$ for which $\langle I|\hat{H}|J\rangle$ and $\langle I|\hat{S}^2|J\rangle$ are not zero due to Slater-Condon's rules. We have modified the function computing $\mathbf{H}\,\mathbf{U}$ so that it also computes $\mathbf{S}^2\,\mathbf{U}$ at the same time. Hence, the search of connected pairs is done once for both operations and $\mathbf{S}^2\,\mathbf{U}$ is obtained with no extra computational cost.

Working with spin-adapted wave functions increases the size of the internal space by a factor usually between 2 and 3, but it is particularly important if one is willing to obtain excited states.[19,173,174,181] Therefore, the default in Quantum Package is to use spin-adapted wave functions.

5.6 Parallelism

In Quantum Package, multiple parallelism layers are implemented: a fine-grained layer to benefit from shared memory, an intermediate layer to benefit from fast communication within a group of nodes, and a coarse-grained layer to interconnect multiple groups of nodes. Fine-grained parallelism is performed with OpenMP[249] in almost every single routine. Then, to go beyond a single compute node, Quantum Package does not use the usual single program/multiple data (SPMD) paradigm. A task-based parallelism framework is implemented with the ZeroMQ library.[250] The single-node instance runs a compute process as well as a task server process, while helper programs can be spawned asynchronously on different (heterogeneous) machines to run a distributed calculation. The helper programs can connect via ZeroMQ to the task server at any time, and contribute to a running calculation. As the ZeroMQ library does not take full advantage of the low latency hardware present in HPC facilities, the helper programs are parallelized also with the message passing interface[251] (MPI) for fast communication among multiple client nodes, typically for fast broadcasting of large data structures.

Hence, we have 3 layers of parallelism in Quantum Package: OpenMP, MPI and

ZeroMQ. This allows for an elastic management of resources: a running calculation taking too much time can be dynamically accelerated by plugging in more computing resources, by submitting more jobs in the queue or possibly in the cloud, i.e. outside of the HPC facility. This scheme has the advantage that it is not necessary to wait for all the nodes to be free to start a calculation, and hence minimizes the waiting time in the batch queue. It also gives the possibility to use altogether different helper programs. For instance, one could use a specific GPU-accelerated helper program on a GPU node while CPU-only helpers run on the CPU-only partition of the cluster. It is also possible to write a helper program that helps only one PT2/selection step and then exit, allowing to gather resources after the PT2/selection has started, and freeing them for the following diagonalization step.

The current limitation of Quantum Package is the memory of the single-node instance. We have not yet considered the possibility to add more compute nodes to increase the available memory, but this can be done by transforming the main program into an MPI program using scattered data structures.

We now describe how the Davidson and PT2/selection steps are parallelized.

5.6.1 Davidson diagonalization

In the direct Davidson diagonalization method, the computational bottleneck is the matrix product $\mathbf{W} = \mathbf{H}\,\mathbf{U}$, and only this step needs to be distributed. The calculation is divided into independent tasks where each task builds a unique piece of $\mathbf{W}$ containing 40 000 consecutive determinants. Communicating the result of all the tasks scales as $\mathcal{O}(N_{\text{det}})$, independently of the number of parallel processes. On the other hand, $\mathbf{U}$ needs to be broadcast efficiently at the beginning of the calculation to each slave process.

The computation of a task is parallelized with OpenMP, looping in a way that guarantees a safe write access to $\mathbf{W}$, avoiding the need of a lock. When idle, a slave process requests a task to the ZeroMQ task server, computes the corresponding result and sends it to the *collector* thread of the master instance via ZeroMQ. As the OpenMP tasks are not guaranteed to be balanced, we have used a dynamic scheduling, with a chunk size of 64 elements. The reason for this chunk size is to force that multiple threads access to $\mathbf{W}$ at memory addresses

far apart, avoiding the so-called *false sharing* performance degradation that occurs when multiple threads write simultaneously in the same cache line.[252] When the task is fully computed, the computed piece of **W** is sent back to the master process and a new task is requested, until the task queue is empty.

The **U** and **W** arrays are shared among threads, as well as all the large constant data needed for the calculation such as the ERIs. Sharing **U** also provides the benefit to reduce the amount of communication since **U** needs to be fetched only once for each node, independently of the number of cores. To make the broadcast of **U** efficient, the slave helper program is parallelized with MPI in a SPMD fashion, and each node runs a single MPI process. The **U** matrix is fetched from the ZeroMQ server by the process with rank zero, and then it is broadcast to the other slave processes within the same MPI job via MPI primitives. Then, each MPI process behaves independently and communicates via ZeroMQ with the task server, and with the master node which collects the results. A schematic view of the communication is presented in Fig 16.

5.6.2 CIPSI selection and PT2 energy

In the computation of $E^{(2)}$ and the CIPSI selection, each task corresponds to the computation of one ε_J or ε_K in Eq. (5.23), together with the selection of the associated external determinants. To establish the list of tasks, the Monte Carlo sampling is pre-computed on the master node. We associate to each task the number of drawn Monte Carlo samples such that running averages can be computed when the results of the tasks have been received by the collector thread. When the convergence criterion is reached, the task queue is emptied and the collector waits for all the running tasks to terminate.

As opposed to the Davidson implementation where each task is parallelized with OpenMP, here each OpenMP thread handles independently a task computed on a single core. Hence, there are multiple ZeroMQ clients per node, typically one per core, requesting tasks to the task server and sending the results back to the collector thread (see Fig. 17). Here, all the OpenMP threads are completely independent during the whole selection, and this explains the pleasing scaling properties of our implementation, as shown in Sec. 5.7.3. As in the

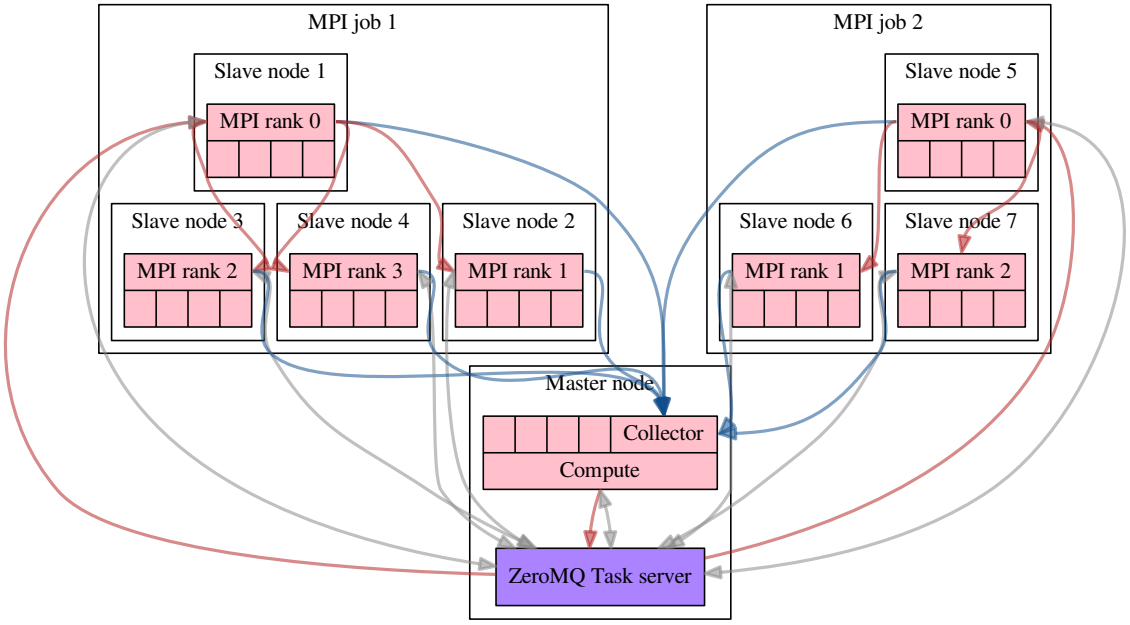

Figure 16: Communications in the Davidson diagonalization for a calculation with a master node and two helper MPI jobs, each using 4 cores for the computation. Red arrows represent the broadcast of **U** starting from the compute process of the master node, gray arrows the exchange of ZeroMQ messages with the task server and blue arrows the collection of the results.

Davidson distributed scheme, when the helper programs are run with MPI all the common data are communicated once from the ZeroMQ server to the rank-zero MPI process. Then, the data is broadcast to all the other processes with MPI primitives (there is one MPI process per node).

5.7 Results

5.7.1 Capabilities of Quantum Package

Before illustrating the new features of Quantum Package in the next subsection. We propose to give an overview of what can be achieved (in terms of system and basis set sizes) with the current implementation of Quantum Package. To do so we propose to review some of our very recent studies.

In Ref. [19], we studied 18 small molecules (water, hydrogen sulfide, ammonia, hydrogen chloride, dinitrogen, carbon monoxide, acetylene, ethylene, formaldehyde, methanimine, thioformaldehyde, acetaldehyde, cyclopropene, diazomethane, formamide, ketene, nitrosomethane, and the smallest streptocyanine) with sizes ranging from 1 to 3 non-hydrogen atoms. For such systems, using sCI expansions of several million determinants, we were able to compute more than hundred highly accurate vertical excitation energies with typically augmented triple-ζ basis sets. It allowed us to benchmark a series of 12 state-of-the-art excited-state wave function methods accounting for double and triple excitations.

Even more recently,[181] we provided accurate reference excitation energies for transitions involving a substantial amount of double excitation using a series of increasingly large diffuse-containing atomic basis sets. Our set gathered 20 vertical transitions from 14 small- and medium-size molecules (acrolein, benzene, beryllium atom, butadiene, carbon dimer and trimer, ethylene, formaldehyde, glyoxal, hexatriene, nitrosomethane, nitroxyl, pyrazine, and tetrazine). For the smallest molecules, we were able to obtain well converged excitation energies with augmented quadruple-ζ basis set while only augmented double-ζ bases were manageable for the largest systems (such as acrolein, butadiene, hexatriene and benzene).

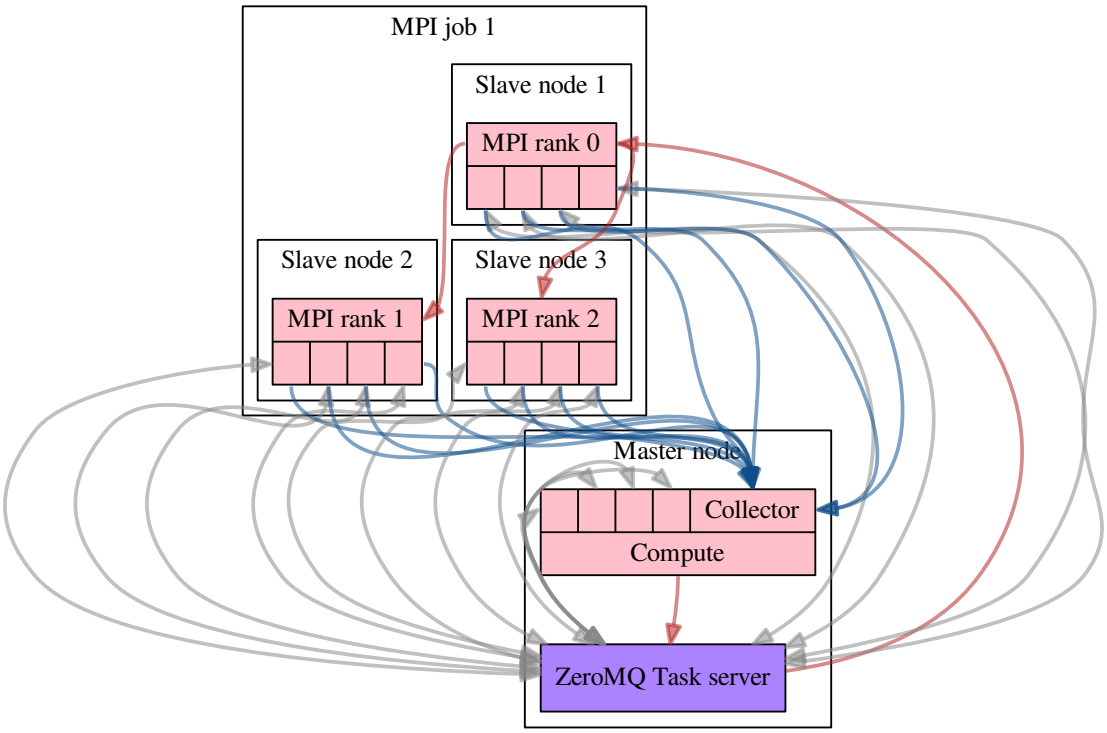

Figure 17: Communications in the stochastic selection for a calculation with a master node and one helper MPI job, each using 4 cores for the computation. Red arrows represent the broadcast of the common data starting from the compute process of the master node, gray arrows the exchange of ZeroMQ messages with the task server and blue arrows the collection of the results.

Note that the largest sCI expansion considered in this study had more than 200 million determinants.

In Ref. [196], Giner *et al.* studied even larger systems containing transition metals: $[CuCl_4]^{2-}$, $[Cu(NH_3)_4]^{2+}$ and $[Cu(H_2O)_4]^{2+}$. They were able, using large sCI expansions, to understand the physical phenomena that determine the relative energies of three of the lowest electronic states of each of these square-planar copper complexes.

5.7.2 Extrapolation

To illustrate the extrapolation procedure described in Sec. 5.3.4, we consider a cyanine dye[253] $H_2N{-}CH{=}NH_2^+$ (labeled as CN3 in the remaining) in both its ground state and first excited state.[18,254,255] The geometry is the equilibrium geometry of the ground state optimized at the PBE0/cc-pVQZ level.[255] The ground state is a closed shell, well described by a single reference, while the excited state is singly excited and requires, at least, two determinants to be properly modeled. The calculations were performed in the aug-cc-pVDZ basis set with state-averaged natural orbitals obtained from an initial CIPSI calculation. All the electrons were correlated, so the FCI space which is explored corresponds to a CAS(24,114) space. The reference excitation energy, obtained at the CC3/ANO-L-VQZP level is 7.18 eV[254] (see also Ref. [18]). Note that this particular transition is fairly insensitive to the basis set as long as at least one set of diffuse functions is included. For example, we have obtained 7.14 and 7.13 eV at the CC3/aug-cc-pVDZ and CC3/aug-cc-pVTZ levels, respectively.[19]

In Fig. 18, we plot the energy convergence of the ground state (GS) and the excited state (ES) as a function of the number of determinants N_{det}, with and without the second-order perturbative contribution. From the data gathered in Table 3, one can see that, although $E^{(2)}$ is still large (roughly 0.02 a.u.), the sCI+PT2 and sCI+rPT2 excitation energies converge to a value of 7.20 eV compatible with the reference energy obtained in a larger basis set. We have also plotted the sCI+rPT2 energy given by $E^{(0)} + ZE^{(2)}$ (see Sec. 5.3.4.2) and we clearly see that this quantity converges much faster than the usual sCI+PT2 energy. Even for very small reference wave function, the energy gap between GS and ES is qualitatively correct. The graph of Fig. 19, which shows the zeroth-order energy $E^{(0)}$ as a function of the

second-order energy $E^{(2)}$ (dotted lines) or its renormalization variant $Z\,E^{(2)}$ (solid lines), also indicates that it is practically much easier to extrapolate to the FCI limit using the rPT2 correction.

As a second test case for rPT2, we consider the widely-studied example of the chromium dimer (Cr_2) in its $^1\Sigma_g^+$ ground state.[14,156,170,256–266] This system is notoriously challenging as it combines dynamic and static correlation effects hence requiring multi-configurational methods and large basis sets in order to have a balanced treatment of these two effects. Consequently, we compute its ground-state energy in the cc-pVQZ basis set with an internuclear distance $R_{\mathrm{Cr-Cr}} = 1.68\,\text{Å}$ close to its experimental equilibrium geometry. Our full-valence calculation corresponds to an active space CAS(28,198) and the computational protocol is similar to the previous example. The second-order corrected value $E^{(0)} + E^{(2)}$ as well as its renormalized version $E^{(0)} + ZE^{(2)}$ as a function of the number of determinants in the reference wave function are reported in Table 4 and depicted in Fig. 20. Here also, we observe that rPT2 is clearly a superior extrapolation framework compared to the standard PT2 version as it yields a much straighter extrapolation curve, even in the case of a strongly correlated system such as Cr_2. The renormalization factor Z [see Eq. (5.17)] mitigates strongly the overestimation of the FCI energy for small wave functions by damping the second-order energy $E^{(2)}$. Linear extrapolations of the PT2 and rPT2 energies based on the two largest wave functions yields extrapolated FCI energies of -2087.734 and -2087.738, respectively (see also Table 4). The difference between these two extrapolated FCI energies provides a qualitative idea of the extrapolation accuracy.

5.7.3 Speedup

In this Section, we discuss the parallel efficiency of the algorithms implemented in Quantum Package. The system we chose for these numerical experiments is the benzene molecule C_6H_6 for which we have performed sCI calculations with the 6-31G* basis set. The frozen-core approximation has been applied and the FCI space that we explore is a CAS(30,90). The measurements were made on GENCI's Irene supercomputer. Each Irene's node is a dual-socket Intel(R) Xeon(R) Platinum 8168 CPU@2.70GHz with 192GiB of RAM,

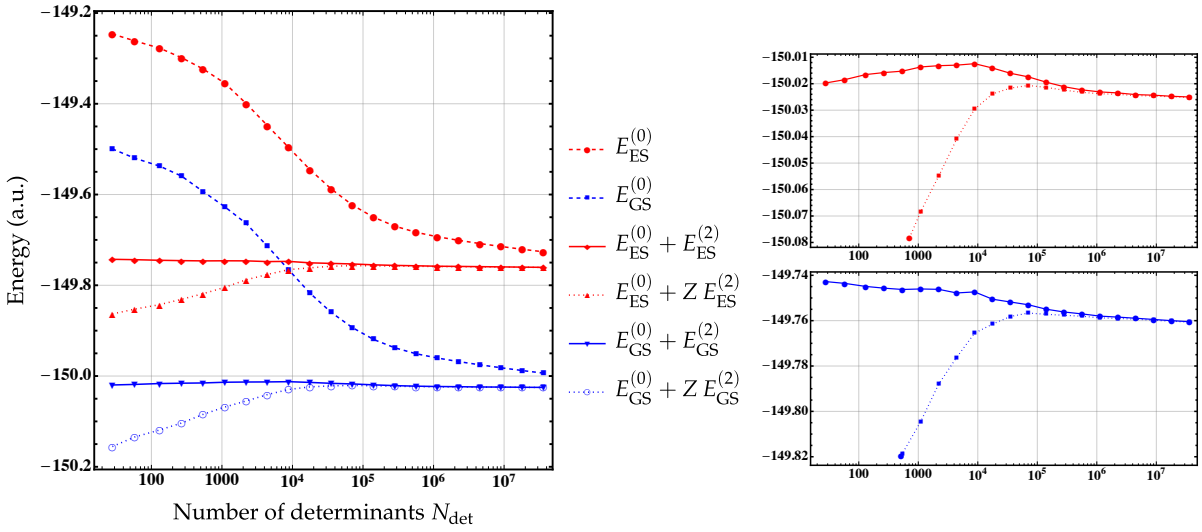

Figure 18: Energy convergence of the ground state (GS, in blue) and excited state (ES, in red) of CN3 with respect to the number of determinants N_{det} in the reference space. The zeroth-order energy $E^{(0)}$ (dashed) , its second-order corrected value $E^{(0)} + E^{(2)}$ (dotted) as well as its renormalized version $E^{(0)} + ZE^{(2)}$ (solid) are represented. See Table 3 for raw data.

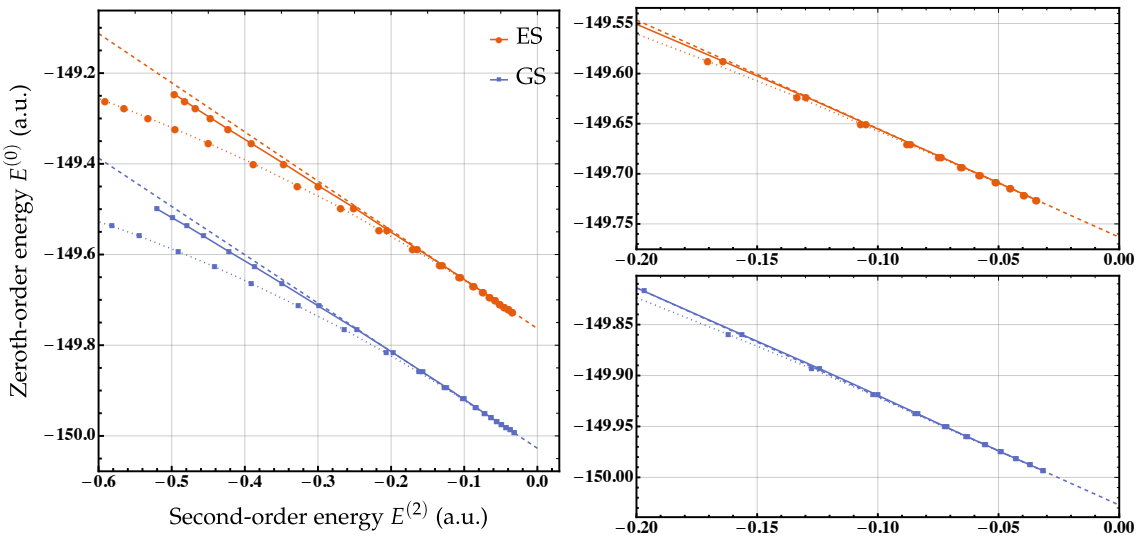

Figure 19: Zeroth-order energy $E^{(0)}$ as a function of the second-order energy $E^{(2)}$ (dotted lines) or its renormalized variant $Z\,E^{(2)}$ (solid lines). A linear fit (dashed lines) of the last 6 points is also reported for comparison. See Table 3 for raw data.

with a total of 48 physical CPU cores. Parallel speedup curves are made up to 12 288 cores (i.e. 256 nodes) for i) a single iteration of the Davidson diagonalization, and ii) the hybrid semistochastic computation of $E^{(2)}$ (which includes the CIPSI selection). The speedup reference corresponds to the single node calculation (48 cores).

First, we measure the time required to perform a single Davidson iteration as a function of the number of CPU cores for the two largest wave functions ($N_{\text{det}} = 25 \times 10^6$ and 100×10^6). The timings are reported in Table 5 while the parallel speedup curve is represented in Fig. 21. The parallel efficiency increases together with N_{det}, as shown in Fig. 21. For the largest wave function, a parallel efficiency of 66% is obtained on 192 nodes (i.e. 9216 cores). We note that the speedup reaches a plateau at 3 072 cores (64 nodes) for $N_{\text{det}} = 25 \times 10^6$. For this wave function, there are 625 tasks computing each 40 000 rows of $\mathbf{W}$. When the number of nodes reaches 64, the number of tasks is too small for the load to be balanced between the nodes, and the computational time is limited by the time taken to compute the longest task. The same situation arises for $N_{\text{det}} = 100 \times 10^6$ with 9 408 cores (192 nodes), with 2 500 tasks to

Table 3: Zeroth-order energy $E^{(0)}$, second-order perturbative correction $E^{(2)}$ and its renormalized version $ZE^{(2)}$ (in hartree) of CN3 for increasingly large wave functions. The excitation energy ΔE (in eV) is the energy difference between the ground state (GS) and the excited state (ES). The statistical error, corresponding to one standard deviation, is reported in parenthesis.

	$E^{(0)}$		$E^{(0)} + E^{(2)}$			$E^{(0)} + ZE^{(2)}$		
N_{det}	GS (a.u.)	ES (a.u.)	GS (a.u.)	ES (a.u.)	ΔE (eV)	GS (a.u.)	ES (a.u.)	ΔE (eV)
28	-149.499 574	-149.246 268	-150.155(1)	-149.863(1)	7.95(5)	-150.020(1)	-149.743(1)	7.54(5)
58	-149.519 908	-149.261 390	-150.134(1)	-149.853(1)	7.67(5)	-150.018(1)	-149.744(1)	7.48(5)
131	-149.537 424	-149.277 496	-150.118(1)	-149.842 7(9)	7.52(4)	-150.017(1)	-149.744 9(9)	7.39(4)
268	-149.559 465	-149.298 484	-150.103 5(9)	-149.830 8(9)	7.42(4)	-150.015 8(9)	-149.745 7(9)	7.35(4)
541	-149.593 434	-149.323 302	-150.084 5(8)	-149.818 6(8)	7.24(4)	-150.015 2(8)	-149.746 3(8)	7.32(4)
1 101	-149.627 202	-149.354 807	-150.068 3(8)	-149.804 5(8)	7.18(3)	-150.013 7(8)	-149.746 0(8)	7.28(3)
2 207	-149.663 850	-149.399 522	-150.054 9(7)	-149.787 9(7)	7.26(3)	-150.013 2(7)	-149.746 2(7)	7.27(3)
4 417	-149.714 222	-149.448 133	-150.040 9(6)	-149.776 2(6)	7.20(3)	-150.013 0(6)	-149.747 8(6)	7.22(3)
8 838	-149.765 886	-149.496 401	-150.029 6(5)	-149.765 5(5)	7.19(2)	-150.012 4(5)	-149.747 3(5)	7.21(2)
17 680	-149.817 301	-149.545 048	-150.023 9(4)	-149.761 5(4)	7.14(2)	-150.014 1(4)	-149.750 5(4)	7.17(2)
35 380	-149.859 737	-149.587 668	-150.021 6(3)	-149.758 2(3)	7.17(1)	-150.016 1(3)	-149.751 8(3)	7.19(1)
70 764	-149.893 273	-149.623 235	-150.020 7(2)	-149.756 6(3)	7.18(1)	-150.017 4(2)	-149.753 0(3)	7.19(1)
141 545	-149.919 463	-149.650 109	-150.021 4(2)	-149.757 2(2)	7.189(8)	-150.019 4(2)	-149.755 0(2)	7.196(8)
283 108	-149.937 839	-149.669 735	-150.022 4(2)	-149.757 6(2)	7.206(7)	-150.021 1(2)	-149.756 2(2)	7.209(7)
566 226	-149.950 918	-149.683 278	-150.023 3(1)	-149.758 0(1)	7.217(6)	-150.022 3(1)	-149.757 0(1)	7.219(6)
1 132 520	-149.960 276	-149.693 053	-150.023 8(1)	-149.758 8(1)	7.212(5)	-150.023 1(1)	-149.758 0(1)	7.214(5)
2 264 948	-149.968 203	-149.700 907	-150.024 0(1)	-149.759 0(1)	7.211(4)	-150.023 5(1)	-149.758 4(1)	7.212(4)
4 529 574	-149.975 230	-149.708 061	-150.024 5(1)	-149.759 4(1)	7.215(4)	-150.024 1(1)	-149.758 9(1)	7.216(4)
9 057 914	-149.981 770	-149.714 526	-150.024 63(9)	-149.759 81(8)	7.206(3)	-150.024 34(9)	-149.759 48(8)	7.207(3)
18 110 742	-149.987 928	-149.720 648	-150.024 95(7)	-149.760 25(8)	7.203(3)	-150.024 74(7)	-149.760 00(8)	7.204(3)
36 146 730	-149.993 593	-149.726 253	-150.025 27(6)	-149.760 65(7)	7.198(3)	-150.025 02(6)	-149.760 47(7)	7.198(3)

Table 4: Zeroth-order energy $E^{(0)}$, second-order perturbative correction $E^{(2)}$ and its renormalized version $ZE^{(2)}$ (in hartree) as a function of the number of determinants N_{det} for the ground-state of the chromium dimer Cr_2 computed in the cc-pVQZ basis set. The statistical error, corresponding to one standard deviation, is reported in parenthesis.

N_{det}	$E^{(0)}$	$E^{(0)} + E^{(2)}$	$E^{(0)} + ZE^{(2)}$
1 631	-2086.742 321	-2087.853(3)	-2087.679(2)
3 312	-2086.828 496	-2087.821(2)	-2087.688(1)
6 630	-2086.920 161	-2087.792(1)	-2087.694(1)
13 261	-2087.008 701	-2087.764(1)	-2087.694(1)
26 562	-2087.091 669	-2087.743(1)	-2087.692(1)
53 129	-2087.165 533	-2087.725(1)	-2087.689(1)
106 262	-2087.234 564	-2087.710 2(9)	-2087.685 0(8)
212 571	-2087.293 488	-2087.703 0(8)	-2087.685 0(7)
425 185	-2087.343 762	-2087.697 3(7)	-2087.684 4(7)
850 375	-2087.386 276	-2087.697 8(6)	-2087.688 1(6)
1 700 759	-2087.422 707	-2087.698 9(6)	-2087.691 6(5)
3 401 504	-2087.454 427	-2087.700 7(5)	-2087.695 1(5)
6 802 953	-2087.482 238	-2087.703 2(4)	-2087.698 8(4)
13 605 580	-2087.506 838	-2087.705 6(4)	-2087.702 2(4)
27 210 163	-2087.528 987	-2087.709 2(4)	-2087.706 4(4)
54 415 174	-2087.549 261	-2087.711 6(3)	-2087.709 5(3)
Extrap.		-2087.734	-2087.738

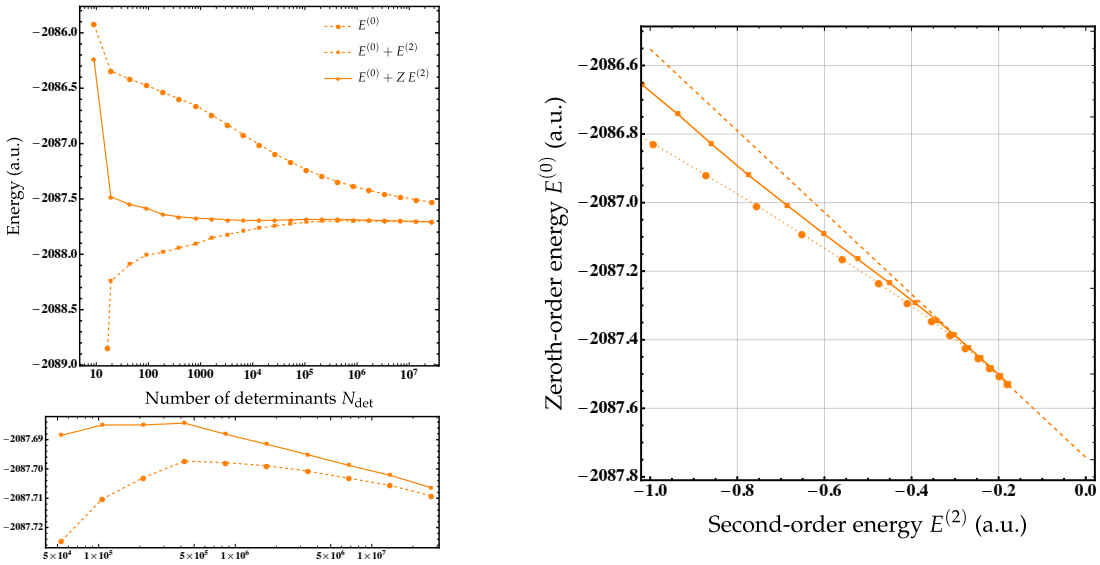

Figure 20: Left: Energy convergence of the ground state of Cr_2 with respect to the number of determinants N_{det} in the reference space. The zeroth-order energy $E^{(0)}$ (dashed) , its second-order corrected value $E^{(0)} + E^{(2)}$ (dotted) as well as its renormalized version $E^{(0)} + ZE^{(2)}$ (solid) are represented. Right: Zeroth-order energy $E^{(0)}$ as a function of the second-order energy $E^{(2)}$ (dotted lines) or its renormalization variant $Z\,E^{(2)}$ (solid lines). A linear fit (dashed lines) of the last 2 points is also reported for comparison. See Table 4 for raw data.

compute.

Second, we analyze the parallel efficiency of the calculation of $E^{(2)}$ for the sCI wave function with $N_{\text{det}} = 25 \times 10^6$. The stopping criterion during the calculation of $E^{(2)}$ is given by a relative statistical error below 2×10^{-3} of the current $E^{(2)}$ value. The speedups are plotted in Fig. 21 (see also Table 5). For 192 nodes, one obtains a parallel efficiency of 89%. The present parallel efficiency is not as good as the one presented in the original paper.[17] The reason behind this is a faster computation of $e_\alpha^{(2)}$, which reduces the parallel efficiency by increasing the ratio communication/computation.

5.8 Developing in Quantum Package

5.8.1 The Quantum Package philosophy

Quantum Package is a standalone easy-to-use library for developers. The main goals of Quantum Package are to i) facilitate the development of new quantum chemistry methods, ii) minimize the dependency on external programs/libraries, and iii) encourage the collaborative and educative work through human readable programs. Therefore, from the developer point of view, Quantum Package can be seen as a standalone library containing all important quantities needed to perform quantum chemistry calculations, both involving wave function theory, through the determinant driven algorithms, and DFT methods, thanks to the presence of a quadrature grid for numerical integrations and basic functionals. These appealing features are made more concrete thanks to the organization of Quantum Package in terms of core modules and plugins (see Sec. 5.8.3) together with its programming language (see Sec. 5.8.2), which naturally creates a very modular environment for the programmer.

Although Quantum Package is able to perform all the required steps from the calculation of the one- and two-electron integrals to the computation of the sCI energy, interfacing Quantum Package, at any stage, with other programs is relatively simple. For example, canonical or CASSCF molecular orbitals can be imported from GAMESS,[96] while atomic and/or molecular integrals can be read from text files like fcidump. Thanks to this flexibility,

Table 5: Wall-clock time (in seconds) to perform a single Davidson iteration and a second-order correction $E^{(2)}$ calculation (which also includes the CIPSI selection) with an increasing number of 48-core compute nodes N_{nodes}. The statistical error obtained on $E^{(2)}$, defining the stopping criterion, is 0.17×10^{-3} a.u.

N_{nodes}	Wall-clock time (in seconds)		
	Davidson for $N_{\text{det}} = 25 \times 10^6$	Davidson for $N_{\text{det}} = 100 \times 10^6$	PT2/selection $N_{\text{det}} = 25 \times 10^6$
1	3 340	65 915	406 840
32	142	2 168	12 711
48	109	1 497	8 515
64	93	1 181	6 421
96	93	834	4 323
128	93	674	3 287
192	96	522	2 435
256	96	519	1 996

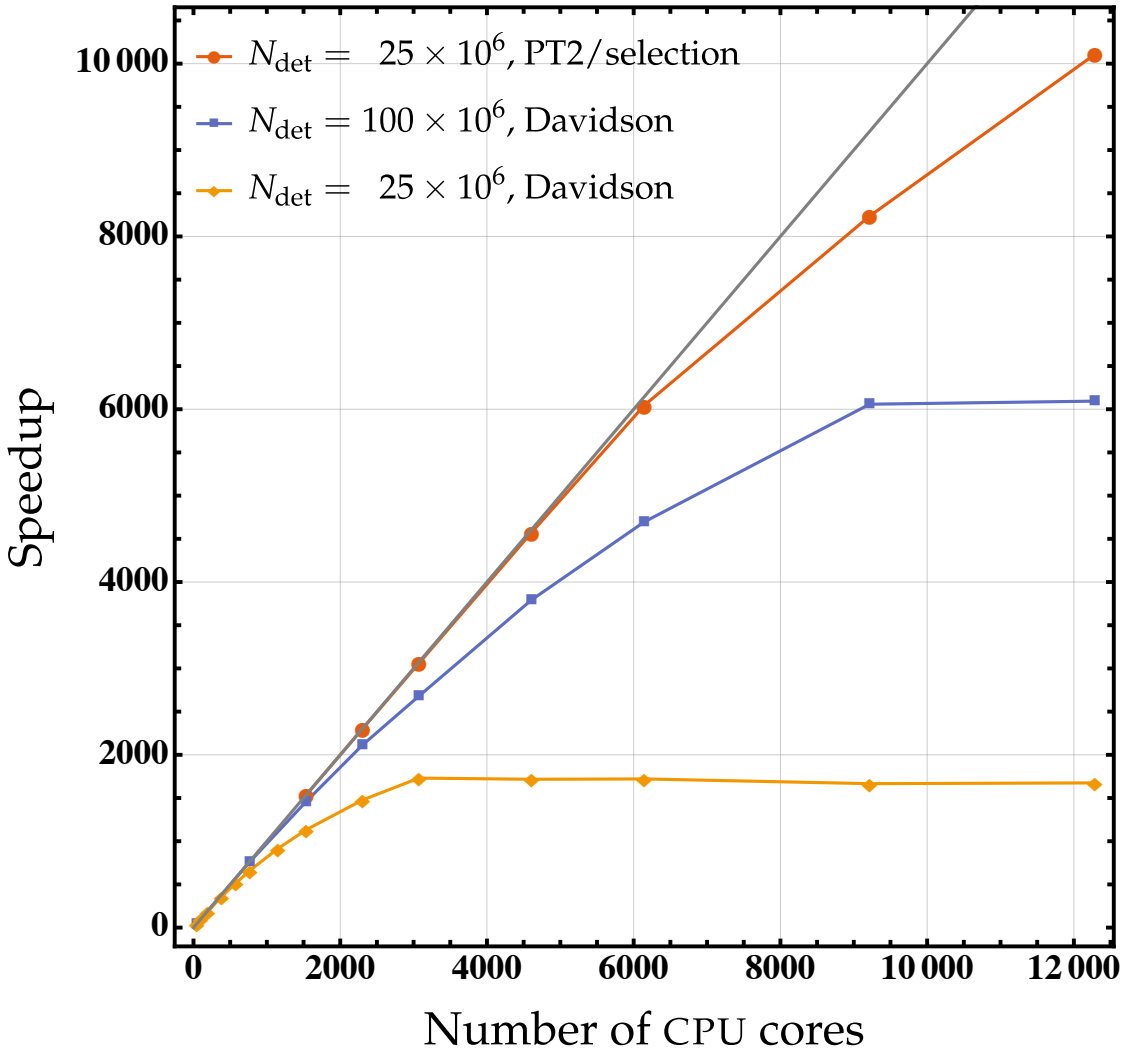

Figure 21: Speedup obtained for a single Davidson iteration (blue and yellow curves) and the combination of CIPSI selection and PT2 calculation (red curve) as a function of the number of CPU cores. For the Davidson diagonalization, two sizes of reference wave functions are reported ($N_{det} = 25 \times 10^6$ and 100×10^6), while for the PT2/selection calculation only results corresponding to the smallest wave function ($N_{det} = 25 \times 10^6$) are reported. See Table 5 for raw data.

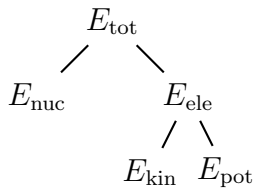

Figure 22: Production tree of the energy computed by IRPF90.

some of us are currently developing plugins for performing sCI calculations for periodic systems.

5.8.2 The IRPF90 code generator

It is not a secret that large scientific codes written in Fortran (or in similar languages) are difficult to maintain. The program's complexity originates from the inter-dependencies between the various entities of the code. As the variables are more and more coupled, the programs become more and more difficult to maintain and to debug. To keep a program under control, the programmer has to be aware of all the consequences of any source code modification within all possible execution paths. When the code is large and written by multiple developers, this becomes almost impossible. However, a computer can easily handle such a complexity by taking care of all the dependencies between the variables, in a way similar to how GNU Make handles the dependencies between source files.

IRPF90 is a Fortran code generator.[267] Schematically, the programmer only writes computation kernels, and IRPF90 generates the glue code linking all these kernels together to produce the expected result, handling all relationships between variables. To illustrate in a few words how IRPF90 works, let us consider the simple example which consists of calculating the total energy of a molecular system as the sum of the nuclear repulsion and the electronic energy $E_{\text{tot}} = E_{\text{nuc}} + E_{\text{ele}}$. The electronic energy is the sum of the kinetic and potential energies, i.e., $E_{\text{ele}} = E_{\text{kin}} + E_{\text{pot}}$.

The *production tree* associated with the computation of the total energy is shown in

Fig. 22. Within the IRPF90 framework, the programmer writes a *provider* for each *entity*, i.e., a node of the production tree. The provider is a subroutine whose only goal is to compute the value associated with the entity, assuming the values of the entities on which it depends are computed and valid. Hence, when an entity is used somewhere in the program (in a subroutine, a function or a provider), a call to its provider is inserted in the code before it is used such that the corresponding value is guaranteed to be valid.

Quantum Package is a library of providers designed to make the development of new wave function theory and DFT methods simple. Only a few programs using these providers are part of the core modules of Quantum Package, such as the sCI module using the CIPSI algorithm or the module containing the semi-stochastic implementation of the second-order perturbative correction. The main goal of Quantum Package is to be used as a library of providers, and programmers are encouraged to develop their own modules using Quantum Package.

5.8.3 The plugin system

External programmers should not add their contributions by modifying directly Quantum Package's core, but by creating their own modules in independent repositories hosted and distributed by themselves. This model gives more freedom to the developers to distribute modules as we do not enforce them to follow any rule. The developers are entirely responsible for their own plugins. This model has the advantage to redirect immediately the users to the right developer for questions, installation problems, bug reports, etc.

Quantum Package integrates commands to download external repositories and integrate all the plugins of these repositories into the current installation of Quantum Package. External plugins appear exactly as if they were part of Quantum Package, and if a plugin is useful for many users, it can be easily integrated in Quantum Package's core after all the coding and documentation standards are respected.

Multiple external plugins were developed by the authors. For instance, one can find a multi-reference coupled cluster program,[17,268] interfaces with the quantum Monte Carlo programs QMC=Chem,[192] QMCPack[194] and CHAMP,[269] an implementation of the shifted-Bk

method,[18] a program combining CIPSI with RSDFT,[270] a four-component relativistic RSDFT code,[271] and many others.

In particular, Quantum Package also contains the basic tools to use and develop range-separated density-functional theory (RSDFT, see, e.g., Refs. [272] , [273]) which allows to perform multi-configurational density-functional theory (DFT) calculations within a rigorous mathematical framework. In the core modules of Quantum Package, single-determinant approximations of RSDFT are available, which fall into the so-called range-separated hybrid[274,275] (RSH) approximation. These approaches correct for the wrong long-range behavior of the usual hybrid approximations thanks to the inclusion of the long-range part of the HF exchange. Quantum Package contains all necessary integrals to perform RSDFT calculations, including the long-range interaction integrals and Hartree-exchange-correlation energies and potentials derived from the short-range version of the local-density approximation (LDA)[276] and a short-range generalized-gradient approximation (GGA) based on the Perdew-Burke-Ernzerhof (PBE) functional.[277] All numerical integrals are performed using the standard Becke quadrature grid[278] associated with the improved radial grids of Mura et *al.*[279] With these tools, more evolved schemes based on RSDFT have been developed, such as an energy correlation functional with multideterminantal reference depending on the on-top pair density[280] or a basis set correction.[270] The corresponding source code can be found as external plugins (see, for example, https://gitlab.com/eginer/qp_plugins_eginer).

5.9 Conclusion

Significant improvements were brought to the second version of Quantum Package. Some were single-core optimizations, and others focused on the algorithm adaptation to large-scale parallelism (load balancing in particular). Currently, the code has a parallel efficiency that enables routinely to realize runs on roughly 2 000 CPU cores, with tens of millions of determinants in the reference space. Moreover, we have been able to push up to 12 288 cores (256 nodes) on GENCI's supercomputer Irene. Such a gain in efficiency has and will lead to many more challenging chemical applications.[19,165–169,174,181,197,198]

The Davidson diagonalization, which is at the center of sCI and FCI methods, suffers from the impossibility to fully store the Hamiltonian in the memory of a single node. The solution we adopted was to resort to *direct methods*, i.e., recomputing *on the fly* the matrix elements at each iteration. While an extremely fast method was already available to detect zero matrix elements,[281] the former implementation still had to search over the $\mathcal{O}(N_{\text{det}}^2)$ matrix elements for interacting determinant pairs. Now, determinants are split in disjoint sets entirely disconnected from each other. Thus, only a small fraction of the matrix elements need to be explored, and an algorithm with $\mathcal{O}(N_{\text{det}}^{3/2})$ scaling was proposed. While the parallelization of this method was somewhat challenging due to the extremely unbalanced nature of the elementary tasks, a distributed implementation was realized with satisfying parallel speedups (typically 35 for 50 nodes) with respect to the 48-core single-node reference.

Significant improvements were also realized in the computation of the second-order perturbative correction, $E^{(2)}$. A natural idea was to take into account the tremendous number of tiny contributions via a stochastic Monte Carlo approach. $E^{(2)}$ being itself an approximate quantity used for estimating the FCI energy, its exact value is indeed not required, as long as the value is unbiased and the statistical error is kept under control. Our scheme allows to compute $E^{(2)}$ with a small error bar for a few percent of the cost of the fully deterministic computation.

Similarly, the CIPSI selection is now performed stochastically alongside the PT2 calculation. Therefore, the selection part of the new stochastic CIPSI selection is virtually free as long as one is interested in the second-order perturbative correction.

Finally, efforts have been made to make this software as developer friendly as possible thanks to a very modular architecture that allows any developer to create his/her own module and to directly benefit from all pre-existing work.

5.10 License

Quantum Package is licensed under GNU Affero General Public License (AGPLv3).

5.11 Acknowledgements

The authors would like to thank the *Centre National de la Recherche Scientifique* (CNRS) for funding and Cyrus Umrigar for carefully reading the manuscript. Funding from *Projet International de Coopération Scientifique* (PICS08310) is also acknowledged. This work was performed using HPC resources from CALMIP (Toulouse) under allocation 2019-18005 and from GENCI-TGCC (Grant 2018-A0040801738). A.B. was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division, as part of the Computational Materials Sciences Program and Center for Predictive Simulation of Functional Materials. K.G. acknowledges support from grant number CHE1762337 from the U.S. National Science Foundation.

5.12 Implementation details

5.12.1 Efficiency of integral storage

The efficiency of the storage as a hash table was measured on a dual socket Intel Xeon E5-2680 v2@2.80GHz processor, taking the water molecule with the cc-pVQZ basis set (115 MOs). The time to access all the integrals was measured by looping over the entire set of ERIs using different loop orderings. The results are given in Table 6, the reference being the storage as a plain four-dimensional array.

In the array storage, the value of 170 ns/integral in the random access case is typical of the latency to fetch a value in the RAM modules, telling that the requested integral is almost never present in any level of cache. When the data is accessed with a stride of one (i, j, l, k storage), the hierarchical architecture of the cache levels accelerates the access by a factor of 18, down to 9.71 ns/integral, corresponding mostly to the overhead of the function call, the retrieval of the data being negligible.

With the hash table, the random access is only 2.18 times slower than the random access in the array. Indeed, two random accesses are required: one for the first element of the key

Table 6: Time to access integrals (in nanoseconds/integral) with different access patterns. The time to generate random numbers (measured as 67 ns/integral) was not counted in the random access results.

Access	Array	Hash table
i, j, k, l	9.72	125.79
i, j, l, k	9.72	120.64
i, k, j, l	10.29	144.65
l, k, j, i	88.62	125.79
l, k, i, j	88.62	120.64
Random	170.00	370.00

bucket to do the search, and one for the value of the integral. The remaining time corresponds to the binary search. The results show that data locality is exploited: when the access is done with a regular access pattern, the data is fetched roughly 3 times faster than using a random access, giving a latency below the latency of a random access in the array.

A CIPSI calculation was run once with the array storage, and once with the hash table storage. With the hash storage, the total wall clock time was increased only by a factor of two. To accelerate the access to the most frequently used integrals and reduce this overhead, we have implemented a software cache. All the integrals involving the 128 MOs closest to the Fermi level are copied in a dense array of 128^4 elements (2 GiB), and benefit from the fastest possible access.

5.12.2 Internal representation of determinants

Determinants can be conveniently written as a string of creation operators applied to the vacuum state $|\rangle$, e.g., $a_i^\dagger a_j^\dagger a_k^\dagger|\rangle = |I\rangle$. Because of the fermionic nature of electrons, a permutation of two contiguous creation operators results in a sign change $a_j^\dagger a_i^\dagger = -a_i^\dagger a_j^\dagger$, which makes their ordering relevant, e.g., $a_j^\dagger a_i^\dagger a_k^\dagger|\rangle = -|I\rangle$. A determinant can be broken down into two pieces of information: i) a set of creation operators corresponding to the set

of occupied spinorbitals in the determinant, and ii) an ordering of the creation operators responsible for the sign of the determinant, known as *phase factor*. Once an ordering operator $\hat{\mathcal{O}}$ is chosen and applied to all determinants, the phase factor may simply be included in the CI coefficient.

The determinants are built using the following order: i) spin-up ($\uparrow$) spinorbitals are placed before spin-down ($\downarrow$) spinorbitals, as in the Waller-Hartree double determinant representation[235] $\hat{\mathcal{O}}\,|I\rangle = \hat{I}|\rangle = \hat{I}_\uparrow \hat{I}_\downarrow|\rangle$, and ii) within each operator $\hat{I}_\uparrow$ and $\hat{I}_\downarrow$, the creation operators are sorted by increasing indices. For instance, let us consider the determinant $|J\rangle = a_j^\dagger a_k^\dagger a_{\bar{i}}^\dagger a_i^\dagger|\rangle$ built from the set of spinorbitals $\{i_\uparrow, j_\uparrow, k_\uparrow, i_\downarrow\}$ with $i < j < k$. If we happen to encounter such a determinant, our choice of representation imposes to consider its re-ordered expression $\hat{\mathcal{O}}\,|J\rangle = -a_i^\dagger a_j^\dagger a_k^\dagger a_{\bar{i}}^\dagger|\rangle = -\,|J\rangle$, and the phase factor must be handled.

The indices of the creation operators (or equivalently the spinorbital occupations), are stored using the so-called *bitstring* encoding. A bitstring is an array of bits; typically, the 64-bit binary representation of an integer is a bitstring of size 64. Quite simply, the idea is to map each spinorbital to a single bit with value set to its occupation number. In other words, 0 and 1 are associated with the *unoccupied* and *occupied* states, respectively.

For simplicity and performance considerations, the occupations of the spin-up and spin-down spinorbitals are stored in different bitstrings, rather than interleaved or otherwise merged in the same one. This allows to straightforwardly map orbital index p to bit index $p-1$ (orbitals are usually indexed from 1, while bits are indexed from 0). This makes the representation of a determinant a tuple of two bitstrings, associated with respectively spin-up and spin-down orbitals. A similar parity representation of the fermionic operators is commonly used in quantum computing.[282]

The storage required for a single determinant is, in principle, one bit per spinorbital, or $2 \times N_{\text{orb}}$ bits. However, because CPUs are designed to handle efficiently 64-bit integers, each spin part is stored as an array of 64-bit integers, the unused space being padded with zeros. The actual storage needed for a determinant is $2 \times 64 \times N_{\text{int}}$ bits, where $N_{\text{int}} = \lfloor (N_{\text{orb}} - 1)/64 \rfloor + 1$ is the number of 64-bit integers needed to store one spin part.

Taking advantage of low-level hardware instructions,[281] we are able, given two arbitrary determinants $|I\rangle$ and $|J\rangle$, to find with a minimal cost the excitation operator $\hat{T}$ such that

$|J\rangle = \hat{T}\,|I\rangle$. This is a necessary step to obtain the (i, j, k, l) indices of the two-electron integral(s) involved in the Hamiltonian matrix element between $|I\rangle$ and $|J\rangle$. Then, fetching the values of the integrals can be done quickly using the hash table presented in Sec. 5.4.1.

Because the data structure used to store determinants implies an ordering of the MOs, we also need to compute a phase factor. Here, we propose an algorithm to perform efficiently the computation of the phase factor. For a determinant $|I\rangle$ that is going to be used repeatedly for phase calculations, we introduce a *phase mask* represented as a bitstring:

$$P_I[i] = 1 \wedge \sum_{k=0}^{i} I[k], \tag{5.27}$$

where $\wedge$ denotes the *and* bitwise operation, and $I[k]$ is the kth bit of bitstring I, corresponding to the $(k+1)$th spinorbital of determinant $|I\rangle$ (remember that the orbital indices start at 1 and the bit indices start at 0). In other words, the ith bit of the phase mask is set to 1 if the number of electrons occupying the $i+1$ lowest spinorbitals is odd, and 0 otherwise. When an electron of determinant $|I\rangle$ is excited from orbital h to p, the associated phase factor is

$$\begin{cases} +(-1)^{P_I[h-1] \oplus P_I[p-1]}, & \text{if } p > h, \\ -(-1)^{P_I[h-1] \oplus P_I[p-1]}, & \text{if } h > p, \end{cases} \tag{5.28}$$

where $\oplus$ denotes the exclusive or (*xor*) operation. So if the phase mask is available, the computation of the phase factor only takes a few CPU cycles. Another important aspect is to create efficiently the phase masks. We propose Algorithm 3, which computes it in a logarithmic time for groups of 64 MOs, taking advantage of the associativity of the exclusive *or* operator. Indeed, the "for" loop executes 6 cycles to update the mask for $2^6 = 64$ MOs.

Function `PhasemaskOfDet(`I`)`**:**

Data: I : 64-bit string representation of $|I\rangle$

Result: P : phase mask associated with $|I\rangle$, as a 64-bit string.

for $\sigma \in \{\uparrow, \downarrow\}$ **do**

$r \leftarrow 0$;

for $i \leftarrow 0, N_{int} - 1$ **do**

$P_\sigma[i] \leftarrow I_\sigma[i] \oplus (I_\sigma[i] \ll 1)$;

for $d \leftarrow 0, 5$ **do**

$P_\sigma[i] \leftarrow P_\sigma[i] \oplus (P_\sigma[i] \ll (1 \ll d))$;

end

$P_\sigma[i] \leftarrow P_\sigma[i] \oplus r$;

if $(\|I_\sigma[i]\| \wedge 1) = 1$ **then**

$r \leftarrow \neg r$;

end

end

end

return P ;

$\|I\|$	: number of bits set to 1 in I (*popcnt*),
$\wedge$	: bitwise *and*,
$\oplus$	: bitwise *xor*,
$(I \ll k)$	: shift I by k bits to the left,
$\neg$	: bitwise negation.

Algorithm 3: Function that returns a phase mask as a bitstring.

5.12.3 Davidson diagonalization

Within QUANTUM PACKAGE, the Davidson diagonalization algorithm is implemented in its multi-state version. Algorithmically, the expensive part of the Davidson diagonalization is the computation of the matrix product $\mathbf{H}\,\mathbf{U}$. As mentioned above (see Sec. 5.3), two determinants $|I\rangle$ and $|J\rangle$ are connected via $\mathbf{H}$ (i.e. $\langle I|\hat{H}|J\rangle \neq 0$) only if they differ by no more than two spinorbitals. Therefore, the number of non-zero elements per row in $\mathbf{H}$ is equal to the number of single and double excitation operators, namely $\mathcal{O}(N_\uparrow^2(N_{\text{orb}} - N_\uparrow)^2)$. As $\mathbf{H}$ is symmetric, the number of non-zero elements per column is identical. This makes $\mathbf{H}$ very sparse. However, for large basis sets, the whole matrix may still not fit in a single node memory, as the number of non-zero entries to be stored is of the order of $N_{\text{det}}N_\uparrow^2(N_{\text{orb}} - N_\uparrow)^2$. One possibility would be to distribute the storage of $\mathbf{H}$ among multiple compute nodes, and use a distributed library such as PBLAS[283] to perform the matrix-vector operations. Another approach is to use a so-called *direct* algorithm, where the matrix elements are computed *on the fly*, and this is the approach we have chosen in QUANTUM PACKAGE. This effectively means iterating over all pairs of determinants $|I\rangle$ and $|J\rangle$, checking whether $|I\rangle$ and $|J\rangle$ are connected by $\mathbf{H}$ and if they are, accessing the corresponding integral(s) and computing the phase factor. Even though it is possible to compute the excitation degree between two determinants very efficiently,[281] the number of such computations scales as N_{det}^2, which becomes rapidly prohibitively high. To get an efficient determinant-driven implementation it is mandatory to filter out all pairs of determinants that are not connected by $\mathbf{H}$, and iterate only over connected pairs. To reach this goal, we have implemented an algorithm similar to the *Direct Selected Configuration Interaction Using Strings* (DISCIUS) algorithm.[187]

The determinants of the internal space are re-ordered in linear time as explained in Ref. [193], such that the wave function can be expressed as

$$|\Psi^{(0)}\rangle = \sum_{I}^{N_{\text{det}}^{\uparrow}} \sum_{J}^{N_{\text{det}}^{\downarrow}} C_{IJ}\,|I_\uparrow J_\downarrow\rangle\,, \tag{5.29}$$

where we take advantage of the Waller-Hartree double determinant representation.[235]

Moving along a row or a column of $\mathbf{C}$ keeps the spin-up or spin-down determinants fixed, respectively. For a given determinant, finding the entire list of same-spin single and

double excitations can be performed in $\mathcal{O}(N_{\text{det}}^{\uparrow}) = \mathcal{O}(N_{\text{det}}^{\downarrow}) = \mathcal{O}(\sqrt{N_{\text{det}}})$, while finding the opposite-spin double excitations is done via a two-step procedure. First, we look for all the spin-up single excitations. Then, starting from this list of spin-up single excitations, we search for the spin-down single excitation such that the resulting opposite-spin doubly-excited determinant belongs to $\Psi^{(0)}$. Hence, the formal scaling is reduced to $\mathcal{O}(N_{\text{det}}^{3/2})$. It could be further reduced to $\mathcal{O}(N_{\text{det}})$ at the cost of storing the list of all singly- and doubly-excited determinants for each spin-up and spin-down determinant, but we preferred not to follow this path in order to reduce the memory footprint as much as possible.

5.12.4 CIPSI selection and PT2 energy

There are multiple ways to compute the $e_\alpha^{(2)}$'s. One way is to loop over pairs of internal determinants $|I\rangle$ and $|J\rangle$, generate the list of external determinants $\{|\alpha\rangle\}$ connecting $|I\rangle$ and $|J\rangle$ and increment the corresponding values $e_\alpha^{(2)}$ stored in a hash table. Using a hash table to store in memory a list of $|\alpha\rangle$'s without duplicates and their contributions $e_\alpha^{(2)}$ is obviously not a reasonable choice since the total number of $|\alpha\rangle$'s scales as $\mathcal{O}(N_{\text{det}} N_\uparrow^2 (N_{\text{orb}} - N_\uparrow)^2)$. To keep the memory growth in check, we must design a function that can build a stream of unique external determinants, compute their contribution $e_\alpha^{(2)}$ and retain in memory only the few most significant pairs $(|\alpha\rangle, e_\alpha^{(2)})$.

In Quantum Package, we build the stream of unique external determinants as follows. We loop over the list of internal determinants (the *generators*) sorted by decreasing c_I^2. For each generator $|I\rangle$, we generate all the singly- and doubly-excited determinants $\{|\alpha\rangle\}$, removing from this set the internal determinants and the determinants connected to any other generator $|J\rangle$ such that $J < I$. This guarantees that the $|\alpha\rangle$'s are considered only once, without any additional memory requirement.

For each generator $|I\rangle$, before generating its set of $|\alpha\rangle$'s, we pre-compute the diagonal of the Fock matrix associated with $|I\rangle$. This enables to compute the diagonal elements $\langle\alpha|\hat{H}|\alpha\rangle$ involved in Eq. (5.8) for a few flops.[284] The computation of $\langle\Psi^{(0)}|\hat{H}|\alpha\rangle = \sum_J c_J \langle J|\hat{H}|\alpha\rangle$ is more challenging than the diagonal term since, at first sight, it appears to involve the N_{det} internal determinants. Fortunately, most of the terms amongst this sum vanish due

to Slater-Condon's rules. Indeed, we know that the terms where $|J\rangle$ is more than doubly excited with respect to $|\alpha\rangle$ vanish, and these correspond to the determinants $|J\rangle$ which are more than quadruply excited with respect to $|I\rangle$.[284] To compute efficiently $\langle\Psi^{(0)}|\hat{H}|\alpha\rangle$, for each generator $|I\rangle$, we create a filtered wave function $|\Psi_I^{(0)}\rangle$ by projecting $|\Psi^{(0)}\rangle$ on a subset $\mathcal{J}_I$ of internal determinants $\{|J\rangle\}$ where $\langle J|\hat{H}|\alpha\rangle$ is possibly non-zero. This yields $\langle\Psi^{(0)}|\hat{H}|\alpha\rangle = \langle\Psi_I^{(0)}|\hat{H}|\alpha\rangle$, where $\Psi_I^{(0)}$ is a much smaller determinant expansion than $\Psi^{(0)}$. In addition, as we have defined the $|\alpha\rangle$'s in such a way that they do not interact with $|J\rangle$ when $J < I$, all these $|J\rangle$'s can also be excluded from $\mathcal{J}_I$. This pruning process yielding to $|\Psi_I^{(0)}\rangle$ will be referred to as the *coarse-grained* filtering. A *fine-grained* filtering of $|\Psi_I^{(0)}\rangle$ is performed in a second stage to reduce even more the number of determinants, as we shall explain later.

To make the coarse-grained filtering efficient, we first filter out the determinants that are more than quadruply excited in the spin-up and spin-down sectors separately. Using the representation shown in Eq. (5.29), this filtering does not need to run through all the internal determinants and scales as $\mathcal{O}(N_{\text{det}}^{\uparrow}) = \mathcal{O}(\sqrt{N_{\text{det}}})$. It is important to notice that, at this stage, the size of $\mathcal{J}_I$ is bounded by the number of possible quadruple excitations in both spin sectors, and does not scale any more as $\mathcal{O}(N_{\text{det}})$. Next, we remove the determinants that are i) quadruply excited in one spin sector and excited in the other spin sector, ii) triply excited in one spin sector and more than singly excited in the other spin sector, and iii) all the determinants that are doubly excited in one spin sector and more than doubly excited in the other spin sector.

The external determinant contributions are computed in batches. A batch I_{pq} is defined by a doubly-ionized generator $|I_{pq}\rangle = a_p a_q |I\rangle$. When a batch is created, the fine-grained filtering step is applied to $\mathcal{J}_I$ to produce $\mathcal{J}_{I_{pq}}$ and $\Psi_{I_{pq}}^{(0)}$, such that $\langle\Psi_{I_{pq}}^{(0)}|\hat{H}|\alpha\rangle = \langle\Psi_I^{(0)}|\hat{H}|\alpha\rangle$.

Each external determinant produced in the batch I_{pq} is characterized by two indices r and s with $\hat{\mathcal{O}} a_r^\dagger a_s^\dagger a_p a_q |I\rangle = |I_{pq}^{rs}\rangle$. The contribution associated with each determinant of a given batch will be computed incrementally in a two-dimensional array $A(r, s)$ as follows. A first loop is performed over all the determinants $|J\rangle$ belonging to the filtered internal space $\mathcal{J}_{I_{pq}}$. Comparing $|J\rangle$ to $|I_{pq}\rangle$ allows to quickly identify if $|J\rangle$ will be present in the list of external determinants, and consequently tag the corresponding cell $A(r, s)$ as *banned*. Banned

cells will not be considered for the computation of $e_\alpha^{(2)}$ nor the determinant selection, as they correspond to determinants already belonging to the internal space. A second loop over all the $|J\rangle \in \mathcal{J}_{I_{pq}}$ is then performed. During this loop, all the (r,s) pairs where $|I_{pq}^{rs}\rangle$ is connected to $|J\rangle$ are generated, and the corresponding cells $A(r,s)$ are incremented with $c_J \langle J|\hat{H}|I_{pq}^{rs}\rangle$. After this second loop, $A(r,s) = \langle\Psi|\hat{H}|I_{pq}^{rs}\rangle$ and all the contributions $e_\alpha^{(2)}$ of the batch can be obtained using $A(r,s)$. The running value of $E^{(2)}$ is then incremented, and the N_{det} most significant determinants are kept in an array sorted by decreasing $|e_\alpha^{(2)}|$.

Figure 23 shows the number of determinants retained in $\Psi_I^{(0)}$ or $\Psi_{I_{pq}}^{(0)}$ after filtering out disconnected determinants of the ground state of the CN3 molecule with 935 522 determinants (see Sec. 5.12). This example shows that, starting from $\Psi^{(0)}$, the coarse-grained process which consists of removing the determinants more than quadruply excited with respect to the generator $|I\rangle$ produces wave functions $\Psi_I^{(0)}$ with a typical size of 120 000 determinants, a reduction by a factor 8. Then, starting from $\Psi_I^{(0)}$, the fine-grained filtering, specific to the batch generating $\Psi_{I_{pq}}^{(0)}$, reduces even more the number of determinants (by a factor 3), down to a typical size of 40 000 determinants, which represents only 4% of the total wave function $\Psi^{(0)}$.

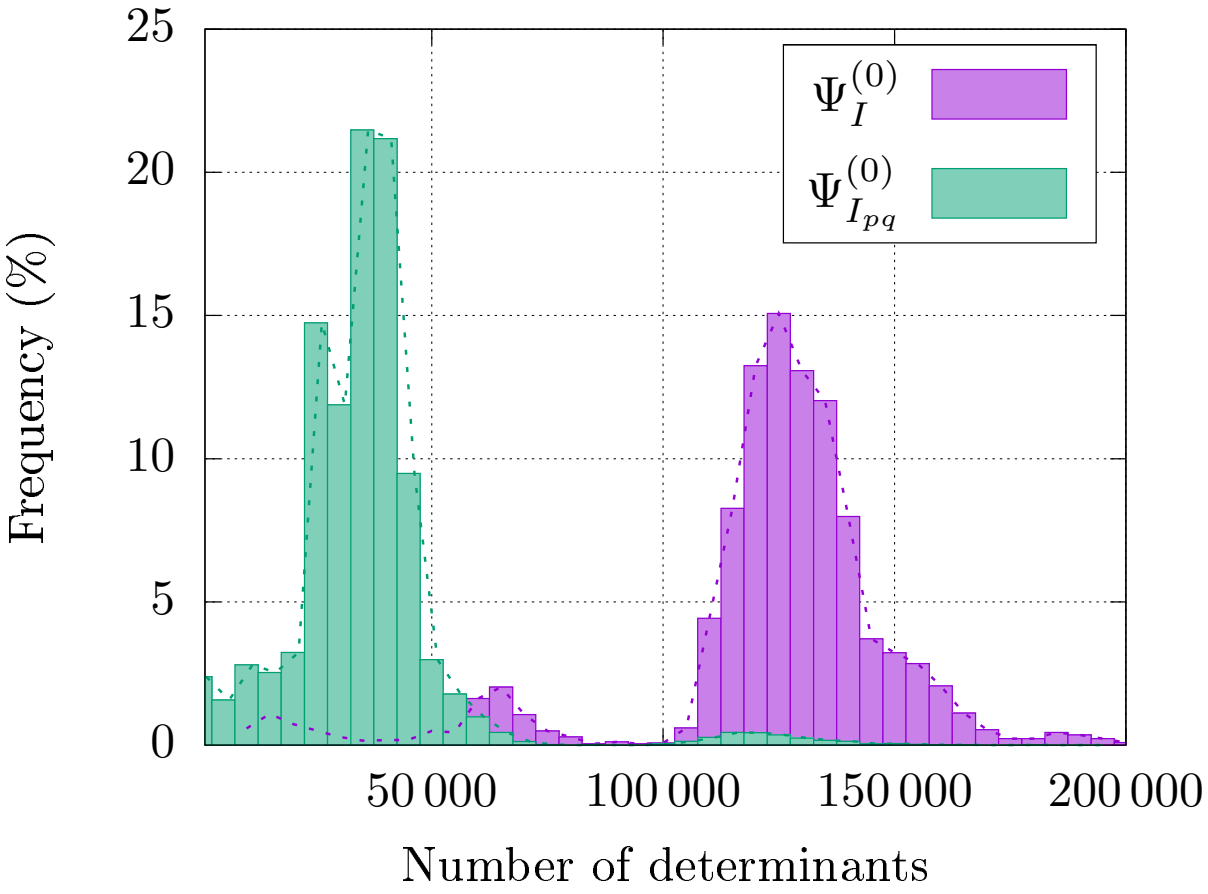

Figure 23: Histograms representing the number of determinants remaining after the coarse-grained (purple) and fine-grained (green) filtering processes applied to the ground state of the CN3 molecule with $N_{\text{det}} = 935\,522$.

Appendix A Strategy for Creating Rational Fraction Fits to Stabilization Graph Data on Metastable Electronic States

The text and figures in this chapter have been reprinted with permission from Gasperich, K.; Jordan, K. D.; Simons, J. Strategy for creating rational fraction fits to stabilization graph data on metastable electronic states. *Chem. Phys.* **2018**, *515*, 342–349, DOI: 10.1016/j.chemphys.2018.07.019. The author's contribution to the work included writing code to generate data, generating/editing figures, and editing/revising the manuscript.

A.1 Summary

An exactly soluble model of two diabatic electronic states interacting through a coupling of strength V is used to generate data for testing the rational fraction analytic continuation technique for determining the energies and widths of metastable states of anions. By making analytical connections between the coefficients in the rational fraction and the parameters of the model, we are able to suggest how to choose the orders of the polynomials and the range of the scaling parameter, Z, within which to compute the energies for a given precision. This analysis also allows us to specify the range of Z-values to use in constructing the rational fraction in a manner that allows determination of all parameters of the model for a given precision. The constraint on the Z-value ranges can be used as a guide for constructing rational fractions of data obtained in electronic structure calculations on actual resonance states.

A.2 Introduction

When treating metastable electronic states of atomic and molecular anions, the stabilization technique introduced by Hazi and Taylor[286] has proven to be very useful. In its most commonly employed form, the energies of several electronic states of the excess-electron system are computed for a range of values of a parameter (Z) that controls the radial extent of the basis functions used in the calculations.[287] These energies are then plotted as functions of Z to generate stabilization plots such as that shown in Fig. 24.

Such stabilization plots typically display three characteristics that merit attention:

1. One or more plateau regions within which the energy of one of the branches changes slowly as the scaling parameter is varied. In Fig. 24, such plateaus occur at energies near 1.5 and 4.1 eV. The energies of the plateau regions approximate the energies of the metastable states being studied.
2. A series of states whose energies change more rapidly as the scaling parameter is varied; these energies describe pseudo-continuum states that correspond to a "free" particle in a pseudo-continuum orbital. In Fig. 24 the energies of these states increase with the scaling parameter Z, which controls the radial extent of the basis set.
3. As Z is varied, one encounters regions where two types of states approach one another and undergo avoided crossings. The regions of these avoided crossings play a central role in determining the lifetime of the metastable state. In Fig. 24 we see that the plateau regions are interrupted by a series of avoided crossing thus limiting the range of Z-values over which any given plateau persists.

There is another class of stabilization methods that involves adding a stabilizing potential that converts the resonance into a bound state followed by analytically continuing the bound-state energy into the resonance region;[288] however, we do not consider these approaches here.

The example illustrated in Fig. 24 shows how the energies of several excess-electron states vary as the scaling parameter is changed. Electronic structure methods such as configuration interaction, equation-of-motion coupled cluster (EOM-CC),[289,290] Koopmans' theorem,[291]

and many-body Green's functions[292] can be used to extract multiple roots for constructing stabilization plots. Given the energies of several roots as functions of the scaling parameter, it is then usually relatively straightforward to identify regions of avoided-crossings within a range of Z-values. The rational fraction (RF) method, which is the subject of this study, is designed to fit the energy of a single root of the stabilization calculation as a function of the scaling factor, and it is usually a root whose energy lies within a plateau region that is used.[293]

A.3 Extracting the energy and lifetime of the metastable state from a stabilization plot

A.3.1 RF and quadratic equation approaches for fitting stabilization-plot data

In the Siegert picture,[294] a metastable state, also called a resonance, is associated with a complex energy $E_R - i\Gamma/2$, where E_R is the resonance position and Γ the width is proportional to the reciprocal of the lifetime. This complex energy, when substituted into $e^{-iEt/\hbar}$, describes a state that decays in time. Correspondingly, one can view such a state as having an energy uncertainty (or width) Γ. The resonance parameters E_R and Γ can be obtained by analytically continuing the energy as a function of Z into the complex plane, locating the stationary points Z_{sp} where $\partial E/\partial Z = 0$, and then evaluating E at Z_{sp}.[295] In the RF method, analytic continuation is performed after using computed energy values to construct a rational fraction:

$$E(Z) = \frac{N(Z)}{D(Z)} \tag{A.1}$$

where

$$N(Z) = \sum_{j=0}^{n} n_j Z^j \tag{A.2}$$

and

$$D(Z) = \sum_{j=0}^{d} d_j Z^j \tag{A.3}$$

When the coefficients in the numerator and denominator are determined from the coefficients of a Taylor series expansion of a function, the RF is also referred to as a Padé

approximant;[296] when the coefficients are determined using numerical data giving E at various Z-values, the term RF is preferred.

The $[n, d]$ approximant has $n + d + 2$ parameters; however, only $n + d + 1$ of these are independent, because the energy depends on the ratio N/D rather than on N and D individually. Often, one opts to set $d_0 = 1$, but other choices are possible. After using the computed energies to determine the values $\{n_j, d_j\}$ of the expansion coefficients, the derivative of the energy with respect to Z is evaluated and set equal to zero,

$$\frac{\partial E}{\partial Z} = 0 \tag{A.4}$$

Solving Eq. (A.4) for the Z-values at which this equation holds gives the complex stationary points Z_{sp}, which are then substituted into the rational fraction expression to generate complex stationary energies

$$E_{sp} = E_R \pm i\Gamma/2 \tag{A.5}$$

Although the primary goal of this paper is to analyze the RF method, it is useful to also consider the alternative quadratic equation (QE) approach,[287,295,297] in which one introduces the following expression for how the energy E varies with the scaling parameter Z:

$$P(Z)\left[E\left(Z\right)\right]^2 + Q(Z)E(Z) + R(Z) = 0 \tag{A.6}$$

where P, Q, and R are polynomials in Z:

$$P(Z) = \sum_{j=0}^{p} p_j Z^j \tag{A.7}$$

$$Q(Z) = \sum_{j=0}^{q} q_j Z^j \tag{A.8}$$

$$R(Z) = \sum_{j=0}^{r} r_j Z^j \tag{A.9}$$

This expression has $p + q + r + 3$ total parameters, but, as with the RF, the energy is unchanged when all polynomials in Z are scaled by a constant factor; therefore, the number

of independent parameters is $p+q+r+2$, and a common choice is to set $p_0 = 1$. After the polynomial coefficients $\{p_j, q_j, r_j\}$ are determined by fitting, one solves for E, obtaining

$$E(Z) = -\left[\frac{Q(Z)}{2P(Z)}\right] \pm \sqrt{\left[\frac{Q(Z)}{2P(Z)}\right]^2 - \frac{R(Z)}{P(Z)}} \tag{A.10}$$

The derivative of the energy expression in Eq. (A.10) with respect to Z is then set to zero to determine (complex) values of Z at which $E(Z)$ is stationary, which are then substituted into Eq. (A.10) to generate the complex energies associated with the stationary points, yielding $E_{sp} = E_R \pm i\Gamma/2$ as discussed above. In general, the stationary points Z_{sp} associated with a resonance arising from a pair of coupled diabatic states are not far from the complex branch points associated with the avoided crossing between discrete and pseudo-continuum diabatic states.[295] The branch points of Eq. (A.10) occur at values of Z where $[Q(Z)]^2 - 4P(Z)R(Z) = 0$. The QE framework builds into its working equations the existence of branch points, whereas the RF method does not.

If one utilizes the same number of $E(Z_k)$ data points as one has parameters in either the RF or QE analytic continuation expression, one obtains a system of linear equations to be solved for the polynomial coefficients. In the RF approach, one can cast the problem in the form of a continued fraction, which allows the coefficients to be determined by a recursion relation.[298] Alternatively, one can employ more data points than parameters and use a least-squares procedure to optimize the parameters. The details of how one fits the calculated energy values to either Eq. (A.1) or Eq. (A.6) will not be further discussed in this work; rather our focus will be on how to determine optimal ranges of Z_k values used to compute the energies used in the fits and what order of polynomials should be used in the RF fits.

A.3.2 Selecting data points for RF fits that are not too far from avoided crossings

Pairs of diabatic states of the same symmetry that cross as the scale parameter is varied undergo avoided crossings when they are allowed to interact as shown in Fig. 24, resulting in adiabatic energies that display complex branch points. This behavior is the primary

motivation for introducing the QE analytic continuation procedure. In the QE approach, the data for the fitting generally employs data points from the vicinity of an avoided crossing but may also include values more distant from the avoided crossing; moreover, in the QE approach, the data points can be chosen from a single branch or from both branches involved in the chosen avoided crossing. RFs of the form given in Eq. (A.1) do not properly describe the branch points but generally have poles and zeros at Z-values close to where the diabatic states cross. If one were performing the analytic continuation using a simple power series, one would then have to avoid data points "close" to the crossing point of the diabatic curves because such points might be outside the radius of convergence of the series. With RFs, this is less of an issue as convergence can be achieved even when using data points close to the crossing region, although the inclusion of such points may slow down the rate of convergence, and, in practice, one often avoids using such data points.

As we illustrate later, the RF approach will not be able to accurately describe the resonance if one only uses data points from a stabilization plot that are "far" from the avoided crossing. In that case, the $E(Z)$ vs. Z data contain too little information about the strength of the coupling between the two diabatic states. A main goal of this work is to provide a path by which one can estimate how close to the crossing point one must include $E(Z)$ data points given the precision to which one knows the Z-variation of the energies contained in the stabilization plot. Alternatively, we show to what precision one must, if feasible, determine the E-values for a given choice of Z-values.

A.4 Model for which the exact energy and width are known

A.4.1 What is the purpose of introducing an analytically solvable model?

We use $E(Z)$ vs. Z data generated from a model's exact solution and from expansions of the model's exact solution valid through various orders in the coupling strength V to illustrate the problems that arise if one employs data points too far from a crossing point in forming a RF. We provide explicit formulas, in terms of the model's parameters, for the

ranges of Z within which data should be calculated given the precision ε to which variations in the energies $E(Z)$ are known as Z varies and given the order in V to which one wishes to determine the resonance state's width.

We suggest that the lessons learned from testing RF methodology on this exactly soluble model can be applied to ab initio electronic structure stabilization graphs. In particular, by using data from an ab initio stabilization graph's plateau-region and from its region approaching an avoided crossing to make connections to the model's parameters, the analytical expressions obtained for the model can be used to estimate the range of Z-values to use in creating an RF fit to the ab initio data.

A.4.2 The model energy expression and its resonance energy and width

The avoided crossings that pairs of diabatic states undergo can be qualitatively described using a two-state Hamiltonian matrix whose diagonal elements H_{11} and H_{22} describe the energies of the diabatic states as functions of the scaling parameter and whose off-diagonal element V describes the coupling. The two eigenvalues of the resulting matrix are given by $\frac{1}{2}\{(H_{11}+H_{22}) \pm \sqrt{4V^2+(H_{11}-H_{22})^2}\}$. Distant from an avoided crossing, the energies of the diabatic states generally vary monotonically with the scaling factor Z, which suggests that H_{11} and H_{22} can be represented as low-order polynomials in Z.

The most elementary reasonable model[299] of a stabilization plot's avoided crossing region assumes two diabatic states whose energies vary linearly (in the region of their crossing) with the scaling parameter Z

$$H_{11} = -b_1 + a_1 Z \tag{A.11a}$$

$$H_{22} = -b_2 + a_2 Z \tag{A.11b}$$

These diabatic states intersect at the point

$$Z_0 = \frac{b_2 - b_1}{a_2 - a_1} \tag{A.12}$$

where their common energy is

$$H_{11}(Z_0) = -b_1 + a_1 Z_0 = H_{22}(Z_0) = -b_2 + a_2 Z_0 = E^0 \tag{A.13}$$

The parameter Z could be the factor by which selected diffuse atomic basis functions are scaled. Alternatively, it could be $(1/R^2)$, where R is the radius of a spherical box within which continuum radial basis functions are constrained. In any case, it is best to define Z in a manner that makes the Z-dependence of the diabatic pseudo-continuum states as linear as possible.

The energy of the diabatic discrete state (here designated as H_{11}) would be expected to be independent of or only weakly dependent on Z. However, the overlap between the pseudo-continuum basis functions and the discrete state can introduce a Z-dependence to H_{11}. Assuming the two orthogonalized diabatic states couple with an off-diagonal Hamiltonian matrix element, V, solution of the associated 2×2 secular equation gives the expression

$$E = E^0 + a(Z - Z_0) \pm \sqrt{V^2 + \left[\frac{\delta a}{2}(Z - Z_0)\right]^2} \tag{A.14}$$

where $a = \frac{a_1 + a_2}{2}$ and $\frac{\delta a}{2} = \frac{a_2 - a_1}{2}$.

As can be seen from Fig. 24, the diabatic states that undergo an avoided crossing in a stabilization plot do not rigorously vary linearly with the scaling parameter; moreover, although we take V to be constant within our model, in general it will depend on Z, because of the Z-dependence in the coupling between the diabatic states and the impact of the overlap contribution.[300] For these reasons, the analytical results obtained here are certainly approximate representations of stabilization graphs from electronic structure calculations. Our analysis could readily be extended to treat cases in which V depends on Z and the diabatic states' energies vary non-linearly with Z; however, here we will limit most of our discussion to the simplest case in which the diabatic energies (accounting for overlap between the discrete and pseudo-continuum states) are assumed to vary linearly and V is assumed to be constant.

In this paper, we use Eq. (A.14) to generate values of $E(Z_k)$ to use as input data for Eq. (A.1) to subsequently determine the energies and widths of the resonance. We do so for three sets of parameters describing resonances with widths differing by a factor of 10 and resonances with clear plateaus and one in which the plateau has a substantial slope. We will refer to energies computed from Eq. (A.14) as the exact energies for the model problem. We suggest that employing Eq. (A.14) to generate "test data" to use in Eq. (A.1) can provide

valuable insight into the performance of the RF method for different choices of input data because plots of energies obtained from Eq. (A.14) display the essential characteristics of actual stabilization plots. Moreover, as we illustrate later, the functional form given in Eq. (A.14) can offer guidance about what powers of Z to use and what range of Z-values to use in forming an effective RF. We suggest that any RF whose polynomials do not contain at least these minimum powers of Z or that do not use data from the recommended range of Z-values will not only fail to give accurate resonance energies and widths for the model problem used here but will also fail when applied to stabilization plot data for real chemical systems.

The exact stationary points for the above model are

$$Z_{sp} = Z_0 \pm 2iV\frac{a}{\delta a\sqrt{a_1 a_2}} \tag{A.15}$$

with the associated energies being

$$E_{sp} = E^0 \pm 2iV\frac{\sqrt{a_1 a_2}}{\delta a} \tag{A.16}$$

The branch points for the model occur at $Z_{bp} = Z_0 \pm 2i\frac{V}{\delta a}$; hence the stationary points lie further off the real axis than the branch points by a factor of $a/\sqrt{a_1 a_2}$.

For the remainder of this paper, we will assume that

(i) $|a_2| > |a_1|$

(ii) a_2 and a_1 have the same sign

(iii) data from only the branch having the smaller slope (i.e., the plateau branch with slope a_1) is being used to generate the $E(Z_k)$ data employed in the RF analysis.

A similar analysis could be carried out using data from the branch having the larger slope. Moreover, simply for convenience, we will assume that the $\{Z_k\}$ values are selected to the right of the crossing point Z_0 so that all $\delta Z_k = Z_k - Z_0$ values are positive.

A.4.3 Guidance offered by the model on how to select powers of Z and Z-values at which to compute energies

The first thing to point out is that Eq. (A.14) contains five parameters (E^0, Z_0, V, a, and δa). Together, the expression for the exact stationary points, Eq. (A.15), and that for the corresponding resonance energy, Eq. (A.16), require knowledge of all five of these parameters. This suggests that to accurately predict Z_{sp} and E_{sp}, any reasonable RF fit should contain at least five parameters.

As discussed earlier, applications[293] of the RF approach generally utilize $E(Z_k)$ energies at Z_k values chosen distant from the avoided crossing region of the stabilization plot to avoid approaching the branch points. As we make more quantitative below, when forming a RF utilizing only Z-values that are far from the crossing point Z_0 the $E(Z_k)$ data might not be known with sufficient precision to accurately characterize the stationary point. Although most ab initio electronic structure calculations are performed using double precision arithmetic, issues such as the tolerance to which one converges matrix eigenvalues limit the final precision of the stabilization-plot energy data. Based on our experience, a precision of ca. 10^{-5} eV is a reasonable estimate and one that we use in this paper.

As noted above, RFs with coefficients determined from fitting data points are closely related to PAs where the coefficients are determined by reproducing a fixed number of terms in a power series expansion about an appropriate point. In particular, RFs can be viewed as employing coefficients that correspond to use of derivatives evaluated by numerical differentiation. For that reason, we find it useful to expand Eq. (A.14) in a power series about a point Z' chosen to be located approximately in the middle of the set of grid points employed in the RF fit. With the choices of grid points describe above, this necessarily locates Z' to the right of Z_0 (i.e., $Z' > Z_0$). The resulting series expansion through terms of

order $(Z-Z')^4$ is:

$$\begin{aligned} E =& E^0 + a\Delta Z \pm K + \left(a \pm \frac{\left(\frac{\delta a}{2}\right)^2 \Delta Z}{K} \right) (Z - Z') \pm \\ & \frac{\left(\frac{\delta a}{2}\right)^2 V^2 (Z-Z')^2}{2K^3} \mp \frac{\left(\frac{\delta a}{2}\right)^4 V^2 \Delta Z (Z-Z')^3}{2K^5} \mp \\ & \frac{\left(\frac{\delta a}{2}\right)^4 V^2 \Delta Z (5V^2 - 4K^2)(Z-Z')^4}{8K^7} + \ldots \end{aligned} \tag{A.17}$$

where $K = \sqrt{V^2 + \left(\frac{\delta a}{2}\right)^2 \Delta Z^2}$ and $\Delta Z = Z' - Z_0$.

For addressing the question concerning the location of Z', rather than the distribution of data points around Z', we need only substitute Z' into Eq. (A.14) giving

$$E = E^0 + a\Delta Z \pm \sqrt{V^2 + \left(\frac{\delta a}{2}\right)^2 \Delta Z^2} \tag{A.18}$$

If Z' obeys $\left|\frac{2V}{\delta a \Delta Z}\right| < 1$, which it will for points within a plateau region, one can estimate the contributions to E at various powers of V by expanding Eq. (A.18) as

$$E = E^0 + \left(a \pm \frac{\delta a}{2} \right) \Delta Z \pm \frac{V^2}{\delta a \Delta Z} \mp \frac{V^4}{\delta a^3 \Delta Z^3} \pm \ldots \tag{A.19}$$

For the root of Eq. (A.14) having the smaller (plateau) slope a_1 at large-Z this becomes

$$E = E^0 + a_1 \Delta Z - \frac{V^2}{\delta a \Delta Z} + \frac{V^4}{\delta a^3 \Delta Z^3} - \ldots \tag{A.20}$$

This allows us to specify how close Z' must be to Z_0 (i.e., how small ΔZ must be) for terms proportional to V^2 or V^4 to exceed the precision ε to which the electronic structure energies have been computed. In particular, to accurately determine the $\frac{V^2}{\delta a \Delta Z}$ term in the series expansion requires that $\Delta Z < \frac{V}{\delta a}\frac{V}{\varepsilon}$, which is likely achievable in most stabilization calculations as we illustrate later. The next two terms in the energy expansion are $\frac{V^4}{\delta a^3 \Delta Z^3}$ and $\frac{2V^6}{\delta a^5 \Delta Z^5}$. The V^4 and V^6 terms exceed ε in magnitude when $\Delta Z < \frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)^{\frac{1}{3}}$ and $\Delta Z < \frac{V}{\delta a}\left(\frac{2V}{\varepsilon}\right)^{\frac{1}{5}}$, respectively. Later we will show that selecting data in ranges that satisfy $\Delta Z < \frac{V}{\delta a}\frac{V}{\varepsilon}$ is usually straightforward, but to select data that satisfy the V^4 condition $\Delta Z < \frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)^{\frac{1}{3}}$ can be challenging, and to satisfy the V^6 condition is even more so.

A.4.4 Stationary points and energies from the series expansion

The energy expression given by Eq. (A.17) when truncated at order V^2 has the stationary points

$$Z_{sp} = Z_0 \pm \frac{iV}{\sqrt{a_1 \delta a}} \tag{A.21}$$

with the corresponding energy

$$E_{sp} = E^0 \pm 2iV\sqrt{\frac{a_1}{\delta a}} \tag{A.22}$$

Note that even if one knows a_1 from large-Z results, these expressions for E_{sp} and Z_{sp} do not allow one to extract individual values for V or a_2.

If we consider the expansion in Eq. (A.17) through order V^4, we find

$$Z_{sp} = Z_0 \pm \frac{iVJ}{\sqrt{a_1 \delta a}} \tag{A.23}$$

and

$$E_{sp} = E^0 \pm iV\sqrt{\frac{a_1}{\delta a}}\left(J + \frac{1}{J} + \frac{a_1}{\delta a J^3}\right) \tag{A.24}$$

where

$$J = \sqrt{\frac{1 \pm \sqrt{1 + 12\left(\frac{a_1}{\delta a}\right)}}{2}} \tag{A.25}$$

Notice that the V^4 expression for the width involves factors of a_1 and $\frac{V}{\sqrt{\delta a}}$ as well as an expression that depends on the ratio $\frac{a_1}{\delta a}$, so only by reaching the V^4 level in E_{sp} and Z_{sp} is one able to access all three of a_1, a_2, and V (assuming that a_1 is available from large-Z data). We will refer to the energy and half-width of Eq. (A.24) as the values through order V^4 in the expansion of the square root in Eq. (A.14).

A.4.5 Guidance for creating rational fractions

A major advantage of Padé approximants is that they provide approximations to higher order terms in the Taylor series that were not used in the fitting, and, as a result, they can often accurately represent the function at points more distant from the expansion point than can the original truncated Taylor series. However, in application to stabilization calculations, one uses numerical energy data from a grid of points rather than the coefficients of a Taylor series, giving rise to what are termed RFs.

When $Z' \gg Z_0$, and reasonable values are chosen for the various parameters in the model, it is found that the power of Z in the numerator of the PA is essentially one higher than that in the denominator even if one constructs a PA having a higher power denominator. This is illustrated below by examining the [2,2] PA of Eq. (A.17) for the case $Z' = 1.2$ and using the S3 set of (a_1, a_2, V, E^0, and Z_0) parameters that are defined in the next Section.

$$PA[2,2] = \frac{1.727 - 5.069Z - 21.337Z^2}{1 - 5.334Z - 0.165 \times 10^{-4}Z^2} \tag{A.26}$$

Examination of this PA reveals that the coefficient of the Z^2 term in the denominator is essentially 0, effectively reducing this to the [2,1] PA, which is reported along with the [1,1], and [3,2] PAs in the Supplementary Material. Similarly, the [3,3] PA is found to be essentially equivalent to the [3,2] PA. For this reason only $[n+1, n]$ RFs are considered in the subsequent discussion. We note that the conclusion about the power of Z in the numerator being one higher than that in the denominator would be altered were one or more of H_{11}, H_{22}, and V to assume more complicated Z dependencies than assumed here. Another observation to make is that the large-Z slopes of the [2,1] and [3,2] PAs are very close to the exact value of 4.0. This shows that these low-order PAs provide accurate values for the a_1 slope parameter of the model.

Using the exact energies of Eq. (A.14) as numerical input, we will form $[n+1, n]$ RFs of the form

$$\mathrm{RF}_1 = \frac{n_0 + n_1 Z + n_2 Z^2}{1 + d_1 Z} \tag{A.27}$$

and

$$\mathrm{RF}_2 = \frac{n_0 + n_1 Z + n_2 Z^2 + n_3 Z^3}{1 + d_1 Z + d_2 Z^2} \tag{A.28}$$

for the three sets of test data whose parameters are given in Table 7. The goal of this numerical experiment is to determine the choices of Z-values for which the stationary points Z_{sp} and resonance energies of RF_1 and RF_2 approach the V^2 or V^4 or exact values (of the model). Because RFs containing N parameters are designed to represent data better than Taylor series expansions containing N parameters, we do not expect RF_1 results to match V^2 results or RF_2 to match V^4. However, because RF_1 contains four parameters, as does the V^2-level expansion of Eq. (A.19), we expect the RF_1- and V^2-level results to be similar. Likewise, because RF_2 contains six parameters while the V^4-level expansion of Eq. (A.19) and the exact expression of Eq. (A.14) contain only five parameters, we expect RF_2 to be able to match or exceed the V^4-level results.

A.5 Comparing results of RF fits of model data to the exact, V^2, and V^4 results

We created three sets of test data (labeled S1-S3) by inserting three choices of parameters (E^0, Z_0, a_1, a_2, and V) into Eq. (A.14). In all cases, E_0 was taken to be 2.50 eV and Z_0 was set equal to 0.200. The energy ranges and coupling strengths V are typical of low-energy electronic shape resonances in atoms and molecules. In Table 7, we list the V, a_1, and a_2 parameters for each case and give the value of Z for which δZ equals $\frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)^{\frac{1}{3}}$ or $\frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)$ with the energy-precision parameter ε set equal to 10^{-5} eV.

The Z-values shown in the third column are the Z-value bounds beyond which the V^4 contribution to the energy falls below ε. Those shown in the fourth column are the bounds beyond which the V^2 contribution is below this same ε.

Fig. 25 depicts the exact solutions of Eq. (A.14) for the S1 parameter set together with two sets of data points considered in Sec. A.8. One set satisfies the V^4 bound, and the other does not. The figure also depicts the curves obtained from the RF_1 and RF_2 fits, respectively.

For any value of ε, the range of acceptable Z-values is much broader if one only wants to assure that the V^2 contribution exceeds ε. The very large values listed in the right column of Table 7 would likely never be realized in an ab initio stabilization plot since other avoided

crossings would constrain Z to smaller values. For example, in Fig. 24, the blue plateau region with energy near 0.45 eV exists only for Z-values between ca. 1 and 2 with the avoided crossing near $Z = 2$ providing the upper limit to Z. On the other hand, the data in the third column in Table 7 suggest that one has to be careful in selecting Z-values if results accurate to V^4 are required to characterize the resonance.

Let us now examine for these three data sets the performance of RF_1 and RF_2 using various choices for Z-values at which the energies are computed. For comparison, we list in Table 8 the half-widths obtained using Eqs. (A.22) and (A.24) that result from expansions of the square root factor through orders V^2 and V^4, respectively, and the exact values from Eq. (A.16).

The primary differences among cases S1-S3 are as follows:

(i) S1 produces a narrow resonance ($\Gamma/2 = 0.07$ eV) because it has both a small value for V and a large difference between a_1 and a_2

(ii) S2 produces a broad resonance ($\Gamma/2 = 0.70$ eV) because it has a large value for V (it has the same values for a_1 and a_2 as in S1)

(iii) S3 produces a broad resonance ($\Gamma/2 = 0.49$ eV) not because it has a large value for V (it has the same value as in S1) but because its slopes a_1 and a_2 do not differ much.

Because case S3 displays the largest differences among the V^2, V^4, and exact half-widths, it offers the best opportunity to highlight the interplay between the energy precision (ε) and the values of Z used to form the RF fit. For this reason, we will discuss case S3 in detail while placing analogous data for cases S1 and S2 in Sec. A.8.

A.5.1 Results of $[n+1, n]$ RF fits for S3

We now determine the extent to which the stationary points and resonance energies obtained by fitting numerical data from Eq. (A.14) to RF_1 do or do not reproduce the results of V^2 and the extent to which fitting numerical data to RF_2 can yield stationary points and resonance energies close to those of V^4. In Table 9 we show the results of forming an RF_1 using the four Z-values listed. From inspection of the stabilization plot (not shown), it was clear that $Z = 1.6$ is well within the near-linear region while $Z = 0.4$ is in a region of

significant curvature. The real and imaginary parts of the stationary point Z_{sp}, and the real and imaginary values of the resonance energy E_{sp} are indicated for three different choices of the precision (10^{-5}, 10^{-7}, and 10^{-12} eV) in the input data. In specifying the precision, we are indicating the number of figures retained to the right of the decimal point. Although in ab initio calculations the precision is likely to be limited to at most 10^{-5} eV, for our model we also report results for precisions of 10^{-7} and 10^{-12} eV to illustrate how the results would evolve if one had more precise data.

From Table 7 it is seen that the V^2 and V^4 bounds for this case occur at $Z = 500$ and 1.28 (for $\varepsilon = 10^{-5}$ eV). Because all four of the Z-values employed lie well within the V^2 bound, it is no surprise that the half-widths obtained at the three precision levels considered are very close to the V^2 value 0.2828 eV listed in Table 8. Although the [2,1] RF_1 fit does not yield an accurate value for the half-width, it does provide an accurate a_1 (4.0 as we pointed out earlier) and reasonably accurate values for Z_0 (0.18 compared to the exact 0.20) and E_0 (2.4 eV compared to the exact 2.5 eV).

In Table 10 we show the results of forming RF_2 (i.e., [3,2]) fits for the S3 parameter set using six Z-values ranging from 0.4 to 1.4; again, the smaller Z-values lie in the curved region of the stabilization plot while the larger Z-values lie in the near-linear portion of the stabilization plot.

Earlier we noted that the V^4 bound for a precision of 10^{-5} eV is 1.28. We see that even though five of the six Z-values used in forming this RF_2 fit are below 1.28, the half-width obtained using data points at the 10^{-5} precision level is essentially the same as the RF_1 (V^2) value. Even using [4,3], [5,4], or [6,5] RF fits with the above five Z-values below 1.28 together with additional Z-values above 1.28, at a precision of 10^{-5} eV, the same V^2 level half-width was obtained. This shows that it is not the level of the RF but the values of Z that prevent RF_2 from doing better than V^2 level with data at a precision of 10^{-5} eV. It also shows that one needs to have all six of the Z-values below or very near to the V^4 bound because when the precision is increased to 10^{-7} eV (where the V^4 bound is $Z = 5.20$), a half-width significantly better than the V^4 value (0.38 eV) and close to the exact value (0.49 eV) is achieved.

The inability of [3,2] (or higher) fits to achieve half-widths close to (or better than as one might expect for RF fits) the V^4 value with only five of six Z-values below the V^4 bound and

using data at 10^{-5} precision raises the issue that we address in the next Section —namely, how to improve on the choice of Z-values by using results from the RF fits to estimate the V^2 and V^4 bounds and to use these results to form more accurate fits.

A.5.2 How to improve the choice of Z-values to create better $[n+1, n]$ RF_n fits

As noted earlier, not all of the Z-values used to form the [3,2] and higher RF fits whose results are shown in Table 10 are below the V^4 bound for an energy precision of $\varepsilon = 10^{-5}\,\mathrm{eV}$. Moreover, any $[n+1, n]$ fit we tried using five Z-values below the V^4 bound did not improve the situation. This suggests that we need to focus on placing at least six data points at or below the V^4 bound. In the case of the S3 example, we know ahead of time where this bound is because we know the values of V and δa, and, as we showed earlier, these two quantities cannot be obtained from the results of the [2,1] fit. Because the [3,2] fit shown above did not improve on the [2,1] level half-width at a precision of $10^{-5}\,\mathrm{eV}$, we cannot use the results of this [3,2] fit to obtain these parameters. However, there are two routes through which an estimate of the V^4 bound can be made as we now demonstrate.

We know that a [2,1] (i.e., RF_1) fit using four Z-values within the wide range below the V^2 bound should be capable of yielding reasonable values for a_1, Z_0, E^0, and $V^2/\delta a$. The value of a_1 is easily obtained from the slope at large Z, which for the [2,1] RF is 4.0. Knowing a_1, $V^2/\delta a$ can be obtained from the V^2 expression for the half-width $2\sqrt{\frac{V^2 a_1}{\delta a}}$. Using the [2,1] results shown in Table 9 (at any of the levels of precision as they are all essentially the same), we can offer the following estimates: $Z_0 = 0.18$, $E^0 = 2.4\,\mathrm{eV}$, and $4a_1V^2/\delta a = (0.29\,\mathrm{eV})^2$, hence $V^2/\delta a = 5.3 \times 10^{-3}\,\mathrm{eV}$. However, we need one more piece of information (a_2) to estimate the V^4 and V^2 bounds.

If one has sufficient knowledge about the slope of the other diabatic state's energy far from Z_0, one can use that value as a_2 and the [2,1] value of $\frac{V^2}{\delta a}$ plus a_1 obtained from the large-Z slope to estimate the V^4 bound as $\delta Z = \frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)^{1/3}$. This route is the most straightforward and should be followed if a reasonable estimate of a_2 is available. However, in the absence of direct knowledge of a_2, another route must be found.

The fact that the [3,2] RF described in Table 10 as well as results from [4,3] and higher

RFs using the Z-values in Table 11 and higher Z-values did not improve the results suggests that one or more of the Z-values used are above the V^4 bound. We therefore replaced the six Z-values whose [3,2] results are shown in Table 10 by six new values ranging from Z =0.25–1.00. Doing so produced a new [3,2] RF_2 fit the results of which are shown in Table 11.

The first thing to notice is that the half-width obtained at a precision of 10^{-5} eV has moved from 0.29 eV (see Table 10) to 0.49 eV, close to the exact value. Even when we reduced the precision to 10^{-3} or 10^{-4} eV, the half-width changed significantly from the V^2 value of 0.29 eV. The good agreement between the width calculated using data at the 10^{-5} precision level and the exact value for the width is partially fortuitous, as seen by the sizable error in the location of the stationary point and in the calculated position of the resonance. Indeed, as shown in Table 11, the half-width drops down to about 0.44 when calculated using [3,2] fits at higher precision values. This indicates that for the S3 model one needs to accurately characterize contributions higher than V^4 to the energy in order to accurately characterize the resonance. Indeed [4,3] and [5,4] RF fits using appropriately placed Z-values and sufficiently high precision gave half-widths near the exact value of 0.49 eV.

Seeing that placing six Z-values sufficiently low causes the new [3,2] RF to produce a half-width significantly different from that of the [2,1] RF, we now use the ratio of the new [3,2] and [2,1] half-widths to estimate the V^2 and V^4 bounds and to then verify whether at least six data points fall below or near to the V^4 bound. To do so, we note that the exact half-width of the model is $2\sqrt{\frac{V^2 a_1 a_2}{\delta a^2}}$, so the ratio of the exact half-width to the [2,1] value should be $\sqrt{\frac{a_2}{\delta a}}$. Taking the [3,2] half-width of 0.44–0.49 eV obtained using the new set of data points as an estimate to the exact value and using 0.29 eV as the [2,1] half-width gives $\frac{a_2}{\delta a}$ =2.3–2.9, which allows us to solve for a_2 =6.1–6.9, so we estimate δa =2.1–2.9. Finally, using $V^2/\delta a = 5.3 \times 10^{-3}$, we obtain V =0.11–0.12 eV.

We can now make use of these approximations to a_1, a_2, and V, to estimate the V^2 and V^4 bounds, which can then be used to improve upon the choice of Z-values used to verify the validity of the new [3,2] or to create higher RF fits. For example, assuming that we have energy data whose variation is precise to $\varepsilon = 10^{-5}$ eV, we estimate these bounds to be (using the average of the parameter estimates given above, i.e., $V = 0.115$ eV and $\delta a = 2.5$),

we obtain $\delta Z = \frac{V^2}{\varepsilon \delta a} = 529$ (the exact value is 500), and $\delta Z = \frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)^{1/3} = 1.04$ (the exact value is 1.28). This suggests that the Z-values used to form the new [3,2] fit do lie below the V^4 bound. As commented on earlier, the V^2 bound is likely irrelevant in real stabilization calculations as additional avoided crossings will limit the scaling parameter; however, this path does offer a route for estimating the more important V^4 bound.

This process allows us to be confident that the final six lowest Z-values in Table 11 used to form the [3,2] and higher RF fits were located in a manner that allowed us to obtain V^4 quality (or higher) resonance energies and widths. It used information from the [2,1], [3,2], and higher fits obtained using Z-values some of which were above the V^4 bound to guide us toward using lower Z-values for an improved [3,2] fit. This resulted in a new [3,2] fit of sufficient accuracy to generate an estimated V^4 bound to offer valuable guidance about where to choose Z-values for forming a subsequent series of higher order RF fits.

In the Supplementary Material we present analogous discussions of how to find appropriate Z-values for forming [2,1] and [3,2] RFs for the S1 and S2 cases. As the reader will see, in the S2 case it became clear that one should not locate all of the Z-values too far below the V^4 bound for either the [3,2] or the [2,1] RF because doing so can cause the evaluation of the a_1 slope parameter to fail; one needs at least one data point in the nearly-linear region of the stabilization plot.

A.6 Conclusions and suggestions for application to ab initio stabilization plot data

A five-parameter model of a stabilization graph which is based on two diabatic states undergoing an avoided crossing, is used to generate test data (energy vs. a scaling parameter Z) for creating RF approximants. Our analysis allowed us to conclude that:

1. when forming a RF_1 approximant to achieve results accurate to order V^2, one needs to use four (or more, if least squares fitting to determine parameters) data points below $Z = Z_0 + \frac{V}{\delta a}\frac{V}{\varepsilon}$ where the second order contribution to the energy falls below the precision ε to which the energies are known (ε of ca. 10^{-5} eV range was assumed), but it is also

important to include at least one point in the quasilinear portion of the stabilization plot to extract the a_1 slope parameter;

2. when forming a RF_2 approximant to achieve results accurate to order V^4, one needs to use six (or more, if least squares fitting to determine parameters) data points below or very near $Z_0 + \frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)^{1/3}$ beyond which the fourth order contributions fall below ε, but again it is important to include at least one point in the quasi-linear region to extract a_1.

We also note that stabilization graphs for real molecular resonances often require fits with eight or more parameters to extract an accurate value of the resonance width.[300] This indicates the importance of higher-order V-dependence than V^4, which would force one to select data points even closer Z_0 than $Z_0 + \frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)^{1/3}$. For example, if terms of the order V6 were important, one would have to choose points at or below $Z_0 + \frac{V}{\delta a}\left(\frac{2V}{\varepsilon}\right)^{1/5}$, which would be Z =0.28, 1.48, and 0.56 for S1, S2, and S3, respectively. In our opinion, these facts argue in favor of using the QE-type methods rather than RFs that emphasize the quasilinear large-Z regions of stabilization graphs.

Based on these observations, we suggest a strategy to use in constructing a $[n+1, n]$ RF representation of data on a single branch of a stabilization plot involving ab initio data in a manner that begins with first identifying a quasi-linear plateau region. For such cases, one generally does not know ahead of time how to select points optimal for accurately determining the metastable state's energy and width because one does not know how close to the more curved region of the stabilization plot one must characterize to achieve results of reasonable accuracy. However, the results obtained here can provide guidance if the ab initio stabilization plot displays two essential features that our model relies upon —(i) a portion that varies approximately linearly with Z at large-Z (described in our model by the terms $E^0 + a_1(Z - Z_0)$) and (ii) a part (arising in our model from the term $\pm\sqrt{V^2 + \left[\frac{\delta a}{2}(Z - Z_0)\right]^2}$ with curvature that increases in magnitude as Z moves closer to the avoided-crossing point Z_0.

For a stabilization plot that shows these characteristics, we suggest the following pathway can allow one to confidently evaluate the reliability of a RF and of the location of its data points.

i. First, one should search the energy data set for a region sufficiently far from Z_0 for which the energy varies approximately linearly with Z. The slope in this region can be associated with a_1 of our model. We suggest looking at our discussion of the S2 case in the Supplementary Material where our first choice of Z-values did not meet this criterion, and consequently, reasonable values of the resonance parameters did not result.

ii. Then, one should examine the data set at smaller Z-values until finding a region where the data begin to deviate significantly from the near-linear form found in step i. Using at least one data point in the near-linear region and the remaining data points in the region of significant curvature, one can form a RF_1 approximant.

iii. Since the stationary points of RF_1 are expected to occur near the V^2 value of $Z_{sp} = Z_0 \pm \frac{iV}{\sqrt{a_1 \delta a}}$ where the complex energy is $E_{sp} = E^0 \pm 2iV\sqrt{\frac{a_1}{\delta a}}$, one can use Z_{sp} and E_{sp} to estimate three more model-system parameters (with a_1 having been estimated from the near-linear region's slope)

$$\frac{V^2}{\delta a} = a_1 \left[\mathrm{Im}\left(Z_{sp}\right)\right]^2 \tag{A.29}$$

$$E^0 = \mathrm{Re}(E_{sp}) \tag{A.30}$$

and

$$Z_0 = \mathrm{Re}(Z_{sp}) \tag{A.31}$$

This knowledge then allows one to compute $Z = Z_0 + \frac{V}{\delta a}\frac{V}{\varepsilon}$, which one can use to verify whether the four data points used to create the ab initio RF_1 lie below the V^2 bound. If not, it is recommended that one adjust the choice of Z-values to form a new RF_1. However, as noted earlier, it is likely that all Z-values between successive avoided crossings lie within this V^2 bound in ab initio stabilization graphs.

iv. If one has a reasonable estimate of a_2, one can multiply the V^2 width by the ratio $\sqrt{\frac{a_2}{a_2 - a_1}}$(see Eq. (A.16)) to obtain an estimate of the exact width. Moreover, knowing a_2, one has enough knowledge to evaluate the point $Z_0 + \frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)^{1/3}$ at which the V^4 contributions to the energy fall below ε, so one can estimate where to place data points in forming a subsequent RF_2 approximant.

v. If one does not have a good estimate for a_2, to form RF_2, one can search the energy data for a range even closer to the crossing point within which the energy deviates significantly from $E^0 + a_1 (Z - Z_0) - \frac{V^2}{\delta a (Z - Z_0)}$ (using the values of E^0, a_1, Z_0 and $V^2/\delta a$ obtained in step iii). One can then use at least six data points in the region where the data deviates from $E^0 + a_1 (Z - Z_0) - \frac{V^2}{\delta a (Z - Z_0)}$ (being careful to include one point as near as possible to the near-linear region) to form a RF_2 approximant. As outlined earlier, the ratio of the half-width obtained from the RF_2 to that from the RF_1 is approximately $\sqrt{\frac{a_2}{\delta a}}$, which allows one to estimate a_2.

Knowing all five parameters of a model derived from the ab initio RF_1 and RF_2 then allows one to compute the point at which the fourth order contributions to the energy fall below ε $Z_0 + \frac{V}{\delta a} \left(\frac{V}{\varepsilon}\right)^{1/3}$ and to thus verify whether all of the Z-values used to form the RF_2 approximant in step v lie below $Z_0 + \frac{V}{\delta a} \left(\frac{V}{\varepsilon}\right)^{1/3}$. Knowing even approximate values for $\frac{V}{\delta a} \frac{V}{\varepsilon}$ and (especially) $\frac{V}{\delta a} \left(\frac{V}{\varepsilon}\right)^{1/3}$ would allow one to wisely choose Z_k values in forming any higher-order $[n+1, n]$ RF approximant of the ab initio data, and it is likely that such higher-order RFs would then produce the most reliable E_{sp} and Z_{sp} values to use in ab initio determinations of resonance-state energies and lifetimes.

A.7 Acknowledgments

KG acknowledges the support of a graduate fellowship from the Pittsburgh Quantum Institute. This research was supported by grant number CHE1762337 from the U.S. National Science Foundation. We also acknowledge the use of computational resources in the University of Pittsburgh's Center for Research Computing.

A.8 Supplementary material

A.8.1 Expressions for Padé approximants to Eq. (A.17) in the case of the S3 set of parameters and expansion about $Z = 1.2$

$$PA[1,1] = \frac{1.68195325 + 4.07888304Z}{1 + 0.00123177369Z} \tag{A.32}$$

$$PA[2,1] = \frac{1.72676374 - 4.94765568Z - 21.0525664Z^2}{1 - 5.26315789Z} \tag{A.33}$$

$$PA[2,2] = \frac{1.72748245 - 5.06898903Z - 21.3365122Z^2}{1 - 5.33422746Z + 0.0000164636478Z^2} \tag{A.34}$$

$$PA[3,2] = \frac{1.72334449 - 12.655213Z + 4.03651879Z^2 + 101.007557Z^3}{1 - 9.72292191Z + 25.2518892Z^2} \tag{A.35}$$

As noted in the main text, the $[2,2]$ RF is essentially a $[2,1]$ RF, and the large-Z slopes of these RFs are all 4.0, the correct a_1 value for the S3 case.

A.8.2 Results for S1 and S2 test cases

Below we present stationary points (Z_{sp}) and resonance energies (E_{sp}) obtained from RF fits to data generated using Eq. (A.14) with the V, a_1, and a_2 parameters listed in Table 7 for the test cases labeled S1 and S2. In each Table, we specify the precision ε to which the energies were evaluated, and we list the values of Z used. We also give the exact values of Z_{sp} and E_{sp} (from Eqs. (A.15) and (A.16), respectively) as well as the half-widths correct through orders V^2 (Eq. (A.22)) and V^4 (Eq. (A.24)). Recall from earlier discussion in the main text that RF_1-level results are expected to be similar to, but not match exactly, the V^2-level results while RF_2-level results should be close to V^4-level results.

For the S1 test case, the exact stationary-point and resonance energy values are: $Z_{sp} = 0.200 - 0.0386501i$ and $E_{sp} = 2.50 - 0.0702728i$. The V^2 half-width is 0.0667 eV, the V^4 half-width is 0.0697 eV, and the exact half-width is 0.0703 eV. The value of Z at which the V^4 contribution falls below 10^{-5} eV is 0.44 and the corresponding V^2 cut-off is 111. The three largest Z-values used for the RF_1 and RF_2 fits lie outside the $\varepsilon = 10^{-5}$ eV V^4 cut-off. Beginning with a Z-value of 1.6, where the stabilization plot was judged to be nearly linear

and placing three more data points at smaller Z-values, we generated the [2,1] fits for which the extracted stationary points and associated energies are summarized in Table 12.

We see that the [2,1] RF cannot achieve the V^4 half-with of 0.0697 eV even if data of very high accuracy are employed (e.g., for $\varepsilon = 10^{-7}$ eV the V^4 bound is $Z = 2.42$ below which all four Z-values in Table 12 lie). As discussed in the text, the functional form of the [2,1] RF contains only four parameters and is only capable of achieving V^2-level results (half-width of 0.0667 eV) as it does using four Z-values within the V^2 cut-off. The fact that the [2,1] RF gives a half width near the V^2 value suggests that the four data points used were below the V^2 bound and that they included enough information from the near-linear region to describe the a_1 parameter.

Next, we chose six Z-values including at least one from the near-linear region and the remaining below that and formed the [3,2] RF whose results appear in Table 13. This [3,2] RF contains six parameters and should be capable of achieving close to V^4-level results if the energy data employed is of sufficient precision. The V^4-level bound is $Z = 0.44$ at $\varepsilon = 10^{-5}$ eV, but three of the six Z-values used do not fall below this value and, as a result, the 0.0784 eV half-width obtained differs considerably from the V^4 value of 0.0697 eV. However, when using energy data precise to $\varepsilon = 10^{-7}$ eV, where the V^4 cut-off is $Z = 24$, the [3,2] RF achieves a half-width (0.0705 eV) very close to the V^4 value and to the exact value 0.0703 eV.

Seeing that the $\varepsilon = 10^{-5}$ eV half-width changed from 0.066 to 0.078 eV when moving from RF_1 to RF_2 suggests that some of the six Z-values used in the RF_2 lie below or near the V^4 bound, but the discussion of case S3 in the main text informs us that one likely needs all six Z-values to be below this bound making sure to have at least one data point in the near-linear region. For this reason, we next consider selecting six even smaller Z-values (but keeping one within the near-linear region) whose results are shown in Table 14. The V^4 cut-off is $Z = 0.44$ at $\varepsilon = 10^{-5}$ eV, and five of the Z-values fall below this bound, and one is just above it. As a result, this [3,2] RF achieves a half-width between the V^4 and exact values. For the test case labeled S2, the exact stationary-point and resonance energy values are: $Z_{sp} = 0.200 - 0.386501i$ and $E_{sp} = 2.50 - 0.702728i$. The V^2 half-width is 0.6667 eV, the V^4 half-width is 0.6973 eV, and the exact half-width is 0.7027 eV. The value of Z at which the V^4 contribution falls below 10^{-5} eV is 5.4, and the corresponding V^2 cut-off is $Z = 11.000$

(again, this V^2 bound is likely irrelevant in ab initio cases).

Because S1 and S2 have the same a_1 and a_2 slope parameters and differ only in their V values, we initially used the same four Z-values as used for Table 12. This [2,1] RF would be expected to closely approximate the V^2 half-width of 0.6667 eV since it has all four Z data points well inside the V^2 bound. However, it is not expected to come close to the V^4 half-width of 0.6973 eV even though all of its Z data points are inside the V^4 bound, because of its [2,1] functional form. Regardless of the precision value ε, a half-width of 0.736 eV is obtained, which is far from either the V^2 or V^4 value. Moreover, the real part of Z_{sp} (ca. 0.11) is not very close to the correct value of 0.20, so something appears to be wrong in this case.

If we cluster the four Z-values closer to the crossing point ($Z = 0.20$) and farther below the V^4 Z-value cut-off, we obtain even worse half-widths and $\mathrm{Re}(Z_{sp})$ regardless of the value of ε. This suggests that having all four data points far below the V^4 Z cut-off of 5.4 can be counterproductive. After further examining a plot of the S2 dataset, we realized that neither of the sets of Z-values considered above included any points in the near-linear region where the $a_1(Z - Z_0)$ term is dominant, a result of which is that the a_1 slope parameter could not be accurately determined. We therefore selected four new Z-values including some in the region where the relevant curve in the stabilization plot is essentially linear. This generated the [2,1] fit described in Table 16.

The RF fit at $\varepsilon = 10^{-5}$ eV has a large-Z slope of 0.999, which gives us the correct value of the a_1 parameter. Also, the half-width of 0.685 eV is close to the V^2 value of 0.697 eV, reinforcing the conclusion that one must not concentrate all four Z-values so far below the V^4 bound that the large-Z slope cannot be determined.

We now consider [3,2] RFs using six Z-values (Table 17). As noted above, for $\varepsilon = 10^{-5}$ eV, the V^4 cut-off is $Z = 5.4$, thus it is no surprise that the [3,2] RF obtains the V^4 half-width because all six of its data points are below 5.4, and it appears that at least one of the data points is within the near-linear region, so the a_1 parameter could be obtained.

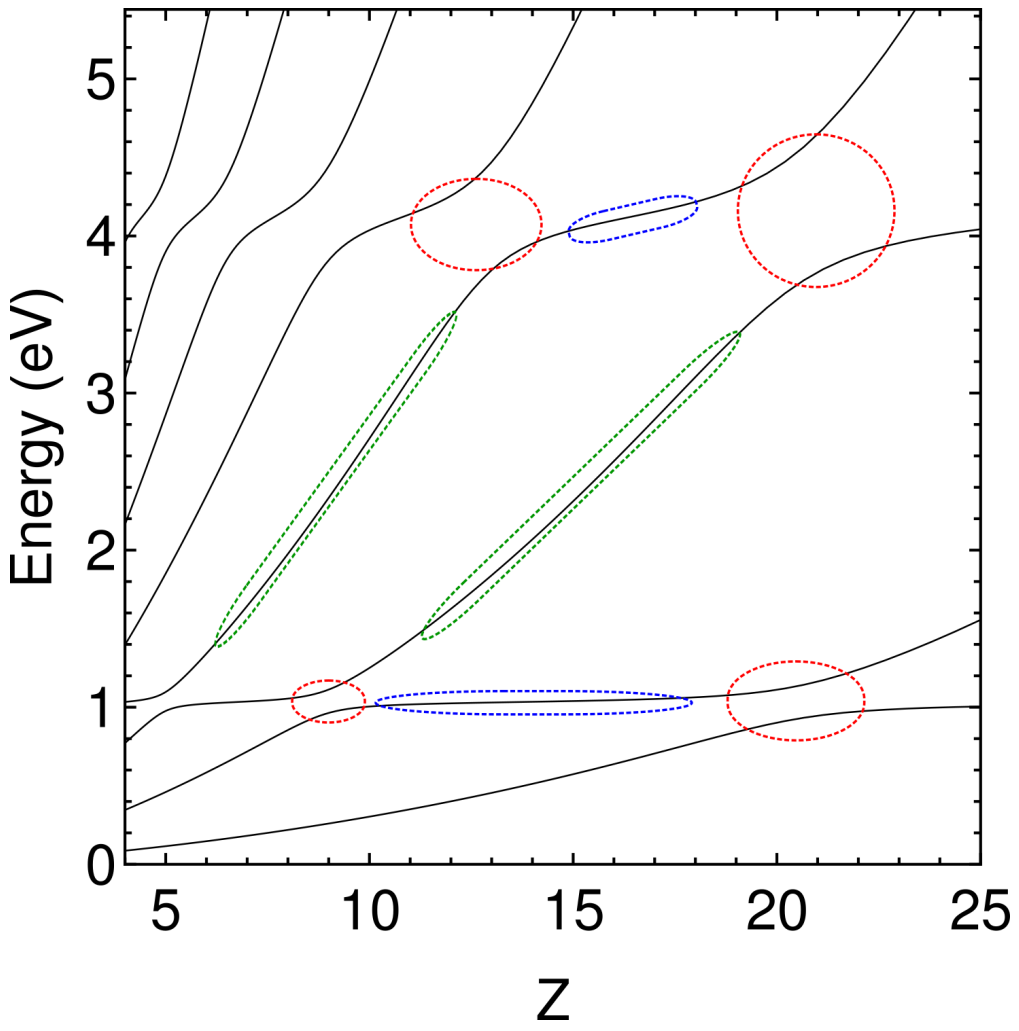

Figure 24: Example of a stabilization plot from a 1D model potential. The parameter Z scales the radial extent of the basis set. The colored dashed regions show examples of the features described in Section A.2. The regions surrounded in blue illustrate two of the plateaus, those in red illustrate four of the avoided crossings, and those in green illustrate two of the pseudo-continuum regions (see text).

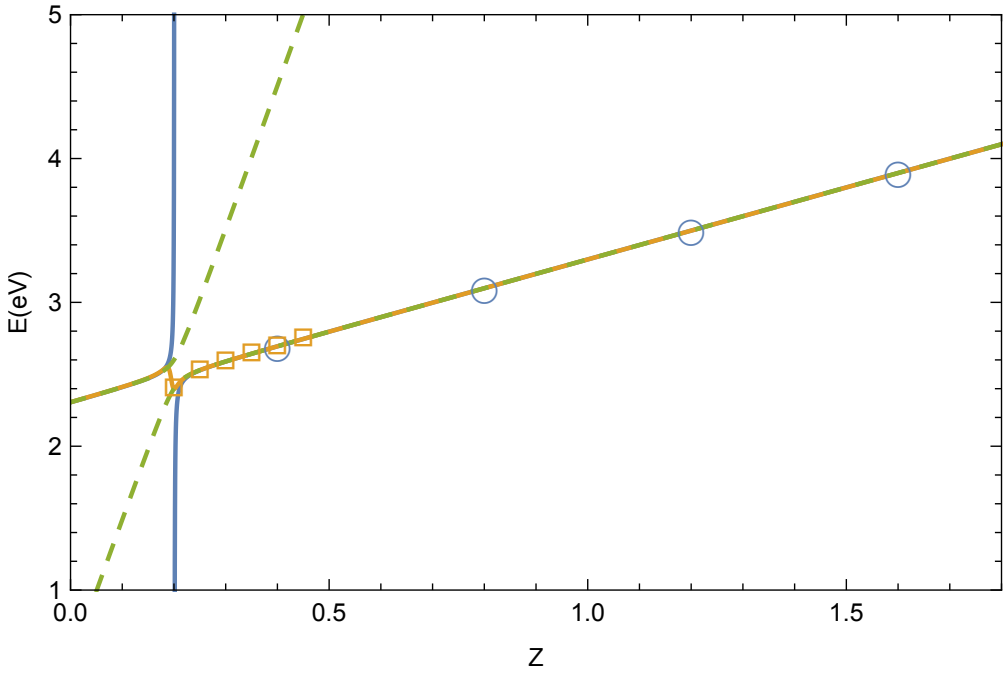

Figure 25: Resonance model (Eq. (A.14)) with the S1 parameter set. The dashed green curves represent the exact energies. The circles denote the data points listed in Table 12, and the blue curves the RF_1 fit to those points. The squares denote the data points listed in Table 14 and the orange curves are the RF_2 fits to these points.

Table 7: Description of the S1, S2, and S3 parameter sets and the upper limits for Z at the V^4 and V^2 levels for the model given by Eq. (A.14) with $\varepsilon = 10^{-5}$ eV.

Test Case	V(eV);a_1;a_2	$\frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right)^{1/3} + Z_0$	$\frac{V}{\delta a}\left(\frac{V}{\varepsilon}\right) + Z_0$
S1	0.1;1;10	0.439	111
S2	1.0;1;10	5.36	11100
S3	0.1;4;6	1.28	500

Table 8: Resonance half-widths (eV) for the model given by Eq. (A.14) using the S1-S3 parameter sets defined in Table 7.

Test Case	V^2 Half-width	V^4 Half-width	Exact Half-width
S1	0.0667	0.0697	0.0703
S2	0.6667	0.6973	0.7027
S3	0.2828	0.3810	0.4899

Table 9: Z_{sp} and E_{sp} (eV) from RF_1 fits to energy values at $Z = 0.4, 0.8, 1.2, 1.6$ for the model described by Eq. (A.14) with the S3 parameter set at different levels of precision ε (eV).

ε	Z_{sp}	E_{sp}
10^{-5}	0.175840862 - 0.0367804471 i	2.40399052 – 0.294207598 i
10^{-7}	0.177067499 - 0.0366038349 i	2.40877481 – 0.292804475 i
10^{-12}	0.17700637 - 0.036613091 i	2.40853738 – 0.292877948 i

Table 10: Z_{sp} and E_{sp} (eV) from RF_2 fits to energy values at $Z = 0.4, 0.6, 0.8, 1.0, 1.2, 1.4$ for the model described by Eq. (A.14) with the S3 parameter set at different levels of precision ε (eV).

ε	Z_{sp}	E_{sp}
10^{-5}	0.185646153 - 0.0358910308 i	2.44266485 - 0.28711605 i
10^{-7}	0.185462414 - 0.0779571285 i	2.45378066 - 0.430534863 i
10^{-12}	0.188780918 - 0.076937987 i	2.46415775 - 0.426152109 i

Table 11: Z_{sp} and E_{sp} (eV) from RF_2 fits to energy values at $Z = 0.25, 0.40, 0.55, 0.70, 0.85, 1.00$ for the model described by Eq. (A.14) with the S3 parameter set at different levels of precision ε (eV).

ε	Z_{sp}	E_{sp}
10^{-3}	0.123918560 - 0.0481745490 i	2.20867287 - 0.384156255 i
10^{-4}	0.147948561 - 0.0406098407 i	2.29477962 - 0.324741394 i
10^{-5}	0.128442525 - 0.0851233013 i	2.27808877 - 0.489209000 i
10^{-7}	0.177648487 - 0.0797605779 i	2.42936543 - 0.439177037 i
10^{-10}	0.177045193 - 0.0803243612 i	2.42773751 - 0.441438257 i

Table 12: RF_1 [2,1] results using $Z = 0.4, 0.8, 1.2, 1.6$.

ε (eV)	Z_{sp}	E_{sp} (eV)
10^{-5}	0.200790514 - 0.0331593265 i	2.50076741 - 0.0663195508 i
10^{-7}	0.199009626 - 0.0333702074 i	2.49901165 - 0.0667403774 i
10^{-12}	0.198967011 - 0.0333753786 i	2.49896962 - 0.066750703 i

Table 13: RF_2 [3,2] results using $Z = 0.2, 0.3, 0.4, 0.50, 0.60, 0.80$.

ε (eV)	Z_{sp}	E_{sp} (eV)
10^{-5}	0.187449988 - 0.0438278308 i	2.49338587 - 0.0784035553 i
10^{-7}	0.198130369 - 0.0384601071 i	2.49898306 - 0.0704764399 i
10^{-12}	0.198139183 - 0.0384389399 i	2.49898494 - 0.0704588588 i

Table 14: RF_2 [3,2] results using $Z = 0.20, 0.25, 0.30, 0.35, 0.40, 0.45$.

ε (eV)	Z_{sp}	E_{sp} (eV)
10^{-5}	0.196801783 - 0.0362839933 i	2.49717195 - 0.0695614783 i
10^{-7}	0.197970877 - 0.0393868362 i	2.49915003 - 0.0711634188 i
10^{-12}	0.197973989 - 0.0394145454 i	2.49916333 - 0.071180649 i

Table 15: S2 RF_1 [2,1] results using $Z = 0.4, 0.8, 1.2, 1.6$.

ε (eV)	Z_{sp}	E_{sp} (eV)
10^{-5}	0.107211538 - 0.3702366 i	2.42892618 - 0.735843273 i
10^{-7}	0.107300967 - 0.37014767 i	2.42893615 - 0.735687203 i
10^{-12}	0.107301691 - 0.370146837 i	2.42893599 - 0.735685833 i

Table 16: S2 RF_1 [2,1] results using $Z = 0.4, 2.07, 3.73, 5.4$.

ε (eV)	Z_{sp}	E_{sp} (eV)
10^{-5}	0.137723706 - 0.342471258 i	2.43952012 - 0.684829025 i
10^{-7}	0.13786842 - 0.342360368 i	2.43962073 - 0.684611524 i
10^{-12}	0.137867479 - 0.342361078 i	2.43962003 - 0.684612928 i

Table 17: S2 RF_2 [3,2] results using $Z = 0.4, 0.8, 1.2, 1.6, 2.0, 2.4$.

ε	Z_{sp}	E_{sp}
10^{-5}	0.184618417 - 0.364902059 i	2.48487808 - 0.693311191 i
10^{-7}	0.187777833 - 0.388309498 i	2.4947076 - 0.706424753 i
10^{-12}	0.187767906 - 0.388327002 i	2.49470624 - 0.706439517 i

Bibliography

[1] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622, DOI: 10.1103/PhysRev.46.618 (pages 1, 25).

[2] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483, DOI: 10.1016/S0009-2614(89)87395-6 (pages 1, 26, 37).

[3] Reynolds, P. J.; Ceperley, D. M.; Alder, B. J.; Lester, W. A. Fixed-node quantum Monte Carlo for molecules a) b). *J. Chem. Phys.* **1982**, *77*, 5593–5603, DOI: 10.1063/1.443766 (pages 1, 19, 26).

[4] Umrigar, C. J.; Nightingale, M. P.; Runge, K. J. A diffusion Monte Carlo algorithm with very small time-step errors. *J. Chem. Phys.* **1993**, *99*, 2865–2890, DOI: 10.1063/1.465195 (pages 1, 19, 26, 29).

[5] Foulkes, W. M. C.; Mitas, L.; Needs, R. J.; Rajagopal, G. Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.* **2001**, *73*, 33–83, DOI: 10.1103/RevModPhys.73.33 (pages 1, 19, 26, 36, 52).

[6] Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Molecular Physics* **2017**, *115*, 2315–2372, DOI: 10.1080/00268976.2017.1333644 (page 2).

[7] Ivanic, J.; Ruedenberg, K. Identification of deadwood in configuration spaces through general direct configuration interaction. *Theoretical Chemistry Accounts* **2001**, *106*, 339–351, DOI: 10.1007/s002140100285 (page 14).

[8] Bender, C. F.; Davidson, E. R. Studies in Configuration Interaction: The First-Row Diatomic Hydrides. *Phys. Rev.* **1969**, *183*, 23–30, DOI: 10.1103/physrev.183.23 (pages 14, 63, 67, 69).

[9] Whitten, J. L.; Hackmeyer, M. Configuration Interaction Studies of Ground and Excited States of Polyatomic Molecules. I. The CI Formulation and Studies of Formaldehyde. *J. Chem. Phys.* **1969**, *51*, 5584–5596, DOI: 10.1063/1.1671985 (pages 14, 63, 67, 69).

[10] Huron, B.; Malrieu, J. P.; Rancurel, P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *J. Chem. Phys.* **1973**, *58*, 5745–5759, DOI: 10.1063/1.1679199 (pages 14, 63, 67, 69, 70).

[11] Booth, G. H.; Thom, A. J. W.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **2009**, *131*, 054106, DOI: 10.1063/1.3193710 (pages 14, 63, 69).

[12] Booth, G. H.; Alavi, A. Approaching chemical accuracy using full configuration-interaction quantum Monte Carlo: A study of ionization potentials. *J. Chem. Phys.* **2010**, *132*, 174104, DOI: 10.1063/1.3407895 (pages 14, 63, 69).

[13] Cleland, D.; Booth, G. H.; Alavi, A. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **2010**, *132*, 041103, DOI: 10.1063/1.3302277 (pages 14, 63, 69, 76).

[14] Sharma, S.; Holmes, A. A.; Jeanmairet, G.; Alavi, A.; Umrigar, C. J. Semistochastic Heat-Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory. *J. Chem. Theory Comput.* **2017**, *13*, 1595–1604, DOI: 10.1021/acs.jctc.6b01028 (pages 14, 63, 69, 70, 86).

[15] Evangelista, F. A. Adaptive multiconfigurational wave functions. *J. Chem. Phys.* **2014**, *140*, 124114, DOI: 10.1063/1.4869192 (pages 14, 63, 69).

[16] Schriber, J. B.; Evangelista, F. A. Communication: An adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy. *J. Chem. Phys.* **2016**, *144*, 161106, DOI: 10.1063/1.4948308 (pages 14, 63, 69).

[17] Garniron, Y.; Giner, E.; Malrieu, J.-P.; Scemama, A. Alternative definition of excitation amplitudes in multi-reference state-specific coupled cluster. *J. Chem. Phys.* **2017**, *146*, 154107, DOI: 10.1063/1.4980034 (pages 14, 76, 92, 96).

[18] Garniron, Y.; Scemama, A.; Giner, E.; Caffarel, M.; Loos, P. F. Selected Configuration Interaction Dressed by Perturbation. *J. Chem. Phys.* **2018**, *149*, 064103, DOI: 10.1063/1.5044503 (pages 14, 63, 69, 70, 85, 97).

[19] Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. A Mountaineering Strategy to Excited States: Highly Accurate Reference Energies and Benchmarks. *J. Chem. Theory Comput.* **2018**, *14*, 4360–4379, DOI: 10.1021/acs.jctc.8b00406 (pages 14, 63, 64, 69, 70, 79, 83, 85, 97).

[20] Garniron, Y.; Applencourt, T.; Gasperich, K.; Benali, A.; Ferté, A.; Paquier, J.; Pradines, B.; Assaraf, R.; Reinhardt, P.; Toulouse, J.; Barbaresco, P.; Renon, N.; David, G.; Malrieu, J.-P.; Véril, M.; Caffarel, M.; Loos, P.-F.; Giner, E.; Scemama, A. Quantum Package 2.0: An Open-Source Determinant-Driven Suite of Programs. *J. Chem. Theory Comput.* **2019**, *15*, 3591–3609, DOI: 10.1021/acs.jctc.9b00176 (pages 14, 24, 61).

[21] Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087, DOI: 10.1063/1.1699114 (page 17).

[22] Hastings, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **1970**, *57*, 97–109, DOI: 10.1093/biomet/57.1.97 (page 17).

[23] Kalos, M. H.; Levesque, D.; Verlet, L. Helium at zero temperature with hard-sphere and other forces. *Phys. Rev. A* **1974**, *9*, 2178–2195, DOI: 10.1103/PhysRevA.9.2178 (page 18).

[24] Umrigar, C. J. Observations on variational and projector Monte Carlo methods. *J. Chem. Phys.* **2015**, *143*, 164105, DOI: 10.1063/1.4933112 (page 18).

[25] Anderson, J. B. A random-walk simulation of the Schrödinger equation:H_3^+. *J. Chem. Phys.* **1975**, *63*, 1499–1503, DOI: 10.1063/1.431514 (pages 19, 26, 36, 52).

[26] Anderson, J. B. 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.* **1976**, *65*, 4121–4127, DOI: 10.1063/1.432868 (pages 19, 52).

[27] Needs, R. J.; Towler, M. D.; Drummond, N. D.; López Ríos, P. Continuum variational and diffusion quantum Monte Carlo calculations. *J. Phys. Condens. Matter* **2010**, *22*, 023201, DOI: 10.1088/0953-8984/22/2/023201 (pages 19, 26, 28, 55).

[28] Austin, B. M.; Zubarev, D. Y.; Lester, W. A. Quantum Monte Carlo and Related Approaches. *Chem. Rev.* **2012**, *112*, 263–288, DOI: 10.1021/cr2001564 (pages 19, 26, 52).

[29] Mella, M.; Anderson, J. B. Intermolecular forces and fixed-node diffusion Monte Carlo: A brute force test of accuracies for He_2 and He-LiH. *J. Chem. Phys.* **2003**, *119*, 8225–8228, DOI: 10.1063/1.1612479 (pages 21, 26, 52).

[30] Dubecký, M.; Jurečka, P.; Derian, R.; Hobza, P.; Otyepka, M.; Mitas, L. Quantum Monte Carlo Methods Describe Noncovalent Interactions with Subchemical Accuracy. *J. Chem. Theory Comput.* **2013**, *9*, 4287–4292, DOI: 10.1021/ct4006739 (pages 21, 26, 52).

[31] Jastrow, R. Many-Body Problem with Strong Forces. *Phys. Rev.* **1955**, *98*, 1479–1484, DOI: 10.1103/PhysRev.98.1479 (pages 22, 27, 38).

[32] Drummond, N. D.; Towler, M. D.; Needs, R. J. Jastrow correlation factor for atoms, molecules, and solids. *Phys. Rev. B* **2004**, *70*, 235119, DOI: 10.1103/PhysRevB.70.235119 (pages 22, 27, 38, 55).

[33] López Ríos, P.; Seth, P.; Drummond, N. D.; Needs, R. J. Framework for constructing generic Jastrow correlation factors. *Phys. Rev. E* **2012**, *86*, 036703, DOI: 10.1103/PhysRevE.86.036703 (pages 22, 27).

[34] Gasperich, K.; D. Jordan, K., Diffusion Monte Carlo Study of the Parallel Displaced Form of the Benzene Dimer. In *Recent Prog. Quantum Monte Carlo*, Tanaka, S., Roy, P.-N., Mitas, L., Eds.; ACS Symposium Series, Vol. 1234; American Chemical Society: 2016, pp 107–117, DOI: 10.1021/bk-2016-1234.ch007 (page 25).

[35] Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. Estimates of the Ab Initio Limit for π-π Interactions: The Benzene Dimer. *J. Am. Chem. Soc.* **2002**, *124*, 10887–10893, DOI: 10.1021/ja025896h (page 25).

[36] Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. Origin of Attraction and Directionality of the π/π Interaction: Model Chemistry Calculations of Benzene Dimer Interaction. *J. Am. Chem. Soc.* **2002**, *124*, 104–112, DOI: 10.1021/ja0105212 (page 25).

[37] Sinnokrot, M. O.; Sherrill, C. D. Highly Accurate Coupled Cluster Potential Energy Curves for the Benzene Dimer: Sandwich, T-Shaped, and Parallel-Displaced Configurations. *J. Phys. Chem. A* **2004**, *108*, 10200–10207, DOI: 10.1021/jp0469517 (page 25).

[38] Zhao, Y.; Truhlar, D. G. Multicoefficient Extrapolated Density Functional Theory Studies of $\pi \cdots \pi$ Interactions: The Benzene Dimer. *J. Phys. Chem. A* **2005**, *109*, 4209–4212, DOI: 10.1021/jp050932v (page 25).

[39] Park, Y. C.; Lee, J. S. Accurate ab Initio Binding Energies of the Benzene Dimer. *J. Phys. Chem. A* **2006**, *110*, 5091–5095, DOI: 10.1021/jp0582888 (page 25).

[40] Podeszwa, R.; Bukowski, R.; Szalewicz, K. Potential Energy Surface for the Benzene Dimer and Perturbational Analysis of π-π Interactions. *J. Phys. Chem. A* **2006**, *110*, 10345–10354, DOI: 10.1021/jp064095o (page 25).

[41] Puzder, A.; Dion, M.; Langreth, D. C. Binding energies in benzene dimers: Nonlocal density functional calculations. *J. Chem. Phys.* **2006**, *124*, 164105, DOI: 10.1063/1.2189229 (page 25).

[42] Lee, E. C.; Kim, D.; Jurečka, P.; Tarakeshwar, P.; Hobza, P.; Kim, K. S. Understanding of Assembly Phenomena by Aromatic-Aromatic Interactions: Benzene Dimer and the Substituted Systems. *J. Phys. Chem. A* **2007**, *111*, 3446–3457, DOI: 10.1021/jp068635t (page 25).

[43] Slipchenko, L. V.; Gordon, M. S. Electrostatic energy in the effective fragment potential method: Theory and application to benzene dimer. *J. Comput. Chem.* **2007**, *28*, 276–291, DOI: 10.1002/jcc.20520 (page 25).

[44] Bludský, O.; Rubeš, M.; Soldán, P.; Nachtigall, P. Investigation of the benzene-dimer potential energy surface: DFT/CCSD(T) correction scheme. *J. Chem. Phys.* **2008**, *128*, 114102, DOI: 10.1063/1.2890968 (page 25).

[45] Pitoňák, M.; Neogrády, P.; Řezáč, J.; Jurečka, P.; Urban, M.; Hobza, P. Benzene Dimer: High-Level Wave Function and Density Functional Theory Calculations. *J. Chem. Theory Comput.* **2008**, *4*, 1829–1834, DOI: 10.1021/ct800229h (page 25).

[46] Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. An Assessment of Theoretical Methods for Nonbonded Interactions: Comparison to Complete Basis Set Limit Coupled-Cluster Potential Energy Curves for the Benzene Dimer, the Methane Dimer, Benzene-Methane, and Benzene-H_2S. *J. Phys. Chem. A* **2009**, *113*, 10146–10159, DOI: 10.1021/jp9034375 (page 25).

[47] Tkatchenko, A.; DiStasio, R. A.; Head-Gordon, M.; Scheffler, M. Dispersion-corrected Møller–Plesset second-order perturbation theory. *J. Chem. Phys.* **2009**, *131*, 094106, DOI: 10.1063/1.3213194 (page 25).

[48] Pitoňák, M.; Heßelmann, A. Accurate Intermolecular Interaction Energies from a Combination of MP2 and TDDFT Response Theory. *J. Chem. Theory Comput.* **2010**, *6*, 168–178, DOI: 10.1021/ct9005882 (pages 25, 26).

[49] Kim, H.; Choi, J.-M.; Goddard, W. A. Universal Correction of Density Functional Theory to Include London Dispersion (up to Lr, Element 103). *J. Phys. Chem. Lett.* **2012**, *3*, 360–363, DOI: 10.1021/jz2016395 (page 25).

[50] Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993, DOI: 10.1039/B600027D (pages 26, 29).

[51] Casimir, H. B. G.; Polder, D. The Influence of Retardation on the London-van der Waals Forces. *Phys. Rev.* **1948**, *73*, 360–372, DOI: 10.1103/PhysRev.73.360 (page 26).

[52] Szabo, A.; Ostlund, N. S. The correlation energy in the random phase approximation: Intermolecular forces between closed-shell systems. *J. Chem. Phys.* **1977**, *67*, 4351–4360, DOI: 10.1063/1.434580 (page 26).

[53] Cybulski, S. M.; Lytle, M. L. The origin of deficiency of the supermolecule second-order Møller-Plesset approach for evaluating interaction energies. *J. Chem. Phys.* **2007**, *127*, 141102, DOI: 10.1063/1.2795693 (page 26).

[54] Heßelmann, A. Improved supermolecular second order Møller–Plesset intermolecular interaction energies using time-dependent density functional response theory. *J. Chem. Phys.* **2008**, *128*, 144112, DOI: 10.1063/1.2905808 (page 26).

[55] Miliordos, E.; Aprà, E.; Xantheas, S. S. Benchmark Theoretical Study of the π–π Binding Energy in the Benzene Dimer. *J. Phys. Chem. A* **2014**, *118*, 7568–7578, DOI: 10.1021/jp5024235 (pages 26, 27, 29).

[56] Deible, M. J.; Kessler, M.; Gasperich, K. E.; Jordan, K. D. Quantum Monte Carlo calculation of the binding energy of the beryllium dimer. *J. Chem. Phys.* **2015**, *143*, 084116, DOI: 10.1063/1.4929351 (pages 26, 36, 51).

[57] Trail, J. R.; Needs, R. J. Norm-conserving Hartree–Fock pseudopotentials and their asymptotic behavior. *J. Chem. Phys.* **2005**, *122*, 014112, DOI: 10.1063/1.1829049 (page 27).

[58] Trail, J. R.; Needs, R. J. Smooth relativistic Hartree–Fock pseudopotentials for H to Ba and Lu to Hg. *J. Chem. Phys.* **2005**, *122*, 174109, DOI: 10.1063/1.1888569 (page 27).

[59] Xu, J.; Deible, M. J.; Peterson, K. A.; Jordan, K. D. Correlation Consistent Gaussian Basis Sets for H, B–Ne with Dirac–Fock AREP Pseudopotentials: Applications in Quantum Monte Carlo Calculations. *J. Chem. Theory Comput.* **2013**, *9*, 2170–2178, DOI: 10.1021/ct300983b (page 27).

[60] Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652, DOI: 10.1063/1.464913 (pages 27, 53).

[61] Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211, DOI: 10.1139/p80-159 (pages 27, 53).

[62] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789, DOI: 10.1103/PhysRevB.37.785 (pages 27, 53).

[63] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional

Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627, DOI: 10.1021/j100096a001 (pages 27, 53).

[64] Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023, DOI: 10.1063/1.456153 (pages 27, 38).

[65] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806, DOI: 10.1063/1.462569 (page 27).

[66] Casula, M.; Moroni, S.; Sorella, S.; Filippi, C. Size-consistent variational approaches to nonlocal pseudopotentials: Standard and lattice regularized diffusion Monte Carlo methods revisited. *J. Chem. Phys.* **2010**, *132*, 154113, DOI: 10.1063/1.3380831 (page 27).

[67] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision B.01., Wallingford, CT, 2009 (pages 28, 55).

[68] Azadi, S.; Cohen, R. E. Chemical accuracy from quantum Monte Carlo for the benzene dimer. *J. Chem. Phys.* **2015**, *143*, 104301, DOI: 10.1063/1.4930137 (page 29).

[69] López Ríos, P.; Ma, A.; Drummond, N. D.; Towler, M. D.; Needs, R. J. Inhomogeneous backflow transformations in quantum Monte Carlo calculations. *Phys. Rev. E* **2006**, *74*, 066701, DOI: 10.1103/PhysRevE.74.066701 (page 29).

[70] Dubecký, M.; Mitas, L. Personal Communication., 2015 (page 29).

[71] Dubecký, M.; Derian, R.; Jurečka, P.; Mitas, L.; Hobza, P.; Otyepka, M. Quantum Monte Carlo for noncovalent interactions: an efficient protocol attaining benchmark accuracy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 20915–20923, DOI: 10.1039/C4CP02093F (page 29).

[72] Sorella, S.; Casula, M.; Rocca, D. Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods. *J. Chem. Phys.* **2007**, *127*, 014105, DOI: 10.1063/1.2746035 (page 30).

[73] Casula, M.; Filippi, C.; Sorella, S. Diffusion Monte Carlo Method with Lattice Regularization. *Phys. Rev. Lett.* **2005**, *95*, 100201, DOI: 10.1103/PhysRevLett.95.100201 (page 30).

[74] Casula, M.; Attaccalite, C.; Sorella, S. Correlated geminal wave function for molecules: An efficient resonating valence bond approach. *J. Chem. Phys.* **2004**, *121*, 7110–7126, DOI: 10.1063/1.1794632 (pages 30, 52, 53).

[75] Casula, M.; Sorella, S. Geminal wave functions with Jastrow correlation: A first application to atoms. *J. Chem. Phys.* **2003**, *119*, 6500–6511, DOI: 10.1063/1.1604379 (page 30).

[76] Gasperich, K.; Deible, M.; Jordan, K. D. H_4 : A model system for assessing the performance of diffusion Monte Carlo calculations using a single Slater determinant trial function. *J. Chem. Phys.* **2017**, *147*, 074106, DOI: 10.1063/1.4986216 (page 35).

[77] Pozun, Z. D.; Su, X.; Jordan, K. D. Establishing the Ground State of the Disjoint Diradical Tetramethyleneethane with Quantum Monte Carlo. *J. Am. Chem. Soc.* **2013**, *135*, 13862–13869, DOI: 10.1021/ja406002n (page 36).

[78] Anderson, J. B. Quantum chemistry by random walk: H_4 square. *Int. J. Quantum Chem.* **1979**, *15*, 109–120, DOI: 10.1002/qua.560150111 (pages 36, 37, 39).

[79] Filippi, C.; Umrigar, C. J. Multiconfiguration wave functions for quantum Monte Carlo calculations of first-row diatomic molecules. *J. Chem. Phys.* **1996**, *105*, 213–226, DOI: 10.1063/1.471865 (page 36).

[80] Zen, A.; Coccia, E.; Luo, Y.; Sorella, S.; Guidoni, L. Static and Dynamical Correlation in Diradical Molecules by Quantum Monte Carlo Using the Jastrow Antisymmetrized Geminal Power Ansatz. *J. Chem. Theory Comput.* **2014**, *10*, 1048–1061, DOI: 10.1021/ct401008s (page 36).

[81] Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **1980**, *48*, 157–173, DOI: 10.1016/0301-0104(80)80045-0 (page 36).

[82] Ceperley, D. M. Fermion nodes. *J. Stat. Phys.* **1991**, *63*, 1237–1267, DOI: 10.1007/BF01030009 (pages 36, 42, 52).

[83] Loos, P.-F.; Bressanini, D. Nodal surfaces and interdimensional degeneracies. *J. Chem. Phys.* **2015**, *142*, 214112, DOI: 10.1063/1.4922159 (pages 36, 42).

[84] Mitas, L. Structure of Fermion Nodes and Nodal Cells. *Phys. Rev. Lett.* **2006**, *96*, 240402, DOI: 10.1103/PhysRevLett.96.240402 (pages 36, 42).

[85] Bajdich, M.; Mitas, L.; Drobný, G.; Wagner, L. K. Approximate and exact nodes of fermionic wavefunctions: Coordinate transformations and topologies. *Phys. Rev. B* **2005**, *72*, 075131, DOI: 10.1103/PhysRevB.72.075131 (pages 36, 42).

[86] Bressanini, D. Implications of the two nodal domains conjecture for ground state fermionic wave functions. *Phys. Rev. B* **2012**, *86*, 115120, DOI: 10.1103/PhysRevB.86.115120 (page 36).

[87] Bressanini, D.; Morosi, G. On the nodal structure of single-particle approximation based atomic wave functions. *J. Chem. Phys.* **2008**, *129*, 054103, DOI: 10.1063/1.2963501 (page 36).

[88] Bressanini, D.; Reynolds, P. J. Unexpected Symmetry in the Nodal Structure of the He Atom. *Phys. Rev. Lett.* **2005**, *95*, 110201, DOI: 10.1103/PhysRevLett.95.110201 (page 36).

[89] Scott, T. C.; Lüchow, A.; Bressanini, D.; Morgan, J. D. Nodal surfaces of helium atom eigenfunctions. *Phys. Rev. A* **2007**, *75*, 060101, DOI: 10.1103/PhysRevA.75.060101 (page 36).

[90] Bressanini, D.; Morosi, G.; Tarasco, S. An investigation of nodal structures and the construction of trial wave functions. *J. Chem. Phys.* **2005**, *123*, 204109, DOI: 10.1063/1.2128672 (pages 36, 52, 53).

[91] Bressanini, D.; Ceperley, D. M.; Reynolds, P. J., WHAT DO WE KNOW ABOUT WAVE FUNCTION NODES? In, 2002, pp 3–11, DOI: 10.1142/9789812775696_0001 (pages 36, 52).

[92] Silver, D. M.; Stevens, R. M. Reaction paths on the H_4 potential energy surface. *J. Chem. Phys.* **1973**, *59*, 3378–3394, DOI: 10.1063/1.1680481 (page 38).

[93] Guest, M. F.; Saunders, V. R. On methods for converging open-shell Hartree-Fock wave-functions. *Mol. Phys.* **1974**, *28*, 819–828, DOI: 10.1080/00268977400102171 (page 38).

[94] Esler, K.; Kim, J.; Ceperley, D.; Shulenburger, L. Accelerating Quantum Monte Carlo Simulations of Real Materials on GPU Clusters. *Comput. Sci. Eng.* **2012**, *14*, 40–51, DOI: 10.1109/MCSE.2010.122 (page 38).

[95] Umrigar, C. J.; Filippi, C. Energy and Variance Optimization of Many-Body Wave Functions. *Phys. Rev. Lett.* **2005**, *94*, 150201, DOI: 10.1103/PhysRevLett.94.150201 (page 38).

[96] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363, DOI: 10.1002/jcc.540141112 (pages 38, 55, 92).

[97] Kim, J.; Esler, K. P.; McMinis, J.; Morales, M. A.; Clark, B. K.; Shulenburger, L.; Ceperley, D. M. Hybrid algorithms in quantum Monte Carlo. *J. Phys. Conf. Ser.* **2012**, *402*, 012008, DOI: 10.1088/1742-6596/402/1/012008 (pages 38, 55).

[98] Bondybey, V. E.; English, J. H. Laser vaporization of beryllium: Gas phase spectrum and molecular potential of Be_2. *J. Chem. Phys.* **1984**, *80*, 568–570, DOI: 10.1063/1.446434 (page 52).

[99] Bondybey, V. E. Electronic structure and bonding of Be_2. *Chem. Phys. Lett.* **1984**, *109*, 436–441, DOI: 10.1016/0009-2614(84)80339-5 (page 52).

[100] Bondbey, V. E. Laser-Induced Fluorescence and Bonding of Metal Dimers. *Science* **1985**, *227*, 125–131, DOI: 10.1126/science.227.4683.125 (page 52).

[101] Merritt, J. M.; Bondybey, V. E.; Heaven, M. C. Beryllium Dimer–Caught in the Act of Bonding. *Science* **2009**, *324*, 1548–1551, DOI: 10.1126/science.1174326 (pages 52, 53, 57).

[102] Meshkov, V. V.; Stolyarov, A. V.; Heaven, M. C.; Haugen, C.; LeRoy, R. J. Direct-potential-fit analyses yield improved empirical potentials for the ground $X^1\Sigma_g^+$ state of Be_2. *J. Chem. Phys.* **2014**, *140*, 064315, DOI: 10.1063/1.4864355 (page 52).

[103] Bender, C. F.; Davidson, E. R. Theoretical Calculation of the Potential Curves of the Be_2 Molecule. *J. Chem. Phys.* **1967**, *47*, 4972–4978, DOI: 10.1063/1.1701748 (page 52).

[104] Robb, M. A.; Wilson, S. Comparison within the algebraic approximation of many-body perturbation theory and configuration interaction for the beryllium dimer. *Mol. Phys.* **1980**, *40*, 1333–1340, DOI: 10.1080/00268978000102321 (page 52).

[105] Blomberg, M. R. A.; Siegbahn, P. E. M.; Roos, B. O. The ground-state potential curve of the beryllium dimer. *Int. J. Quantum Chem.* **1980**, *18*, 229–247, DOI: 10.1002/qua.560180826 (page 52).

[106] Liu, B.; McLean, A. D. Ab initio potential curve for Be_2 ($^1\Sigma_g{}^+$) from the interacting correlated fragments method. *J. Chem. Phys.* **1980**, *72*, 3418–3419, DOI: 10.1063/1.439528 (page 52).

[107] Røeggen, I.; Almlöf, J. Interatomic potential for the $X^1\Sigma_g^+$ state of Be_2. *Int. J. Quantum Chem.* **1996**, *60*, 453–466, DOI: 10.1002/(SICI)1097-461X(1996)60:1<453::AID-QUA44>3.0.CO;2-A (page 52).

[108] Martin, J. M. L. The ground-state spectroscopic constants of Be_2 revisited. *Chem. Phys. Lett.* **1999**, *303*, 399–407, DOI: 10.1016/S0009-2614(99)00214-6 (page 52).

[109] Stärck, J.; Meyer, W. The ground state potential of the beryllium dimer. *Chem. Phys. Lett.* **1996**, *258*, 421–426, DOI: 10.1016/0009-2614(96)00657-4 (page 52).

[110] Füsti-Molnár, L.; Szalay, P. G. New Versions of Approximately Extensive Corrected Multireference Configuration Interaction Methods. *J. Phys. Chem.* **1996**, *100*, 6288–6297, DOI: 10.1021/jp952840j (page 52).

[111] Kaledin, L. A.; Kaledin, A. L.; Heaven, M. C.; Bondybey, V. E. Electronic structure of Be_2: theoretical and experimental results. *J. Mol. Struct. THEOCHEM* **1999**, *461-462*, 177–186, DOI: 10.1016/S0166-1280(98)00425-4 (page 52).

[112] Gdanitz, R. J. Accurately solving the electronic Schrödinger equation of atoms and molecules using explicitly correlated (r_{12}-)MR-CI. The ground state of beryllium dimer (Be_2). *Chem. Phys. Lett.* **1999**, *312*, 578–584, DOI: 10.1016/S0009-2614(99)00985-9 (page 52).

[113] Harrison, R. J.; Handy, N. C. Full CI results for Be_2 and $(H_2)_2$ in large basis sets. *Chem. Phys. Lett.* **1983**, *98*, 97–101, DOI: 10.1016/0009-2614(83)87105-X (page 52).

[114] Bauschlicher, C. W.; Langhoff, S. R.; Partridge, H. Theoretical study of the BeLi, BeNa, MgLi, MgNa, and AlBe molecules and their negative ions. *J. Chem. Phys.* **1992**, *96*, 1240–1247, DOI: 10.1063/1.462160 (page 52).

[115] Røeggen, I.; Veseth, L. Interatomic potential for the $X^1\Sigma_g^+$ state of Be_2 , revisited. *Int. J. Quantum Chem.* **2005**, *101*, 201–210, DOI: 10.1002/qua.20222 (page 52).

[116] Helal, W.; Evangelisti, S.; Leininger, T.; Monari, A. A FCI benchmark on beryllium dimer: The lowest singlet and triplet states. *Chem. Phys. Lett.* **2013**, *568-569*, 49–54, DOI: 10.1016/j.cplett.2013.03.012 (page 52).

[117] Sheng, X. W.; Kuang, X. Y.; Li, P.; Tang, K. T. Analyzing and modeling the interaction potential of the ground-state beryllium dimer. *Phys. Rev. A* **2013**, *88*, 022517, DOI: 10.1103/PhysRevA.88.022517 (page 52).

[118] El Khatib, M.; Bendazzoli, G. L.; Evangelisti, S.; Helal, W.; Leininger, T.; Tenti, L.; Angeli, C. Beryllium Dimer: A Bond Based on Non-Dynamical Correlation. *J. Phys. Chem. A* **2014**, *118*, 6664–6673, DOI: 10.1021/jp503145u (pages 52, 53, 57).

[119] Patkowski, K.; Spirko, V.; Szalewicz, K. On the Elusive Twelfth Vibrational State of Beryllium Dimer. *Science* **2009**, *326*, 1382–1384, DOI: 10.1126/science.1181017 (page 52).

[120] Harkless, J. A. W.; Irikura, K. K. Multi-determinant trial functions in the determination of the dissociation energy of the beryllium dimer: Quantum Monte Carlo study. *Int. J. Quantum Chem.* **2006**, *106*, 2373–2378, DOI: 10.1002/qua.21014 (pages 52, 53, 57).

[121] Patkowski, K.; Podeszwa, R.; Szalewicz, K. Interactions in Diatomic Dimers Involving Closed-Shell Metals. *J. Phys. Chem. A* **2007**, *111*, 12822–12838, DOI: 10.1021/jp076412c (page 52).

[122] Sharma, S.; Yanai, T.; Booth, G. H.; Umrigar, C. J.; Chan, G. K.-L. Spectroscopic accuracy directly from quantum chemistry: Application to ground and excited states of beryllium dimer. *J. Chem. Phys.* **2014**, *140*, 104112, DOI: 10.1063/1.4867383 (page 52).

[123] Lesiuk, M.; Przybytek, M.; Musiał, M.; Jeziorski, B.; Moszynski, R. Reexamination of the calculation of two-center, two-electron integrals over Slater-type orbitals. III. Case study of the beryllium dimer. *Phys. Rev. A* **2015**, *91*, 012510, DOI: 10.1103/PhysRevA.91.012510 (page 52).

[124] Koput, J. The ground-state potential energy function of a beryllium dimer determined using the single-reference coupled-cluster approach. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20311, DOI: 10.1039/c1cp22417d (page 52).

[125] Schmidt, M. W.; Ivanic, J.; Ruedenberg, K. Electronic Structure Analysis of the Ground-State Potential Energy Curve of Be_2. *J. Phys. Chem. A* **2010**, *114*, 8687–8696, DOI: 10.1021/jp101506t (pages 52, 59).

[126] Toulouse, J.; Umrigar, C. J. Full optimization of Jastrow-Slater wave functions with application to the first-row atoms and homonuclear diatomic molecules. *J. Chem. Phys.* **2008**, *128*, 174101, DOI: 10.1063/1.2908237 (pages 52, 56).

[127] Hammond, B. L.; Lester, W. A.; Reynolds, P. J., *Monte Carlo Methods in Ab Initio Quantum Chemistry*; World Scientific Lecture and Course Notes in Chemistry, Vol. 1; WORLD SCIENTIFIC: 1994, DOI: 10.1142/1170 (pages 52, 53).

[128] Anderson, A. G.; Goddard, W. A. Generalized valence bond wave functions in quantum Monte Carlo. *J. Chem. Phys.* **2010**, *132*, 164110, DOI: 10.1063/1.3377091 (pages 52, 53).

[129] Cleland, D.; Booth, G. H.; Overy, C.; Alavi, A. Taming the First-Row Diatomics: A Full Configuration Interaction Quantum Monte Carlo Study. *J. Chem. Theory Comput.* **2012**, *8*, 4138–4152, DOI: 10.1021/ct300504f (page 52).

[130] Prascher, B. P.; Woon, D. E.; Peterson, K. A.; Dunning, T. H.; Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. VII. Valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg. *Theor. Chem. Acc.* **2011**, *128*, 69–82, DOI: 10.1007/s00214-010-0764-0 (pages 52, 53).

[131] Koput, J.; Peterson, K. A. Ab initio prediction of the potential energy surface and vibration-rotation energy levels of BeH_2. *J. Chem. Phys.* **2006**, *125*, 044306, DOI: 10.1063/1.2212932 (pages 52, 53).

[132] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868, DOI: 10.1103/PhysRevLett.77.3865 (page 53).

[133] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100, DOI: 10.1103/PhysRevA.38.3098 (page 53).

[134] Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170, DOI: 10.1063/1.478522 (page 53).

[135] Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377, DOI: 10.1063/1.464304 (page 53).

[136] Dykstra, C. E. An examination of the Brueckner condition for the selection of molecular orbitals in correlated wavefunctions. *Chem. Phys. Lett.* **1977**, *45*, 466–469, DOI: 10.1016/0009-2614(77)80065-1 (page 53).

[137] Handy, N. C.; Pople, J. A.; Head-Gordon, M.; Raghavachari, K.; Trucks, G. W. Size-consistent Brueckner theory limited to double substitutions. *Chem. Phys. Lett.* **1989**, *164*, 185–192, DOI: 10.1016/0009-2614(89)85013-4 (page 53).

[138] Ma, A.; Towler, M. D.; Drummond, N. D.; Needs, R. J. Scheme for adding electron-nucleus cusps to Gaussian orbitals. *J. Chem. Phys.* **2005**, *122*, 224322, DOI: 10.1063/1.1940588 (page 55).

[139] Flyvbjerg, H.; Petersen, H. G. Error estimates on averages of correlated data. *J. Chem. Phys.* **1989**, *91*, 461–466, DOI: 10.1063/1.457480 (page 55).

[140] Morales, M. A.; McMinis, J.; Clark, B. K.; Kim, J.; Scuseria, G. E. Multideterminant Wave Functions in Quantum Monte Carlo. *J. Chem. Theory Comput.* **2012**, *8*, 2181–2188, DOI: 10.1021/ct3003404 (page 55).

[141] Pachucki, K.; Komasa, J. Relativistic and QED Corrections for the Beryllium Atom. *Phys. Rev. Lett.* **2004**, *92*, 213001, DOI: 10.1103/PhysRevLett.92.213001 (pages 57, 58).

[142] Rasch, K. M.; Mitas, L. Impact of electron density on the fixed-node errors in Quantum Monte Carlo of atomic systems. *Chem. Phys. Lett.* **2012**, *528*, 59–62, DOI: 10.1016/j.cplett.2012.01.016 (page 57).

[143] Moore, G. Cramming More Components onto Integrated Circuits. *Electronics* **1965**, *38*, 114–117 (page 62).

[144] Lists | TOP500 Supercomputer Sites., [Online; accessed 9. Oct. 2018], 2018 (page 62).

[145] Sutter, H.; Larus, J. Software and the concurrency revolution. *Queue* **2005**, *3*, 54, DOI: 10.1145/1095408.1095421 (page 63).

[146] Wulf, Wm. A.; McKee, S. A. Hitting the memory wall: implications of the obvious. *SIGARCH Comput. Archit. News* **1995**, *23*, 20–24, DOI: 10.1145/216585.216588 (page 63).

[147] Khan, H. N.; Hounshell, D. A.; Fuchs, E. R. H. Science and research policy at the end of Moore's law. *Nature Electronics* **2018**, *1*, 14–21, DOI: 10.1038/s41928-017-0005-9 (page 63).

[148] Smith, J. E. T.; Mussard, B.; Holmes, A. A.; Sharma, S. Cheap and Near Exact CASSCF with Large Active Spaces. *J. Chem. Theory Comput.* **2017**, *13*, 5468–5478, DOI: 10.1021/acs.jctc.7b00900 (page 63).

[149] Neuhauser, D.; Rabani, E.; Baer, R. Expeditious Stochastic Approach for MP2 Energies in Large Electronic Systems. *J. Chem. Theory Comput.* **2013**, *9*, 24 (page 63).

[150] Willow, S. Y.; Hirata, S. Stochastic, real-space, imaginary-time evaluation of third-order Feynman–Goldstone diagrams. *J. Chem. Phys.* **2014**, *140*, 024111 (page 63).

[151] Willow, S. Y.; Kim, K. S.; Hirata, S. Stochastic evaluation of second-order many-body perturbation energies. *J. Chem. Phys.* **2012**, *137*, 204122 (page 63).

[152] Johnson, C. M.; Hirata, S.; Ten-no, S. Explicit correlation factors. *Chem. Phys. Lett.* **2017**, *683*, 247 (page 63).

[153] Johnson, C. M.; Doran, A. E.; Zhang, J.; Valeev, E. F.; Hirata, S. Monte Carlo explicitly correlated second-order many-body perturbation theory. *J. Chem. Phys.* **2016**, *145*, 154115 (page 63).

[154] Gruneis, A.; Hirata, S.; Ohnishi, Y.-Y.; Ten-no, S. Perspective: Explicitly correlated electronic structure theory for complex systems. *J. Chem. Phys.* **2017**, *146*, 080901 (page 63).

[155] Doran, A. E.; Hirata, S. Monte Carlo MP2 on Many Graphical Processing Units. *J. Chem. Theory Comput.* **2016**, *12*, 4821 (page 63).

[156] Garniron, Y.; Scemama, A.; Loos, P.-F.; Caffarel, M. Hybrid stochastic-deterministic calculation of the second-order perturbative contribution of multireference perturbation theory. *J. Chem. Phys.* **2017**, *147*, 034101, DOI: 10.1063/1.4992127 (pages 63, 64, 69, 86).

[157] Evangelisti, S.; Daudey, J.-P.; Malrieu, J.-P. Convergence of an improved CIPSI algorithm. *Chemical Physics* **1983**, *75*, 91–102, DOI: 10.1016/0301-0104(83)85011-3 (pages 63, 69, 76).

[158] Cimiraglia, R. Second order perturbation correction to CI energies by use of diagrammatic techniques: An improvement to the CIPSI algorithm. *J. Chem. Phys.* **1985**, *83*, 1746–1749, DOI: 10.1063/1.449362 (pages 63, 69).

[159] Cimiraglia, R.; Persico, M. Recent advances in multireference second order perturbation CI: The CIPSI method revisited. *J. Comput. Chem.* **1987**, *8*, 39–47 (pages 63, 69).

[160] Illas, F.; Rubio, J.; Ricart, J. M. Approximate natural orbitals and the convergence of a second order multireference many-body perturbation theory (CIPSI) algorithm. *J. Chem. Phys.* **1988**, *89*, 6376–6384, DOI: 10.1063/1.455405 (pages 63, 69).

[161] Povill, A.; Rubio, J.; Illas, F. Treating large intermediate spaces in the CIPSI method through a direct selected CI algorithm. *Theor. Chem. Acc.* **1992**, *82*, 229–238, DOI: 10.1007/BF01113255 (pages 63, 69).

[162] Abrams, M. L.; Sherrill, C. D. Important configurations in configuration interaction and coupled-cluster wave functions. *Chem. Phys. Lett.* **2005**, *412*, 121–124, DOI: 10.1016/j.cplett.2005.06.107 (pages 63, 69).

[163] Bunge, C. F.; Carbó-Dorca, R. Select-divide-and-conquer method for large-scale configuration interaction. *J. Chem. Phys.* **2006**, *125*, 014108, DOI: 10.1063/1.2207621 (pages 63, 69).

[164] Bytautas, L.; Ruedenberg, K. A priori identification of configurational deadwood. *Chemical Physics* **2009**, *356*, 64–75, DOI: 10.1016/j.chemphys.2008.11.021 (pages 63, 67, 69).

[165] Giner, E.; Scemama, A.; Caffarel, M. Using perturbatively selected configuration interaction in quantum Monte Carlo calculations. *Can. J. Chem.* **2013**, *91*, 879–885, DOI: 10.1139/cjc-2013-0017 (pages 63, 64, 69, 97).

[166] Caffarel, M.; Giner, E.; Scemama, A.; Ramírez-Solís, A. Spin Density Distribution in Open-Shell Transition Metal Systems: A Comparative Post-Hartree–Fock, Density Functional Theory, and Quantum Monte Carlo Study of the $CuCl_2$ Molecule. *J. Chem. Theory Comput.* **2014**, *10*, 5286–5296, DOI: 10.1021/ct5004252 (pages 63, 64, 69, 97).

[167] Giner, E.; Scemama, A.; Caffarel, M. Fixed-node diffusion Monte Carlo potential energy curve of the fluorine molecule F_2 using selected configuration interaction trial wavefunctions. *J. Chem. Phys.* **2015**, *142*, 044115, DOI: 10.1063/1.4905528 (pages 63, 64, 69, 97).

[168] Caffarel, M.; Applencourt, T.; Giner, E.; Scemama, A. Using CIPSI nodes in diffusion Monte Carlo. **2016**, DOI: 10.1021/bk-2016-1234.ch002 (pages 63, 64, 69, 97).

[169] Caffarel, M.; Applencourt, T.; Giner, E.; Scemama, A. Communication: Toward an improved control of the fixed-node error in quantum Monte Carlo: The case of the water molecule. *J. Chem. Phys.* **2016**, *144*, 151103, DOI: 10.1063/1.4947093 (pages 63, 64, 69, 97).

[170] Holmes, A. A.; Tubman, N. M.; Umrigar, C. J. Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling. *J. Chem. Theory Comput.* **2016**, *12*, 3674–3680, DOI: 10.1021/acs.jctc.6b00407 (pages 63, 69, 70, 86).

[171] Holmes, A. A.; Umrigar, C. J.; Sharma, S. Excited states using semistochastic heat-bath configuration interaction. *J. Chem. Phys.* **2017**, *147*, 164111, DOI: 10.1063/1.4998614 (pages 63, 69, 70).

[172] Chien, A. D.; Holmes, A. A.; Otten, M.; Umrigar, C. J.; Sharma, S.; Zimmerman, P. M. Excited States of Methylene, Polyenes, and Ozone from Heat-Bath Configuration Interaction. *J. Phys. Chem. A* **2018**, *122*, 2714–2722, DOI: 10.1021/acs.jpca.8b01554 (pages 63, 69, 70).

[173] Scemama, A.; Garniron, Y.; Caffarel, M.; Loos, P. F. Deterministic construction of nodal surfaces within quantum Monte Carlo: the case of FeS. *J. Chem. Theory Comput.* **2018**, *14*, 1395, DOI: 10.1021/acs.jctc.7b01250 (pages 63, 64, 69, 70, 79).

[174] Scemama, A.; Benali, A.; Jacquemin, D.; Caffarel, M.; Loos, P.-F. Excitation energies from diffusion Monte Carlo using selected configuration interaction nodes. *J. Chem. Phys.* **2018**, *149*, 034108, DOI: 10.1063/1.5041327 (pages 63, 64, 69, 70, 79, 97).

[175] Schriber, J. B.; Evangelista, F. A. Adaptive Configuration Interaction for Computing Challenging Electronic Excited States with Tunable Accuracy. *J. Chem. Theory Comput.* **2017**, DOI: 10.1021/acs.jctc.7b00725 (pages 63, 69).

[176] Liu, W.; Hoffmann, M. R. iCI: Iterative CI toward full CI. *J. Chem. Theory Comput.* **2016**, *12*, 1169–1178, DOI: 10.1021/acs.jctc.5b01099 (pages 63, 69).

[177] Per, M. C.; Cleland, D. M. Energy-based truncation of multi-determinant wavefunctions in quantum Monte Carlo. *J. Chem. Phys.* **2017**, *146*, 164101, DOI: 10.1063/1.4981527 (pages 63, 69).

[178] Ohtsuka, Y.; Hasegawa, J.-y. Selected configuration interaction method using sampled first-order corrections to wave functions. *J. Chem. Phys.* **2017**, *147*, 034102, DOI: 10.1063/1.4993214 (pages 63, 69).

[179] Zimmerman, P. M. Incremental full configuration interaction. *J. Chem. Phys.* **2017**, *146*, 104102, DOI: 10.1063/1.4977727 (pages 63, 69).

[180] Li, J.; Otten, M.; Holmes, A. A.; Sharma, S.; Umrigar, C. J. Fast semistochastic heat-bath configuration interaction. *J. Chem. Phys.* **2018**, *149*, 214110, DOI: 10.1063/1.5055390 (pages 63, 69).

[181] Loos, P. F.; Boggio-Pasqua, M.; Scemama, A.; Caffarel, M.; Jacquemin, D. Reference energies for double excitations. *J. Chem. Theory Comput.* **2019**, *15*, in press, DOI: 10.1021/acs.jctc.8b01205 (pages 63, 64, 69, 70, 79, 83, 97).

[182] Shih, S.; Butscher W ans Buenker, R.; Peyerimhoff, S. Calculation of vertical electronic-spectrum of nitrogen molecule using mrd-ci method. *Chemical Physics* **1978**, *29*, 241–252, DOI: 10.1016/0301-0104(78)85076-9 (page 63).

[183] Buenker, R.; Peyerimhoff, S.; Butscher, W. Applicability of multi-reference double-excitation CI (MRD-CI) method to calculation of electronic wavefunctions and comparison with related techniques. *Molecular Physics* **1978**, *35*, 771–791, DOI: 10.1080/00268977800100581 (page 63).

[184] Engels, B.; Hanrath, M.; Lennartz, C. Individually selecting multi-reference CI and its application to biradicalic cyclizations. *Computers & Chemistry* **2001**, *25*, 15–38, DOI: 10.1016/S0097-8485(00)00086-3 (page 63).

[185] Musch, P.; Engels, B. DIESEL-MP2: A new program to perform large-scale multi-reference-MP2 computations. *Journal of Computational Chemistry* **2006**, *27*, 1055–1062, DOI: 10.1002/jcc.20416 (page 63).

[186] Quantum Package., [Online; accessed 11. Feb. 2019], 2019 (page 63).

[187] Povill, À.; Rubio, J. An efficient improvement of the string-based direct selected CI algorithm. *Theor. Chem. Acc.* **1995**, *92*, 305–313, DOI: 10.1007/bf01113870 (pages 63, 104).

[188] Olson, R. M.; Bentz, J. L.; Kendall, R. A.; Schmidt, M. W.; Gordon, M. S. A Novel Approach to Parallel Coupled Cluster Calculations: Combining Distributed and Shared Memory Techniques for Modern Cluster Based Systems. *J. Chem. Theory Comput.* **2007**, *3*, 1312 (page 64).

[189] Kjaergaard, T.; Baudin, P.; Bykov, D.; Kristensen, K.; Jorgensen, P. The divide-expand-consolidate coupled cluster scheme. *WIREs Comput. Mol. Sci.* **2017**, *7*, e1319 (page 64).

[190] Kantian, A.; Dolfi, M.; Troyer, M.; Giamarchi, T. Understanding repulsively mediated superconductivity of correlated electrons via massively parallel DMRG. *arXiv* **2019**, 1903.12184 (page 64).

[191] Blase, X.; Duchemin, I.; Jacquemin, D. The Bethe–Salpeter Equation in Chemistry: Relations with TD-DFT, Applications and Challenges. *Chem. Soc. Rev.* **2018**, *47*, 1022–1043, DOI: 10.1039/C7CS00049A (page 64).

[192] Scemama, A.; Caffarel, M.; Oseret, E.; Jalby, W., QMC=Chem: A Quantum Monte Carlo Program for Large-Scale Simulations in Chemistry at the Petascale Level and beyond. In *Lecture Notes in Computer Science*; Springer Berlin Heidelberg: 2013, pp 118–127, DOI: 10.1007/978-3-642-38718-0_14 (pages 64, 96).

[193] Scemama, A.; Applencourt, T.; Giner, E.; Caffarel, M. Quantum Monte Carlo with very large multideterminant wavefunctions. *J. Comput. Chem.* **2016**, *37*, 1866–1875, DOI: 10.1002/jcc.24382 (pages 64, 104).

[194] Kim, J.; Baczewski, A. D.; Beaudet, T. D.; Benali, A.; Bennett, M. C.; Berrill, M. A.; Blunt, N. S.; Borda, E. J. L.; Casula, M.; Ceperley, D. M.; Chiesa, S.; Clark, B. K.; Clay, R. C.; Delaney, K. T.; Dewing, M.; Esler, K. P.; Hao, H.; Heinonen, O.; Kent, P. R. C.; Krogel, J. T.; Kylänpää, I.; Li, Y. W.; Lopez, M. G.; Luo, Y.; Malone, F. D.; Martin, R. M.; Mathuriya, A.; McMinis, J.; Melton, C. A.; Mitas, L.; Morales, M. A.; Neuscamman, E.; Parker, W. D.; Flores, S. D. P.; Romero, N. A.; Rubenstein, B. M.; Shea, J. A. R.; Shin, H.; Shulenburger, L.; Tillack, A. F.; Townsend, J. P.; Tubman, N. M.; Goetz, B. V. D.; Vincent, J. E.; Yang, D. C.; Yang, Y.; Zhang, S.; Zhao, L. QMCPACK: an open source ab initio quantum Monte Carlo package for the electronic structure of atoms, molecules and solids. *J. Phys. Cond. Mat.* **2018**, *30*, 195901, DOI: 10.1088/1361-648x/aab9c3 (pages 64, 96).

[195] Scemama, A.; Applencourt, T.; Giner, E.; Caffarel, M. Accurate nonrelativistic ground-state energies of 3d transition metal atoms. *J. Chem. Phys.* **2014**, *141*, 244110, DOI: 10.1063/1.4903985 (page 64).

[196] Giner, E.; Tew, D. P.; Garniron, Y.; Alavi, A. Interplay between Electronic Correlation and Metal–Ligand Delocalization in the Spectroscopy of Transition Metal Compounds:

Case Study on a Series of Planar Cu^{2+} Complexes. *J. Chem. Theory Comput.* **2018**, *14*, 6240–6252, DOI: 10.1021/acs.jctc.8b00591 (pages 64, 85).

[197] Dash, M.; Moroni, S.; Scemama, A.; Filippi, C. Perturbatively Selected Configuration-Interaction Wave Functions for Efficient Geometry Optimization in Quantum Monte Carlo. *J. Chem. Theory Comput.* **2018**, *14*, 4176–4182, DOI: 10.1021/acs.jctc.8b00393 (pages 64, 97).

[198] Pineda Flores, S. D.; Neuscamman, E. Excited State Specific Multi-Slater Jastrow Wave Functions. *arXiv* **2018** (pages 64, 97).

[199] Löwdin, P. Correlation Problem in Many-Electron Quantum Mechanics I. Review of Different Approaches and Discussion of Some Current Ideas. *Adv. Chem. Phys.* **1959**, *2*, 207 (page 65).

[200] Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Reviews of Modern Physics* **1951**, *23*, 69–89, DOI: 10.1103/revmodphys.23.69 (page 66).

[201] Obara, S.; Saika, A. Efficient recursive computation of molecular integrals over Cartesian Gaussian functions. *J. Chem. Phys.* **1986**, *84*, 3963–3974, DOI: 10.1063/1.450106 (page 67).

[202] Head-Gordon, M.; Pople, J. A. A method for two-electron Gaussian integral and integral derivative evaluation using recurrence relations. *J. Chem. Phys.* **1988**, *89*, 5777–5786, DOI: 10.1063/1.455553 (page 67).

[203] Ten-no, S. An efficient algorithm for electron repulsion integrals over contracted Gaussian-type functions. *Chem. Phys. Lett.* **1993**, *211*, 259–264, DOI: 10.1016/0009-2614(93)85195-t (page 67).

[204] Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. An efficient algorithm for the generation of two-electron repulsion integrals over gaussian basis functions. *Int. J. Quantum Chem.* **1989**, *36*, 269–280, DOI: 10.1002/qua.560360831 (page 67).

[205] Gill, P. M. W.; Pople, J. A. The prism algorithm for two-electron integrals. *Int. J. Quantum Chem.* **1991**, *40*, 753–772, DOI: 10.1002/qua.560400605 (page 67).

[206] E. F. Valeev Libint: A library for the evaluation of molecular integrals of many-body operators over Gaussian functions., http://libint.valeyev.net/, version 2.5.0-beta.1, 2018 (page 67).

[207] Barca, G. M.; Loos, P.-F. Three-and Four-Electron Integrals Involving Gaussian Geminals: Fundamental Integrals, Upper Bounds, and Recurrence Relations. *J. Chem. Phys.* **2017**, *147*, 024103 (page 67).

[208] Zhang, J. libreta: Computerized Optimization and Code Synthesis for Electron Repulsion Integral Evaluation. *J. Chem. Theory Comput.* **2018**, *14*, 572–587, DOI: 10.1021/acs.jctc.7b00788 (page 67).

[209] Wilson, S., Four-Index Transformations. In *Methods in Computational Chemistry*; Springer US: 1987, pp 251–309, DOI: 10.1007/978-1-4899-1983-0_3 (page 67).

[210] Rajbhandari, S.; Rastello, F.; Kowalski, K.; Krishnamoorthy, S.; Sadayappan, P. In *Proceedings of the 22nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming - PPoPP '17*, ACM Press: 2017, DOI: 10.1145/3018743.3018771 (page 67).

[211] Limaye, A. C.; Gadre, S. R. A general parallel solution to the integral transformation and second-order Møller-Plesset energy evaluation on distributed memory parallel machines. *J. Chem. Phys.* **1994**, *100*, 1303–1307, DOI: 10.1063/1.466659 (page 67).

[212] Fletcher, G.; Schmidt, M.; Gordon, M. Developments in parallel electronic structure theory. *Advances in chemical physics* **1999**, *110*, 267–294 (page 67).

[213] Covick, L. A.; Sando, K. M. Four-Index transformation on distributed-memory parallel computers. *J. Comput. Chem.* **1990**, *11*, 1151–1159, DOI: 10.1002/jcc.540111006 (page 67).

[214] Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496–4501, DOI: 10.1063/1.1679012 (page 67).

[215] Eichkorn, K.; Treutler, O.; Oehm, H.; Haeser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* **1995**, *240*, 283, DOI: 10.1016/0009-2614(95)00838-U (page 67).

[216] Schmitz, G.; Madsen, N. K.; Christiansen, O. Atomic-batched tensor decomposed two-electron repulsion integrals. *J. Chem. Phys.* **2017**, *146*, 134112, DOI: 10.1063/1.4979571 (page 67).

[217] Beebe, N. H. F.; Linderberg, J. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *Int. J. Quantum Chem.* **1977**, *12*, 683–705, DOI: 10.1002/qua.560120408 (page 67).

[218] Aquilante, F.; Pedersen, T. B.; Lindh, R. Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals. *J. Chem. Phys.* **2007**, *126*, 194106, DOI: 10.1063/1.2736701 (page 67).

[219] Røeggen, I.; Johansen, T. Cholesky decomposition of the two-electron integral matrix in electronic structure calculations. *J. Chem. Phys.* **2008**, *128*, 194107, DOI: 10.1063/1.2925269 (page 67).

[220] Peng, B.; Kowalski, K. Low-rank Factorization of Electron Integral Tensors and Its Application in Electronic Structure Theory. *Chem. Phys. Lett.* **2017**, *672*, 47 (page 67).

[221] Pham, B. Q.; Gordon, M. S. Compressing the Four-Index Two-Electron Repulsion Integral Matrix using the Resolution-of-the-Identity Approximation Combined with the Rank Factorization Approximation. *J. Chem. Theory Comput.* **2019**, *15*, 2254–2264, DOI: 10.1021/acs.jctc.8b01256 (page 67).

[222] Anderson, J. S. M.; Heidar-Zadeh, F.; Ayers, P. W. Breaking the curse of dimension for the electronic Schrödinger equation with functional analysis. *Comput. Theor. Chem.* **2018**, *1142*, 66–77, DOI: 10.1016/j.comptc.2018.08.017 (page 67).

[223] Knowles, P. J.; Handy, N. C. A new determinant-based full configuration interaction method. *Chem. Phys. Lett.* **1984**, *111*, 315–321, DOI: 10.1016/0009-2614(84)85513-x (page 68).

[224] Greer, J. C. Estimating full configuration interaction limits from a Monte Carlo selection of the expansion space. *J. Chem. Phys.* **1995**, *103*, 1821–1828, DOI: 10.1063/1.469756 (page 69).

[225] Greer, J. C. Monte Carlo Configuration Interaction. *J. Comput. Phys.* **1998**, *146*, 181–202, DOI: 10.1006/jcph.1998.5953 (page 69).

[226] Coe, J. P. Machine Learning Configuration Interaction. *J. Chem. Theory Comput.* **2018**, *14*, 5739, DOI: 10.1021/acs.jctc.8b00849 (page 69).

[227] Nitzsche, L. E.; Davidson, E. R. A perturbation theory calculation on the $^1\pi\pi^*$ state of formamide. *J. Chem. Phys.* **1978**, *68*, 3103–3109, DOI: 10.1063/1.436151 (page 70).

[228] Onida, G.; Reining, L.; Rubio, A. Electronic Excitations: Density-Functional versus Many-Body Green's-Function Approaches. *Rev. Mod. Phys.* **2002**, *74*, 601–659, DOI: 10.1103/RevModPhys.74.601 (page 71).

[229] Reining, L. The GW Approximation: Content, Successes and Limitations: The GW Approximation. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, e1344, DOI: 10.1002/wcms.1344 (page 71).

[230] Loos, P. F.; Romaniello, P.; Berger, J. A. Green Functions and Self-Consistency: Insights From the Spherium Model. *J. Chem. Theory Comput.* **2018**, *14*, 3071, DOI: 10.1021/acs.jctc.8b00260 (page 71).

[231] Veril, M.; Romaniello, P.; Berger, J. A.; Loos, P. F. Unphysical Discontinuities in GW Methods. *J. Chem. Theory Comput.* **2018**, *14*, 5220, DOI: 10.1021/acs.jctc.8b00745 (page 71).

[232] Garniron, Y. Development and parallel implementation of selected configuration interaction methods., Ph.D. Thesis, Université de Toulouse, 2019, DOI: 10.5281/zenodo.2558127 (page 71).

[233] Maurer, W. D.; Lewis, T. G. Hash Table Methods. *ACM Comput. Surv.* **1975**, *7*, 5–19, DOI: 10.1145/356643.356645 (page 72).

[234] Djoudi, L.; Barthou, D.; Carribault, P.; Lemuet, C.; Acquaviva, J.-T.; Jalby, W. In *Workshop on EPIC Architectures and Compiler Technology San Jose, California, United-States*, 2005 (page 72).

[235] Pauncz, R. The Waller-Hartree double determinant in quantum chemistry. *Int. J. Quantum Chem.* **1989**, *35*, 717–719, DOI: 10.1002/qua.560350607 (pages 73, 101, 104).

[236] Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **1975**, *17*, 87–94, DOI: 10.1016/0021-9991(75)90065-0 (page 74).

[237] Liu, B. The simultaneous expansion for the solution of several of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *Numerical Algorithms in Chemistry: Algebraic Method, Lawrence Berkeley Laboratory, University of California, California* **1978**, 49–53 (page 74).

[238] Olsen, J.; Jørgensen, P.; Simons, J. Passing the one-billion limit in full configuration-interaction (FCI) calculations. *Chem. Phys. Lett.* **1990**, *169*, 463–472, DOI: 10.1016/0009-2614(90)85633-N (page 74).

[239] Gadea, F. X. Large matrix diagonalization, comparison of various algorithms and a new proposal. *Chem. Phys. Lett.* **1994**, *227*, 201–210, DOI: 10.1016/0009-2614(94)00796-9 (page 74).

[240] Crouzeix, M.; Philippe, B.; Sadkane, M. The Davidson Method. *SIAM Journal on Scientific Computing* **1994**, *15*, 62–76, DOI: 10.1137/0915004 (page 74).

[241] Giner, E. Coupling Configuration Interaction and quantum Monte Carlo methods: The best of both worlds., Theses, Université de Toulouse, 2014 (page 74).

[242] Ten-no, S. L. Multi-state effective Hamiltonian and size-consistency corrections in stochastic configuration interactions. *J. Chem. Phys.* **2017**, *147*, 244107, DOI: 10.1063/1.5003222 (page 76).

[243] Angeli, C.; Cimiraglia, R.; Persico, M.; Toniolo, A. Multireference perturbation CI I. Extrapolation procedures with CAS or selected zero-order spaces. *Theor. Chem. Acc.* **1997**, *98*, 57–63, DOI: 10.1007/s002140050278 (page 78).

[244] Davidson, E. R.; Nitzche, L. E.; McMurchie, L. E. A modified H_o for Epstein-Nesbet Rayleigh-Schrödinger pertubation theory. *Chemical Physics Letters* **1979**, *62*, 467–468, DOI: 10.1016/0009-2614(79)80742-3 (page 78).

[245] Kozlowski, P. M.; Davidson, E. R. Considerations in constructing a multireference second-order perturbation theory. *J. Chem. Phys.* **1994**, *100*, 3672–3682, DOI: 10.1063/1.466355 (page 78).

[246] Caballol, R.; Malrieu, J. P.; Daudey, J. P.; Castell, O. SCIEL program., Rovira i Virgili University, Tarragona, Spain, 1998 (page 78).

[247] Applencourt, T.; Gasperich, K.; Scemama, A. Spin adaptation with determinant-based selected configuration interaction. *arXiv* **2018** (page 78).

[248] Fales, B. S.; Hohenstein, E. G.; Levine, B. G. Robust and Efficient Spin Purification for Determinantal Configuration Interaction. *J. Chem. Theory Comput.* **2017**, *13*, 4162–4172, DOI: 10.1021/acs.jctc.7b00466 (page 78).

[249] Dagum, L.; Menon, R. OpenMP: An Industry-Standard API for Shared-Memory Programming. *IEEE Comput. Sci. Eng.* **1998**, *5*, 46–55, DOI: 10.1109/99.660313 (page 79).

[250] Hintjens, P., *ZeroMQ*; O'Reilly Media: 2013 (page 79).

[251] Forum, M. P. MPI: A Message-Passing Interface Standard. *University of Tennessee* **1994** (page 79).

[252] Bolosky, W. J.; Scott, M. L. False sharing and its effect on shared memory performance. *USENIX Association* **1993**, 3 (page 81).

[253] Le Guennic, B.; Jacquemin, D. Taking Up the Cyanine Challenge with Quantum Tools. *Acc. Chem. Res.* **2015**, *48*, 530–537, DOI: 10.1021/ar500447q (page 85).

[254] Send, R.; Valsson, O.; Filippi, C. Electronic Excitations of Simple Cyanine Dyes: Reconciling Density Functional and Wave Function Methods. *J. Chem. Theory Comput.* **2011**, *7*, 444–455 (page 85).

[255] Boulanger, P.; Jacquemin, D.; Duchemin, I.; Blase, X. Fast and Accurate Electronic Excitations in Cyanines with the Many-Body Bethe-Salpeter Approach. *J. Chem. Theory Comput.* **2014**, *10*, 1212–1218, DOI: 10.1021/ct401101u (page 85).

[256] Scuseria, G. E.; Schaefer III, H. F. Diatomic Chromium (Cr_2): Application of the Coupled Cluster Method Including All Single and Double Excitation (CCSD). *Chem. Phys. Lett.* **1990**, *174*, 501–503 (page 86).

[257] Roos, B. O.; Andersson, K. Multiconfigurational Perturbation Theory with Level Shift—the Cr_2 Potential Revisited. *Chem. Phys. Lett.* **1995**, *245*, 215–223 (page 86).

[258] Brynda, M.; Gagliardi, L.; Roos, B. O. Analysing the Chromium–Chromium Multiple Bonds Using Multiconfigurational Quantum Chemistry. *Chem. Phys. Lett.* **2009**, *471*, 1–10, DOI: 10.1016/j.cplett.2009.02.006 (page 86).

[259] Coe, J. P.; Murphy, P.; Paterson, M. J. Applying Monte Carlo Configuration Interaction to Transition Metal Dimers: Exploring the Balance between Static and Dynamic Correlation. *Chem. Phys. Lett.* **2014**, *604*, 46–52, DOI: 10.1016/j.cplett.2014.04.050 (page 86).

[260] Purwanto, W.; Zhang, S.; Krakauer, H. An Auxiliary-Field Quantum Monte Carlo Study of the Chromium Dimer. *J. Chem. Phys.* **2015**, *142*, 064302, DOI: 10.1063/1.4906829 (page 86).

[261] Sokolov, A. Y.; Chan, G. K.-L. A Time-Dependent Formulation of Multi-Reference Perturbation Theory. *J. Chem. Phys.* **2016**, *144*, 064102, DOI: 10.1063/1.4941606 (page 86).

[262] Sokolov, A. Y.; Guo, S.; Ronca, E.; Chan, G. K.-L. Time-Dependent *N*-Electron Valence Perturbation Theory with Matrix Product State Reference Wavefunctions for Large Active Spaces and Basis Sets: Applications to the Chromium Dimer and *All-Trans* Polyenes. *J. Chem. Phys.* **2017**, *146*, 244102, DOI: 10.1063/1.4986975 (page 86).

[263] Tsuchimochi, T.; Ten-no, S. Bridging Single- and Multireference Domains for Electron Correlation: Spin-Extended Coupled Electron Pair Approximation. *J. Chem. Theory Comput.* **2017**, *13*, 1667–1681, DOI: 10.1021/acs.jctc.7b00073 (page 86).

[264] Li Manni, G.; Ma, D.; Aquilante, F.; Olsen, J.; Gagliardi, L. SplitGAS Method for Strong Correlation and the Challenging Case of Cr_2. *J. Chem. Theory Comput.* **2013**, *9*, 3375–3384, DOI: 10.1021/ct400046n (page 86).

[265] Vancoillie, S.; Malmqvist, P. Å.; Veryazov, V. Potential Energy Surface of the Chromium Dimer Re-Re-Revisited with Multiconfigurational Perturbation Theory. *J. Chem. Theory Comput.* **2016**, *12*, 1647–1655, DOI: 10.1021/acs.jctc.6b00034 (page 86).

[266] Guo, S.; Watson, M. A.; Hu, W.; Sun, Q.; Chan, G. K.-L. *N* -Electron Valence State Perturbation Theory Based on a Density Matrix Renormalization Group Reference Function, with Applications to the Chromium Dimer and a Trimer Model of Poly(*p*-Phenylenevinylene). *J. Chem. Theory Comput.* **2016**, *12*, 1583–1591, DOI: 10.1021/acs.jctc.5b01225 (page 86).

[267] Scemama, A. IRPF90: a programming environment for high performance computing. *arXiv* **2009** (page 95).

[268] Giner, E.; David, G.; Scemama, A.; Malrieu, J. P. A simple approach to the state-specific MR-CC using the intermediate Hamiltonian formalism. *J. Chem. Phys.* **2016**, *144*, 064101, DOI: 10.1063/1.4940781 (page 96).

[269] Umrigar, C. J.; Filippi, C.; Moroni, S. CHAMP: Cornell-Holland Ab-initio Materials Package., 2018 (page 96).

[270] Giner, E.; Pradines, B.; Ferté, A.; Assaraf, R.; Savin, A.; Toulouse, J. Curing basis-set convergence of wave-function theory using density-functional theory: A systematically improvable approach. *J. Chem. Phys.* **2018**, *149*, 194301, DOI: 10.1063/1.5052714 (page 97).

[271] Paquier, J.; Toulouse, J. Four-component relativistic range-separated density-functional theory: Short-range exchange local-density approximation. *J. Chem. Phys.* **2018**, *149*, 174110 (page 97).

[272] Savin, A., On Degeneracy, Near Degeneracy and Density Functional Theory. In *Recent Developments of Modern Density Functional Theory*, Seminario, J. M., Ed.; Elsevier: Amsterdam, 1996, pp 327–357 (page 97).

[273] Toulouse, J.; Colonna, F.; Savin, A. Long-range–short-range separation of the electron-electron interaction in density-functional theory. *Phys. Rev. A* **2004**, *70*, 062505 (page 97).

[274] Gerber, I. C.; Ángyán, J. G. Hybrid functional with separated range. *Chem. Phys. Lett.* **2005**, *415*, 100–105, DOI: https://doi.org/10.1016/j.cplett.2005.08.060 (page 97).

[275] Ángyán, J. G.; Gerber, I. C.; Savin, A.; Toulouse, J. van der Waals forces in density functional theory: perturbational long-range electron interaction corrections. *Phys. Rev. A* **2005**, *72*, 012510 (page 97).

[276] Paziani, S.; Moroni, S.; Gori-Giorgi, P.; Bachelet, G. B. *Phys. Rev. B* **2006**, *73*, 155111 (page 97).

[277] Goll, E.; Werner, H.-J.; Stoll, H.; Leininger, T.; Gori-Giorgi, P.; Savin, A. A short-range gradient-corrected spin density functional in combination with long-range coupled-cluster methods: Application to alkali-metal rare-gas dimers. *Chem. Phys.* **2006**, *329*, 276 (page 97).

[278] Becke, A. D. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.* **1988**, *88*, 2547–2553, DOI: 10.1063/1.454033 (page 97).

[279] Mura, M. E.; Knowles, P. J. Improved radial grids for quadrature in molecular density-functional calculations. *J. Chem. Phys.* **1996**, *104*, 9848–9858, DOI: 10.1063/1.471749 (page 97).

[280] Ferté, A.; Giner, E.; Toulouse, J. Range-separated multideterminant density-functional theory with a short-range correlation functional of the on-top pair density. *J. Chem. Phys.* **2019**, *150*, 084103, DOI: 10.1063/1.5082638 (page 97).

[281] Scemama, A.; Giner, E. An efficient implementation of Slater-Condon rules. *ArXiv [physics.comp-ph]* **2013**, 1311.6244 (pages 98, 101, 104).

[282] Seeley, J. T.; Richard, M. J.; Love, P. J. The Bravyi-Kitaev transformation for quantum computation of electronic structure. *J. Chem. Phys.* **2012**, *137*, 224109, DOI: 10.1063/1.4768229 (page 101).

[283] Choi, J.; Dongarra, J.; Ostrouchov, S.; Petitet, A.; Walker, D.; Whaley, R. C. *A Proposal for a Set of Parallel Basic Linear Algebra Subprograms*; LAPACK Working Note 100, LAPACK Working Note #100. UT-CS-95-292, May 1995.; Knoxville, TN 37996, USA: Department of Computer Science, University of Tennessee, Knoxville, 1995 (page 104).

[284] Cimiraglia, R. Many-body multireference Møller-Plesset and Epstein-Nesbet perturbation theory: Fast evaluation of second-order energy contributions. *Int. J. Quantum Chem.* **1996**, *60*, 167–171, DOI: 10.1002/(SICI)1097-461X(1996)60:1<167::AID-QUA18>3.0.CO;2-C (pages 105, 106).

[285] Gasperich, K.; Jordan, K. D.; Simons, J. Strategy for creating rational fraction fits to stabilization graph data on metastable electronic states. *Chem. Phys.* **2018**, *515*, 342–349, DOI: 10.1016/j.chemphys.2018.07.019 (page 109).

[286] Hazi, A. U.; Taylor, H. S. Stabilization Method of Calculating Resonance Energies: Model Problem. *Phys. Rev. A* **1970**, *1*, 1109–1120, DOI: 10.1103/PhysRevA.1.1109 (page 110).

[287] Chao, J. S.-Y.; Falcetta, M. F.; Jordan, K. D. Application of the stabilization method to the $N_2^-(1^2\Pi_g)$ and $Mg^-(1^2P)$ temporary anion states. *J. Chem. Phys.* **1990**, *93*, 1125–1135, DOI: 10.1063/1.459176 (pages 110, 112).

[288] Horáček, J.; Paidarová, I.; Čurík, R. Determination of the Resonance Energy and Width of the $^2B_{2g}$ Shape Resonance of Ethylene with the Method of Analytical Continuation in the Coupling Constant. *J. Phys. Chem. A* **2014**, *118*, 6536–6541, DOI: 10.1021/jp503075a (page 110).

[289] Stanton, J. F.; Gauss, J. Perturbative treatment of the similarity transformed Hamiltonian in equation-of-motion coupled-cluster approximations. *J. Chem. Phys.* **1995**, *103*, 1064–1076, DOI: 10.1063/1.469817 (page 110).

[290] Nooijen, M.; Bartlett, R. J. Equation of motion coupled cluster method for electron attachment. *J. Chem. Phys.* **1995**, *102*, 3629–3647, DOI: 10.1063/1.468592 (page 110).

[291] Koopmans, T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. *Physica* **1934**, *1*, 104–113, DOI: 10.1016/S0031-8914(34)90011-2 (page 110).

[292] Schirmer, J.; Cederbaum, L. S.; Walter, O. New approach to the one-particle Green's function for finite Fermi systems. *Phys. Rev. A* **1983**, *28*, 1237–1259, DOI: 10.1103/PhysRevA.28.1237 (page 111).

[293] Haritan, I.; Moiseyev, N. On the calculation of resonances by analytic continuation of eigenvalues from the stabilization graph. *J. Chem. Phys.* **2017**, *147*, 014101, DOI: 10.1063/1.4989867 (pages 111, 118).

[294] Siegert, A. J. F. On the Derivation of the Dispersion Formula for Nuclear Reactions. *Phys. Rev.* **1939**, *56*, 750–752, DOI: 10.1103/PhysRev.56.750 (page 111).

[295] McCurdy, C. W.; McNutt, J. F. On the possibility of analytically continuing stabilization graphs to determine resonance positions and widths accurately. *Chem. Phys. Lett.* **1983**, *94*, 306–310, DOI: 10.1016/0009-2614(83)87093-6 (pages 111–113).

[296] Baker Jr, G. A.; Graves-Morris, P., *Padé Approximants*, 2nd ed.; Cambridge University Press: 1996 (page 112).

[297] Jordan, K. D. Construction of potential energy curves in avoided crossing situations. *Chem. Phys.* **1975**, *9*, 199–204, DOI: 10.1016/0301-0104(75)80130-3 (page 112).

[298] Schlessinger, L. Use of Analyticity in the Calculation of Nonrelativistic Scattering Amplitudes. *Phys. Rev.* **1968**, *167*, 1411–1423, DOI: 10.1103/PhysRev.167.1411 (page 113).

[299] Simons, J. Resonance state lifetimes from stabilization graphs. *J. Chem. Phys.* **1981**, *75*, 2465–2467, DOI: 10.1063/1.442271 (page 115).

[300] Gasperich, K.; Jordan, K. D. unpublished results., 2017 (pages 116, 128).