Theoretical Treatment of Weakly Bound Fermions to Atoms, Molecules, and Clusters

by

Shiv Upadhyay

Bachelor's of Arts, Washington and Jefferson College, 2015

Master of Science, Duquesne University, 2017

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

2023

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Shiv Upadhyay

It was defended on

April 20, 2023

and approved by

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

Geoffrey Hutchison, Associate Professor, Department of Chemistry

Peng Liu, Associate Professor, Department of Chemistry

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

Theoretical Treatment of Weakly Bound Fermions to Atoms, Molecules, and Clusters

Shiv Upadhyay, PhD

University of Pittsburgh, 2023

Certain atoms, molecules, and clusters can bind an excess electron or positron in a diffuse orbital. Given the weak binding energy of these states, usually on the order of tens to hundreds of millielectron volts, the computed binding energies are sensitive to the correlations between fermions in the system. In order to quantitatively describe these systems, one must use computational methods that capture these correlation effects. In this work, we describe several nonvalence anions and diffuse positron bound states, and in the process we gain insight into the nature of correlation.

First, a series of dipole bound anions are treated using stochastic quantum Monte Carlo (QMC) methods. It is found that quantum Monte Carlo methods can accurately describe dipole bound anions in both a first-quantized basis using diffusion Monte Carlo (DMC) and a second-quantized basis using auxiliary field quantum Monte Carlo (AFQMC). Additionally, the use of correlated sampling is shown to increase the efficiency of QMC calculations to resolve the small electron binding energies.

We use the insight gained to develop a strategy for treating nonvalence correlation bound anions. In a follow-up study, a nonvalence anion model $(H_2O)_4$ cluster is treated using QMC methods. The development of an accurate, compact, and relatively computationally inexpensive ansatz for the anionic system is presented, which when used as a trial wave function for DMC calculations performs similarly to more expensive configuration interaction expansions. The importance of higher order correlation effects is explored in both equation of motion coupled cluster calculations and in QMC calculations.

The lessons learned in this work about electron correlation and about the interpretability of correlation effects in changes in the electron density assist in the treatment of the chemisorption of a hydrogen atom on a graphene surface. The weak binding energy of the hydrogen to the graphene sheet poses a theoretical challenge as the energy difference is sen-

sitive to the electron correlation recovered. For example, by changing the functional is used in density functional theory (DFT) calculations can change the predicted binding energy by an order of magnitude. Using DMC calculations, we can resolve this discrepancy and provides a benchmark value of the binding energy. The issues with certain DFT functionals is explored in the context of shifts in the electron density.

In the final portion of this work, the treatment of positronic systems using QMC is explored. Correlation bound positron states are analogues of nonvalence correlation bound (NVCB) anions, and the importance of electron-positron correlation turns out to be even more important in the former case than electron-electron correlation in the NVCB case. This is due to the absence of coulombic repulsion and exchange effects in the positronic case. Using the insight gained with NVCB anions, we develop a similar compact, correlated ansatz for positronic systems, which performs extremely well. The results presented agree extremely well with previous reference values from the literature.

In this work, small energy differences that are extremely sensitive to the treatment of correlations in an accurate and balanced way are described using quantum Monte Carlo methods. The ideas developed here work well for the systems studied and will also scale favorably to larger systems.

Table of Contents

List of Tables

List of Figures

Preface

This thesis is first dedicated to my family – my mother Kailas, my father Navnit, and my sisters Sri and Krupa. I have always found support and encouragement from my parents, and my efforts through graduate school were made possible through this energy. I have looked up to my older sister, Sri, for her dedicated pursuit of her interests both professionally and personally. Krupa, my younger sister, was also an inspiration as she was always empathetic and engaged in interpersonal communication and collectively with her surrounding community as a whole. She will be remembered fondly.

Kenneth D. Jordan was an excellent mentor and provided support and guidance through my entire graduate career. Ken has shaped the way I approach problems as he had an uncanny ability to cut to the heart of an issue and find a simpler system that displays the same characteristics. Former and current graduate students in the Jordan group have been extremely helpful, including but not limited to Amanda Dumi, Kevin Gasperich, Arailym Kairalapova, and Devin Mulvey. There are countless times where I have turned to my peers for advice and found invaluable advice.

I would also like to thank the other theory professors at the University of Pittsburgh, Dr. Chong, Dr. Hutchison, Dr. Liu, Dr. Coalson, and formerly Dr. Lambrecht. I have also benefited from conversations with other theory group graduate students including Eric Berquist, Keith Werling, Dakota Folmsbee, Danielle Hiener, Anthony Bogetti, and Sanjeev Gautam.

Anouar Benali was extremely influential to my development and provided a lot of scientific and financial support. I'd like to thank a number of other collaborators and external friends: Brenda Rubenstein, Hongxia Hao, James Shee, Tina Mihm, Paul Kent, Ye Luo, and Hyeondeok Shin.

I would not have progressed to this point in my scientific career without a lot of guidance along the way. From my undergraduate days at Washington and Jefferson college, interactions with Dr. Iuliucci and Dr. McCracken led me to enjoy science and eventually theoretical chemistry. I began my graduate studies at Duquesne University under Dr. Jeffry D. Madura,

who taught me to be inquisitive, work hard, and to have fun while doing science. His memory will live on in the work of his students. I have to thank the Madura group members who have all helped me along the way particularly Mike Wasko, Matt Srnec, Emily Benner, Kendy Guarinoni, Riley Workman, and Amanda Dumi.

Finally, I want to thank my friends. Amanda Dumi – through many years of graduate school, we became best friends. From our start at Duquesne in the same year, we had a long, and sometimes difficult, path through graduate school. I enjoyed the numerous hours spent together working on problem sets, and later, collaborations. Justin Kinslow, Patrick Brady, Sean Rieger, and Andrew Giel – my close friends of many years. I appreciate your patience as I often flaked on plans through the years, and I look forward to many trips in the future. Rajan Plumley, Riley Workman, Sean Gracy, Nick Gilbert, and Evan Appod – I appreciate the many hours spent virtually to decompress from the stresses of life.

1.0 Introduction

Chemistry involves the study of molecules and materials constructed from nuclei and electrons. The theoretical description of a chemical system requires a consideration of the size of the system and of the target properties. An ideal computational chemistry theoretical model would reproduce experimental observables in a computationally tractable amount of time. In the following sections, nonvalence anions and the related nonvalence positronic bound states are introduced along with the target properties of interest followed by a discussion of the computational chemistry methods used to describe these systems.

1.1 Chemical Systems

1.1.1 Nonvalence Anions

Nonvalence anions are formed when a molecule weakly captures an electron in an extremely diffuse orbital. It is helpful to further differentiate nonvalence anions due to the dominant mechanism of binding into nonvalence electrostatically-bound (NVEB) and nonvalence correlation-bound (NVCB) anions. NVEB anions bind an electron using a permanent electrostatic moment. For example, an electron can be captured by a molecule or molecular cluster with a sufficiently strong dipole moment, creating a NVEB anion also called a dipole bound anion (DBA). Nonvalence correlation-bound anions require electron correlation effects to account for the binding of the electron. This is not to say that the neutral system giving rise to a NVCB lacks any permanent electric moments, but rather, that the uncorrelated description of a system containing these permanent moments will not bind the electron.

1.1.1.1 Nonvalence Electrostatically Bound Anions

DBAs are formed when a molecule with sufficiently large dipole moments (1.625 D in the Born-Oppenheimer approximation) capture an electron in a diffuse, nonvalence orbital..[1–10]

Initially, one may be tempted to think that electron correlation would not play a significant role in such systems given the diffuse nature of the excess electron. However the excess electron is bound far from the valence electrons, there is a dispersion interaction between the excess electron and the neutral molecule.[11,12] However, considering the simplified expression for London dispersion it can be seen that the dispersion interaction is dependent on the polarizability of the two interacting entities.[13]

$$E_{\text{Disp}} = -\left(\frac{3h}{2R^6}\right)(\alpha_i\alpha_j)\left(\frac{\nu_k\nu_j}{\nu_k+\nu_j}\right) \tag{1}$$

where α is the static polarizability, ν is the ionization potential, and R is the separation distance. The diffuse electron is spread over a large volume, and therefore has a large polarizability leading to a large dispersion interaction.

Nonvalence correlation-bound anions are similar to DBAs, but the leading contribution to the binding energy comes from electronic correlation. Often this means NVCB anions do not posses a dipole moment. For DBAs, Hartree-Fock can bind an electron for molecules with dipole moments above the critical threshold, although the electron would be considerably underbound. For NVCB anions, it has been shown that the Hartree-Fock (HF) solution is qualitatively incorrect. Rather than binding the electron, the HF solution collapses onto the continuum yielding a wavefunction for the neutral plus an excess electron in a continuum orbital. When this occurs, methods that build upon the HF result such as single reference perturbation theory, or even coupled cluster methods, will also fail to bind the excess electron. A simple example of this is a model water tetramer cluster that has been studied previously in our group.[14] This model system can be seen in fig. 1. This system has no dipole moment, and the electron is bound by electrostatics and correlation effects. By increasing the separation between the two pairs of water dimers, one can induce a crossover from a nonvalence electrostatically-bound anion to a nonvalence *correlation-bound* anion. DBAs can serve as a pathway to valence-bound anions,[15–18] and nonvalence correlation bound anions can as well.[19]

Fig. 2 shows illustrative potential energy curves for a hypothetical bound anion. The coordinate (R) represents a collective variable which captures the change in geometry (δR) that the neutral molecule undergoes upon capturing an electron. The change in energy

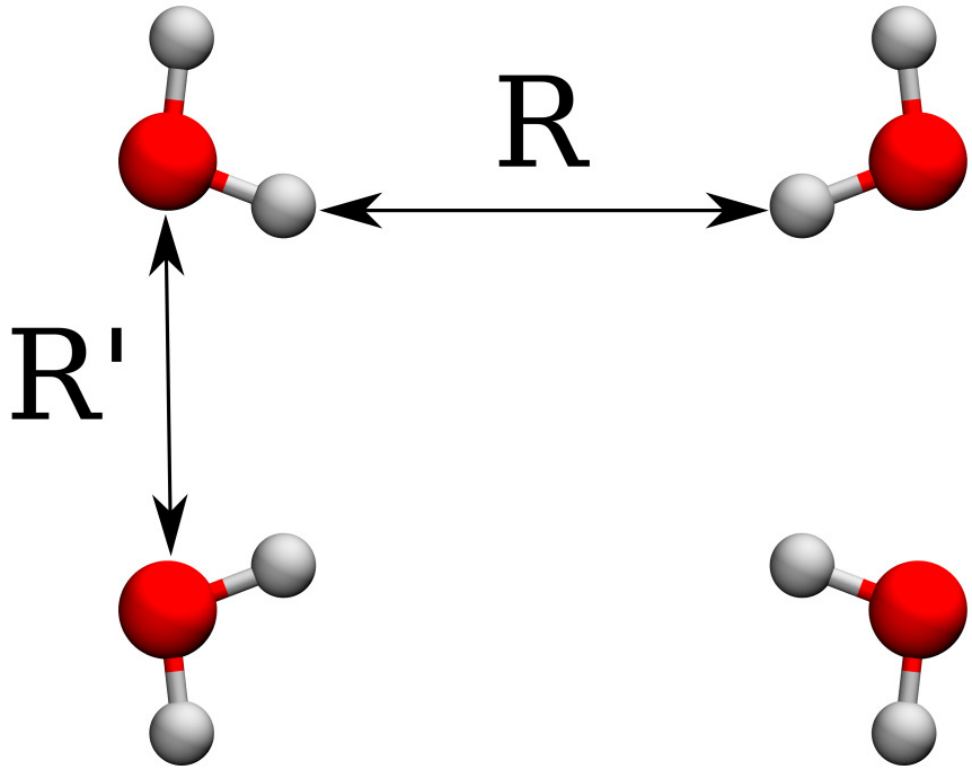


Figure 1: A model $(H_2O)_4$ system. The separation between the dimers R is variable, but the separation of two waters within each dimer $r = 2.775\,14\,\text{Å}$ is fixed.

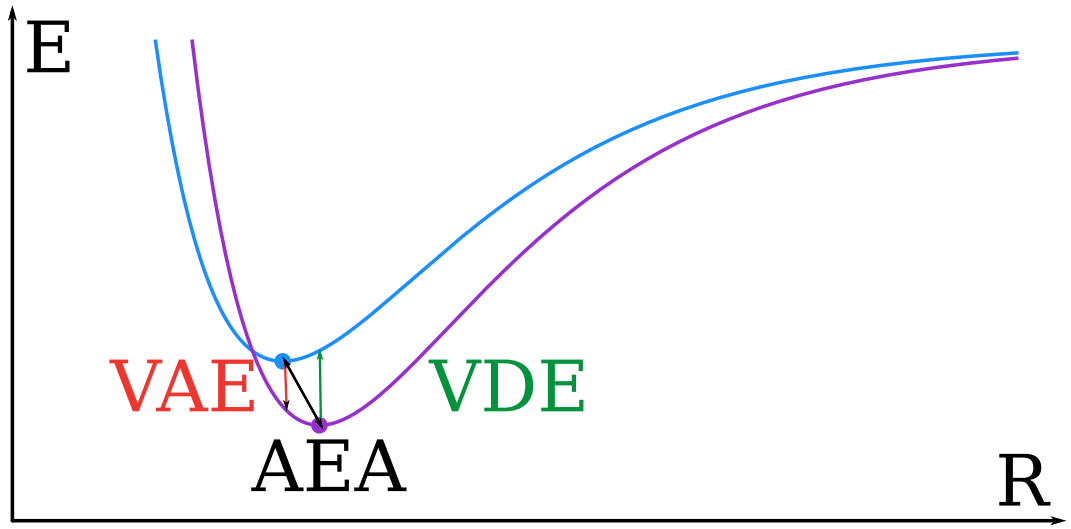


Figure 2: Illustrative potential energy curves for a neutral molecule (blue) and its corresponding dipole-bound or correlation-bound anion (purple) plotted with respect to a fictious collective variable, R, which represents the distortion of the geometry due to the addition of the excess electron. The energy difference at the minimum energy neutral geometry is the vertical attachment energy. The energy difference at the minimum energy anion geometry is the vertical detachment energy. The energy difference between the two respective minima is the **adiabatic electron affinity**.

associated with binding the electron and the subsequent change in geometry is termed the adiabatic electron affinity. The change in energy associated only with binding the electron may occur at either the neutral or anionic geometry and is termed the vertical electron affinity and vertical detachment energy respectively. The extremely diffuse nature of the excess electron in DBAs and NVCB anions often results in a very minor change in geometry. In the limit of zero change in geometry ($\delta R \rightarrow 0$), the adiabatic electron affinity, vertical electron affinity, vertical detachment energy are equal. For the DBAs in this thesis the geometrical distortions are minimal, and therefore it is sufficient to calculate only one of these energies. In practice, and in this work, it is easiest to calculate the vertical electron affinity as this only requires the ground state neutral geometry. The magnitude by which the electron is bound is also referred to as the electron binding energy (EBE). The sign convention adopted is that a positive EBEs indicates that the the electron is bound.

1.1.2 Positron bound states

Since the theoretical prediction[20] and subsequent experimental observation[21] of positrons, the exotic antimatter has seen use in a variety of fields such as positron emission tomography in the medical field or in defect characterization in condensed matter physics.[22–25] Modern applications have also included the formation of antihydrogen as a test of fundamental physics regarding violations of charge, parity, and time-reversal symmetries.[26–30]

In chemistry, one is interested in the behavior of matter. It may seem that the study of antimatter is of little interest to chemistry, however through electron-positron interactions serve as an experimental probe to understand the electronic and vibronic structure of molecules.

In order to understand how positrons yield insight into the behavior of molecules, it helps to consider the interaction of a low energy positron with an idealized molecule. The simplest picture of this process would imply that the positron and electron annihilation rate would be proportional to the number of electrons present.[31] Interestingly, the observed rate of annihilation is orders of magnitude greater for most molecules.[31–37] The major causes of this increased annihilation rate is the formation of positron-molecule bound states and positron-

molecule temporary states (vibrational Feschbach resonances).[31,37–39] Since the process is molecule dependent, the experimental characterization of positron annihilation allows one to characterize different molecules. This also explains the motivation to computationally study positron bound states. The diffuse positronic states are diffuse analogues to nonvalence anions with direct connection to experiment.

1.2 Theoretical Methods

1.2.1 Molecular Quantum Mechanics

Computational chemistry has established itself as a tool for the elucidation and prediction of chemical processes. For a wide variety of problems such as chemical reactivity, molecular structure, and spectroscopic identification, molecular quantum mechanics provides great insight.

1.2.1.1 Schrodinger's Equation

Molecular quantum mechanics involves the solution of a quantum many body problem for a chemical system. Since experimental observables can be represented as quantum mechanical operators, molecular quantum mechanics can both be verified by comparison to experimental results and also provide atomistic insight into experimental results. The work in this thesis will only include nonrelativistic systems and time independent states so the time-independent Schrodinger equation will be the target equation which we seek to solve,

$$H_{mol}\Psi = E\Psi, \tag{2}$$

where

$$H_{mol} = T_{\text{elec}} + T_{\text{nuc}} + V_{\text{elec}-\text{elec}} + V_{\text{elec}-\text{nuc}} + V_{\text{nuc}-\text{nuc}}. \tag{3}$$

The nonrelativistic Schrodinger equation contains the sum over the kinetic energy operators for the electrons and nuclei (T) and their corresponding pairwise interaction potentials (V).

The dimensionality of this problem restricts its solution to simple, few-particle problems such as the hydrogen atom or the hydrogen molecular anion.

1.2.1.2 Born-Oppenheimer Approximation

A common approximation which enables the solution of the Schrodinger equation is the Born-Oppenheimer approximation (BO approximation). This approximation involves an assumption that the nuclei are fixed and the electrons see the static potential of these nuclei. A physical rationalization of this approximation relies on the mass difference between electrons and nuclei as the lightest nucleus, a proton, is roughly 1800 times more massive then an electron. Due to the extreme difference in masses, it can be assumed that any rearrangement of the nuclei results in an instantaneous rearrangement of the electrons. This approach has several practical advantages.

First, the dimensionality of the problem is greatly reduced. By fixing the nuclei, the nuclei degrees of freedom are removed from the the problem. The nuclei's kinetic energy operators are also removed, and the pairwise interaction between the nuclei becomes a constant energy shift.

Second, the BO approximation allows for a logical source for a basis expansion for the electrons. The reason that the hydrogen atom can be solved without invoking the BO approximation is that the Schrödinger equation is separable if the nucleus is assumed to be a point charge. For the hydrogen molecular ion, with prolate spherical coordinates, the Schrödinger equation is again separable.[40] As the number of nuclei and electrons grows attempting to separate the nuclear and electronic equations is impossible. By fixing the nuclei and making the electronic wave function have only a parametric dependence on the nuclei's coordinates, we enable the use of atom centered basis functions. This approximation is therefore central to the popular linear combination of atomic orbitals (LCAO) approach.

Third, the BO approximation gives us the concepts of a molecule and its associated potential energy surface. A potential energy surface is a multidimensional scalar function of a molecule in a particular electronic state in which the energy is a function of the nuclear positions of a molecule. In modern computational chemistry, it is routine that we consider

that we have a molecule and we can optimize the geometry optimization in ground or excited states. One expects that these geometries and the associated observables correspond to ground or excited states of chemical systems. This way of thinking about molecules depends on a decoupling of the nuclei and electron coordinates enabled by the BO approximation.

1.2.2 Basis

Invoking the BO approximation, allows us to define a basis for the electronic degrees of freedom. Since the goal is to represent the electron density that is expected to be localized around the nuclei, a natural basis for representing this density is a set of atomic orbitals (AOs) centered at the nuclei. For computational reasons, Gaussian orbitals are used to represent molecular orbitals (MOs),

$$\psi_{\mathrm{MO}} = \sum_i \Phi_{\mathrm{AO}} = \sum_i N(L,\alpha) Y_{lm}(\theta,\phi) e^{-\alpha|r-R_A|^2}, \tag{4}$$

where r_A is the location of the atomic center where this Gaussian orbital is centered, $Y_{lm}(\theta,\phi)$ is a spherical harmonic, and $N(L,\alpha)$ is a normalization constant. This approach is also called the linear combination of atomic orbitals.

As a simple example for a two electron problem, we write the wave function using the LCAO approach,

$$\Psi = \phi_{MO1}(r_1)\phi_{MO2}(r_2). \tag{5}$$

However, since electrons are Fermions any wave function ansatz must be antisymmetric. The ansatz in eq. 5, also known as the Hartree product, is not antisymmetric,

$$\Psi(1,2) = \phi_{MO1}(r_1)\phi_{MO2}(r_2) \neq -\phi_{MO1}(r_2)\phi_{MO2}(r_1) = -\Psi(2,1). \tag{6}$$

This can be remedied by writing the ansatz wave function as,

$$\Psi = \phi_{MO1}(r_1)\phi_{MO2}(r_2) - \phi_{MO1}(r_2)\phi_{MO2}(r_1), \tag{7}$$

which can be easily verified to be antisymmetric by interchanging the two particles. The expansion in eq. 7 is the expansion of a determinant of a 2×2 matrix. The generalization of the ansatz for N electrons is a Slater determinant (SD),

$$\Psi(r_1, r_2, ..., r_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(r_1) & \phi_2(r_1) & \dots & \phi_N(r_1) \\ \phi_1(r_2) & \phi_2(r_2) & \dots & \phi_N(r_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(r_N) & \phi_2(r_N) & \dots & \phi_N(r_N) \end{vmatrix}. \quad (8)$$

Slater determinants (SDs) serve as the many electron wave function basis for our calculations.

1.2.3 Mean Field Methods

The electronic time independent Schrodinger equation is still a many body problem, and due to its multidimensional nature, it does not lend itself to an easy solution. The common approximation for many body problems in physics is a mean-field (MF) approximation. Mean-field approximations involve mapping an interacting problem onto a noninteracting problem with averaged interactions. The HF method involves mapping the interacting quantum many-body problem onto a noninteracting problem with each electron interacting with the average potential of the other electrons. Since the average potential that each electron experiences is dependent on the density of each of the other electrons, the solution to the HF equations must be found iteratively. This means that the independent particle densities yield averaged interaction potentials that yield new independent particle densities, and this process is repeated. Eventually, the previous iteration's densities are unchanged in an iteration, a condition termed self-consistency. Once this condition has been fulfilled, we stop iterating, and we have found the mean-field solution.

1.2.4 Correlated Methods

Once the HF method, has converged a single Slater determinant wave function is obtained. The description of the wave function captures many of the salient properties of the

system. For example, if we knew the exact nonrelativistic energy for a system, the HF energy accounts for ∼99% of the exact energy. We define the missing energy as the correlation energy,

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}, \tag{9}$$

since it is the portion of the energy missed due to the use of averaged interactions between electrons.

If the correlation energy per electron were constant, then computational chemistry would be much easier. For example, if one were calculating the atomization energy of a water molecule,

$$E_{\text{atom}} = E_{\text{H2O}} - 2E_{\text{H}} - E_{\text{O}}, \tag{10}$$

each term would have a constant shift, and the relative energy difference E_{atom} would be equivalent if calculated with the exact energy or the HF energy. Unfortunately this is not the case, and thus we aim to capture the correlation energy accurately for a chemical system.

In this thesis three main approaches are used to recover electron correlation:

- **Density Function Theory**- Since the Hartree-Fock method neglects e^- correlation by using an averaged potential, density functional theory attempts to modify the effective potential to recover correlation energy.
- **Wave function based methods**- Build upon a Hartree-Fock wavefunction and reincorporate correlation through methods such as perturbation theory.
- **Stochastic methods**- Randomly sample correlated forms of wave functions, which would be difficult or impossible to solve deterministically, to recover correlation.

1.2.4.1 Density Functional Theory Methods

Density Functional Theory (DFT) has seen widespread use since it is an economical, albeit approximate, treatment of electron correlation.[41,42]

DFT in theory is an exact method. The first Hohenberg-Kohn theorem establishes an exact mapping between the interaction potential and the electron density.[43] The second Hohenberg-Kohn theorem shows that a variational principle for the energy as a functional of the density exists.[43] While these theorems and extensions for finite temperature, magnetic

fields, and other generalizations,[44–46] solidify DFT's theoretical footing, they do not indicate how one would practically use DFT.

In practice, the Kohn-Sham equations are used which maps the density-based Schrödinger equation problem on to a fictious noninteracting problem with an effective field,

$$(\hat{T} + \hat{V}_{\text{eff}})\psi(r) = E\psi(r). \tag{11}$$

This formulation resembles a HF equation where the interaction potential has been replaced by an effective potential. This allows one to utilize the standard HF machinery and incorporate electron correlation at minimal cost. The form of the effective potential is what is commonly referred to as the DFT functional. These functionals are constructed and parameterized as the exact functional is not known. This means that although DFT is an exact theory, in practice it is not *ab initio*.

1.2.4.2 Wave Function Based Methods

Wave function based correlated methods build upon a reference, usually the uncorrelated HF reference, and attempt to reincorporate electron correlation. In HF, one assumes the wave function can be written as a single Slater determinant. This single electronic configuration misses electron correlation which can be divided into two categories:

- Static correlation - Some wave functions cannot be written as a single configuration. If there are other electronic configurations of the same symmetry with similar energy, then the true wave function for the system will have a nonneglible contribution from the low lying electronic configurations. A well-known example is the Be atom. The $1s^2 2s^2$ configuration is not significantly lower in energy that the $1s^2 2p^2$ electronic configuration, and therefore these configurations contribute to the true wave function.[47–51]
- Dynamic correlation - Since an averaged interaction potential is used in HF, the instantaneous interactions between electrons are neglected. In a time independent picture this is manifested by a poorly resolved electron-electron cusp. Exact conditions exist for the functional form of the wave function as two electrons coalesce, but HF doesn't obey these Kato cusp conditions.[52] Including additional electronic configurations can improve the description of this cusp.

When a non-minimal basis, meaning there are more basis functions than electrons, is used in HF, the resulting wave function has occupied orbitals and unoccupied virtual orbitals. Other determinants can be constructed from the HF determinant by "exciting" an electron from an occupied to unoccupied orbital. Post-HF methods write an ansatz for the correlated wave function to include these excitations.

1.2.4.3 Full Configuration Interaction

Conceptually the most simple approach is to construct all of the possible excitations from the HF determinant and to write the wave function as a superposition of all of these determinants,

$$|\Psi\rangle = C_0 |\Phi_0\rangle + \sum C |\Phi_\mu^a\rangle + \sum C |\Phi_{\mu\nu}^{ab}\rangle + \cdots . \tag{12}$$

With all possible excitations are generated, the resulting wave function is called the Full Configuration Interaction (FCI) wave function. Once the ansatz has been constructed the coefficients are variationally optimized yielding the correlated wave function. Unfortunately the complexity of the FCI wave function means this method scales as $N!$ where N is the number of electrons. This restricts its usage to small systems with a small number of electrons and orbitals.

1.2.4.4 Truncated Configuration Interaction

The FCI wave function may also be written using excitation operators that generate the excited electronic determinants,

$$|\Psi\rangle = (C_0 \hat{1} + \sum C \hat{R}_\mu^a + \sum C \hat{R}_{\mu\nu}^{ab} + \cdots) |\Phi_0\rangle , \tag{13}$$

where $\hat{R}$ are the excitation operators. Truncated Configuration Interaction (CI) expansions work by limiting the sum of excitation operators. Truncated CI expansions are commonly created by terminating the sum in eq. 13 at a certain excitation level. This forms a family of truncated CI expansions such as Configuration Interaction Singles (CIS), Configuration Interaction Singles and Doubles (CISD), Configuration Interaction Singles, Doubles, and Triples (CISDT), and higher levels.

Although this the most common truncated CI expansion, other truncations exist. For example, senority CI truncates based on the number of unpaired electrons. The most common variant of this type of truncated CI is doubly occupied configuration interaction (DOCI), which is a senority zero CI meaning there are no unpaired electrons.[53–56] DOCI is not a widely used method compared to excitation based truncated CI, but senority based methods have been experiencing renewed interest.[56–63]

The difficulty with truncated CI schemes is that one doesn't have an *a priori* knowledge of the relevant determinants. Selected CI methods, such as perturbatively selected configuration interaction scheme (CIPSI),[64] attempt to address this deficiency by growing the CI space iteratively. In the first iteration, single and double excitations from the HF determinant are considered. The excitations are given a perturbative estimate of their importance, and if the estimated importance is greater than a threshold the excitation is retained. The CI eigenvalue is solved in the space of the retained determinants to obtain variational values of each determinant's contribution. The cycle is then restarted where single and double excitations are generated from the previous iterations retained determinants. The iterative process continues until the CI energy drops below a predefined threshold. This cyclic growth allows for the generation of compact CI expansions with mainly important determinants selected.

1.2.4.5 Coupled cluster methods

Coupled Cluster (CC) methods eschew the linear ansatz of CI, and instead adopt an exponential ansatz,

$$|\Psi\rangle = e^{\hat{T}_1+\hat{T}_2+\cdots}|\Phi_0\rangle\,. \tag{14}$$

This exponential ansatz has the advantage of being size consistent unlike CI methods. Size consistency is a property of quantum mechanical methods where the energy of two noninteracting systems is equal to the sum of each subsystem's energy,

$$E(A+B) = E(A) + E(B). \tag{15}$$

Truncated CI would correlate the subsystems more and yield lower energies for the subsystems. CC by the nature of the exponential ansatz has a multiplicatively separable energy, which leads to size consistency.[65,66]

The extension of coupled cluster to excited states is slight more complicated than the CI case. For CC excited states, the Equation-of-Motion (EOM) formulation has be developed, which conceptually is similar to performing a CI calculation on top of a CC wave function.[67,68] This ansatz is similar to the CI ansatz with excitations applied to the CC wave function,

$$|\Psi\rangle = (C_0\hat{1} + \sum C\hat{R}^a_\mu + \sum C\hat{R}^{ab}_{\mu\nu} + \cdots)e^{\hat{T}_1+\hat{T}_2+\cdots}\,|\Phi_0\rangle\,. \tag{16}$$

In practice, this can be implemented similar to a configuration interaction calculation where the matrix elements come from a similarity transformed Hamiltonian. This method is especially powerful as the excitation operators do not need to conserve particle number so one can calculate ionization potentials and electron affinities as well.[69–72]

1.2.4.6 Stochastic Methods

1.2.4.7 Projector quantum Monte Carlo

This work involves the application of projector quantum Monte Carlo (QMC) methods to locate ground state solutions to the imaginary time Schrödinger equation. The Schrödinger equation in imaginary time ($\tau \to it$) shifted by a constant (E_r) is given by eq. 17.

$$\frac{\partial\,|\Psi\rangle}{\partial\tau} = -(\hat{H} - E_r)\,|\Psi\rangle \tag{17}$$

The formal solution expanded in energy eigenstates is eq. 18.

$$|\Psi(\tau)\rangle = \sum_{i=0}^{\infty} e^{-(\epsilon_i - E_r)\tau}\,|\phi_i\rangle \tag{18}$$

This means any initial state with nonzero ground state overlap converges to the ground state. The constant E_r is introduced to maintain normalization. To illustrate this consider setting $E_r = 0$, all positive energy states would die off exponentially, but all negative energy states would grow exponentially. The ground state would grow the fastest so in the long time limit

it would dominate, but numerically this involves exponential growth. This is mitigated by introducing the E_r, setting it as close to E_0 as possible, and periodically updating it.

The ground state obeys the projection equation (eq. 19).

$$|\Psi_0\rangle \propto \lim_{\tau\to\infty} e^{-\tau(\hat{H}-E_r)}\,|\Psi_T\rangle \tag{19}$$

We now introduce a basis $|B\rangle$ with the only constraint that it is complete. The propagation of a state in imaginary time (multiplying $|B'\rangle$ and inserting a complete basis $\int dB\,|B\rangle\,\langle B| = 1$)

$$\langle B'|\Psi(\tau)\rangle = \int dB\,\langle B'|e^{-(\hat{H}-E_r)\tau}|B\rangle\,\langle B|\Psi(0)\rangle \tag{20}$$

The term $\langle B'|e^{-(\hat{H}-E_r)\tau}|B\rangle$ is the propagator (Green's function), which is unknown. The short time approximation to the Green's function (eq. 21) however can be be constructed by Suzuki-Trotter decomposition of the exponential, which is exact as $\Delta\tau \to 0$.[73,74]

$$\begin{aligned}\langle B'|\Psi(\tau)\rangle = \int dB_n \cdots dB_1 dB\,&\langle B'|e^{-(\hat{H}-E_r)\Delta\tau}|B_n\rangle \\ &\langle B_n|e^{-(\hat{H}-E_r)\Delta\tau}|B_{n-1}\rangle \\ &\cdots \\ &\langle B_2|e^{-(\hat{H}-E_r)\Delta\tau}|B_1\rangle\,\langle B_1|\Psi(0)\rangle\end{aligned} \tag{21}$$

For DMC, $|B\rangle$ is chosen to be real space $|R\rangle$. For auxiliary field quantum Monte Carlo (AFQMC), $|B\rangle$ is chosen to be an over-complete set of nonorthogonal determinants $|D\rangle$. This space may seem complicated, but its use is motivated by the fact that Fermionic antisymmetry is built in. In order to use an exponential propagator, the space must be over-complete and nonorthogonal.[75–77] The methods differ in how the short time approximation to the Green's function is constructed.

1.2.4.8 Diffusion Monte Carlo (DMC)

DMC performs the projection in real space.[78,79]

$$\Psi(R', \tau + \Delta\tau) = \int_{\mathbb{R}^3} G(R \to R', \Delta\tau)\Psi(R, \tau) \tag{22}$$

As previously stated the short time approximation to the Green's function can be constructed (eq. 23).

$$G(R \to R', \Delta\tau) \approx \frac{1}{(2\pi\Delta\tau)^{3N/2}} e^{-\frac{(R-R')^2}{2\Delta\tau}} e^{-\Delta\tau[\frac{V(R)+V(R')}{2} - E_T]} \tag{23}$$

The first exponential term represents a diffusion term and the second represents a branching process. The stochastic evaluation of these processes is known from diffusing species undergoing chemical reaction, hence the name diffusion Monte Carlo.

A problem with this approach is that the lowest energy solution is Bosonic, and therefore propagating with the unmodified projector (eq. 23) leads to the Bosonic ground state. In order to locate the Fermionic ground state, the mixed distribution $f(R, \tau) = \Psi_t(R)\Phi(R, \tau)$ is propagated where Ψ_t is a trial wave function and Φ is the distribution of the "true" distribution. Applying the propagator (eq. 20) to the mixed distribution results in the importance-sampled propagator (eq. 24).

$$\tilde{G}(R \to R', \Delta\tau) = \Psi_t(R')G(R \to R', \Delta\tau)\frac{1}{\Psi_t(R)} \tag{24}$$

The importance-sampled propagator is a similarity transformation of the standard propagator. By introducing the trial wave function Ψ_t, the nodes of the wave function can be fixed allowing one to locate the Fermionic ground state. In practice, this means propagating configurations according to the short time approximation to the importance-sampled Green's function (eq. 25) and rejecting any moves that cross the nodes of the trial wave function Ψ_t.

$$\tilde{G}(R \to R', \Delta\tau) \approx \frac{1}{(2\pi\Delta\tau)^{3N/2}} e^{-\frac{(R-R'-v(R')\Delta\tau)^2}{2\Delta\tau}} e^{-\frac{\Delta\tau}{2}[E_L(R)+E(R')-2E_T]} \tag{25}$$

1.2.4.9 Auxiliary Field Quantum Monte Carlo (AFQMC)

AFQMC performs projection in Slater determinant space. The short time approximation to the propagator consists of Suzuki-Trotter factorizations as shown in eq. 26, where H_1 consists of the one body operators and H_2 is the two body electron interaction term.

$$e^{-\Delta\tau\hat{H}} \approx e^{-\Delta\hat{H}_1/2} e^{-\Delta\tau\hat{H}_2} e^{-\Delta\tau\hat{H}_1/2} \tag{26}$$

In a determinant space, the application of a one body operator (e.g. $\hat{H}_1$) on a Slater determinant results in another Slater determinant, however in general the application of a two body operator (e.g. $\hat{H}_2$) does not. If at each application of the propagator the number of determinants grew, then the algorithm would scale exponentially. To ensure that the propagation of a determinant results in a single determinant rather than a linear combination of determinants the two body operator must be linearized. The first step of the linearization is to write the second quantized two-electron interaction operator as the square of one body terms.

$$\hat{H}_2 = -\frac{1}{2}\sum_{\alpha}\hat{v}_{\alpha}^2 \tag{27}$$

This linearization is completed by applying the Hubbard-Stratonovich transformation to rewrite the two body operator as a collection of one body operators (eq. 28).[80,81]

$$e^{-\Delta\tau\hat{H}} \approx e^{-\Delta\hat{H}_0/2}\left(\int_{-\infty}^{\infty} P(\phi)\hat{B}(\phi)d\phi\right)e^{-\Delta\tau\hat{H}_0/2} \tag{28}$$

where the terms $P(\phi)$ and $\hat{B}(\phi)$ is given by

$$P(\phi) = \frac{1}{2\pi}e^{-\frac{\phi^2}{2}} \tag{29}$$

$$\hat{B}(\phi) = e^{\sqrt{\tau}\phi\hat{v}_{\alpha}} \tag{30}$$

This transformation maps a set of interacting particles to a set of non-interacting particles that interact with fluctuating external (auxiliary) fields. This in practice amounts to a decomposition of the two-body electron interaction, followed by a sampling of the Gaussian distributed auxiliary fields. Following this the short-time approximation to the propagator is constructed directly and applied to some initial wave function.

The sign problem manifests itself in a slightly different way in AFQMC than in DMC. The antisymmetry of the problem is fulfilled implicitly by the nature of the Slater determinant space. However, the decomposition of the two-body operators leads to an ambiguous phase of the wave function. To alleviate the sign problem, a trial wave function can again be introduced leading to the constrained path AFQMC method.[75,76] The problems dealt with here can also be treated by invoking a special case of the constrained path AFQMC where the phase is fixed to the real plane also called phaseless AFQMC.[77]

1.2.5 Multicomponent Methods

Multicomponent methods allow for the treatment of multiple quantum particle types in quantum chemistry. These methods can be used to coupled the electronic degrees of freedom to the nuclear degrees of freedom by treating light nuclei, usually only hydrogen, quantum mechanically. This allows one to study nuclear quantum effects and go beyond the BO approximation. These methods can also be used to calculate the interactions between electrons and antimatter such as positrons. The advantage of such an approach is that the multicomponent methods are conceptually similar to standard electronic structure methods.

The development of multicomponent methods has occurred over many years beginning from the seminal work of Thomas,[82–85] which was soon followed by the application of multicomponent methods to positrons.[86–90] Since then, the application and developments of multicomponent methods have continued for quantum protons,[91–197] positronic systems,[130,131,198–215] and other quantum particle types.[206,216–223]

2.0 Accurate Predictions of Electron Binding Energies of Dipole-Bound Anions via Quantum Monte Carlo Methods

Reprinted (adapted) with permission from Hao, H.; Shee, J.; Upadhyay, S.; Ataca, C.; Jordan, K. D.; Rubenstein, B. M. Accurate Predictions of Electron Binding Energies of Dipole-Bound Anions via Quantum Monte Carlo Methods. *J. Phys. Chem. Lett.* **2018**, *9*, 6185–6190, DOI: 10.1021/acs.jpclett.8b02733. The author's contribution to the work included developing a recipe to generate appropriate trial wave functions for subsequent quantum Monte Carlo calculations, assisting in the generation of figures, and revising the manuscript.

2.1 Summary

Neutral molecules with sufficiently large dipole moments can bind electrons in diffuse nonvalence orbitals with most of their charge density far from the nuclei, forming so-called dipole-bound anions. Because long-range correlation effects play an important role in the binding of an excess electron and overall binding energies are often only of the order of 10-100s of wave numbers, predictively modeling dipole-bound anions remains a challenge. Here, we demonstrate that quantum Monte Carlo methods can accurately characterize molecular dipole-bound anions with near threshold dipole moments. We also show that correlated sampling Auxiliary Field Quantum Monte Carlo is particularly well-suited for resolving the fine energy differences between the neutral and anionic species. These results shed light on the fundamental limitations of quantum Monte Carlo methods and pave the way toward using them for the study of weakly-bound species that are too large to model using traditional electron structure methods.

2.2 Introduction

Dipole-bound anions are intriguing species that bind excess electrons via their molecular dipole moments.[10,225] As the charge-dipole attraction is governed by a long-range potential that behaves as $1/r^2$ at large r, dipole-bound electrons are delicately bound in diffuse orbitals with most of their charge density located far from the atomic centers of their parent molecules.[1–3] Within the Born-Oppenheimer approximation, the critical dipole moment necessary for binding an electron is 1.625 D,[4–6] but increases to 2.5 D or larger when corrections to the Born-Oppenheimer approximation are made.[7–9,226] Beyond being "doorways" to the formation of valence-bound anions,[15–18] dipole-bound anions may be key contributors to the diffuse interstellar bands, a set of absorption peaks emanating from the interstellar medium whose source has yet to be conclusively identified.[227–233] The sheer experimental challenge of resolving the exceedingly small binding energies of such fragile species has motivated spectroscopists to produce dipole-bound species via electron attachment[234,235] and Rydberg electron transfer,[9,236,237] and to study them via field detachment and photoelectron spectroscopy.[238] From the theoretical perspective, dipole-bound anions are of particular interest because they pose a formidable challenge for *ab initio* methods – only high levels of theory, such as coupled cluster theories combined with large, flexible basis sets, are capable of accurately predicting dipole-bound anion electron binding energies that are often of the order of just a few hundred wave numbers in magnitude.[2,10,69,239,240] However, these highly accurate methods scale steeply with system size, severely restricting the size of systems to which they can be applied.

Herein, we explore the accuracy with which Diffusion Monte Carlo (DMC)[79,241,242] and Auxiliary Field Quantum Monte Carlo (AFQMC),[243–248] two highly accurate, stochastic methods that scale as only $O(N^3) - O(N^4)$ with system size, can model dipole-bound anions, with the aim of uncovering a new set of approaches for modeling dipole- and correlation-bound[14] anions of large molecules. Interestingly, despite the different approximations they employ, we find that both methods reproduce experimental results and distinguish molecules that bind an extra electron from those that do not. Furthermore, we find that a newly-developed correlated sampling AFQMC approach (C-AFQMC)[249] is particularly well-suited

for the task of studying energy differences involving weakly-bound species and converges electron binding energies orders of magnitude faster than stochastic methods that do not employ such sampling.

To gauge the viability of characterizing dipole-bound anions using QMC methods, we compute vertical electron affinities for several systems known to form dipole-bound anions. Vertical electron affinities may be obtained by taking the difference between the energies of the neutral and anionic species both calculated at the neutral geometry. However, since the dipole-bound excess electron makes almost no impact on the geometry of the molecule, we can equate these electron affinities to the electron binding energies (EBEs), which are typically defined using the geometry of the anion. In the following paragraphs, we summarize the calculations performed; further details may be found in the Supplemental Information.

2.3 Computational Methods

In order to compute the EBEs, the neutral geometries were first optimized using the MP2 method[250] together with the aug-cc-pVDZ basis set[251,252] in Gaussian 09.[253–257] Hartree-Fock (HF) wave functions were then generated in Gaussian 09, GAMESS,[258,259] or NWChem[260] for use as trial wave functions, which guide sampling and curb the growth of the sign/phase problem, an exponential decay in the signal to noise ratio, in DMC and AFQMC.[261,262] In order to obtain stable dipole-bound anions, it is essential to use flexible basis sets with very diffuse basis functions. Here, we use the aug-cc-pVDZ basis set augmented with a set of diffuse s and p functions, and in one case, also d functions, located near the positive end of the dipole.[2] DMC[78,79,263] and AFQMC[243,264] are performed on all of the species studied. DMC calculations were conducted using the CASINO package.[263,265] Averages were obtained by sampling the configurations of 4650 walkers for 500,000 or more propagation steps. In order to obtain DMC energies in the zero time-step ($\Delta\tau \to 0$) limit, DMC simulations were conducted at three different time step sizes and then linearly extrapolated to zero time-step to yield the final values reported.

Because the calculation of dipole-bound anion vertical binding energies involves en-

Table 1: EBEs and dipole moments of selected species from experiment and Self-Consistent Field [HF], Coupled Cluster [CCSD(T)], DMC, and C-AFQMC calculations.

		Electron Binding Energy (cm^{-1})				
	Dipole Moment (D)	**Experiment**	**ΔSCF**[a]	**CCSD(T)**[a]	**DMC**[b]	**C-AFQMC**[c]
SO	1.55[266]	NOT-BOUND	-3.84	-4.13	-308.20 ± 70.82	-4.54 ± 0.64
HCN	2.98[266]	13[267]	11.00	7.44	46.17 ± 45.30	10.80 ± 2.95
CH_2CHCN	3.87[266]	56 - 87[9,268]	43.30	61.87	106.63 ± 58.12	65.70 ± 11.03
CH_3CN	3.92[266]	93 - 145[9,268]	50.83	103.00	93.83 ± 36.21	95.85 ± 9.73
C_3H_2	4.14[269]	170 ± 50[270]	54.61	162.08	151.22 ± 64.25[d]	132.45 ± 9.43[e]
$C_3H_2O_3$	4.55[266]	194 ± 24[271]	103.13	163.31	213.98 ± 116.15	157.70 ± 17.96

[a] HF and CCSD(T) calculations were preformed using Gaussian 09. [b] DMC calculations of the anion are based on unrestricted Hartree Fock (UHF) trial wave functions obtained from Gaussian 09 or GAMESS. [c] AFQMC calculations were based on UHF trial wave functions obtained from NWChem. [d] The DMC calculations on the C_3H_2 dipole-bound anion used a restricted open-shell Hartree Fock (ROHF) trial wave function. [e] To be consistent, the C-AFQMC calculations on the C_3H_2 dipole-bound anion were also based on an ROHF trial wave function.

ergy differences between two species with identical molecular geometries, we employed C-AFQMC[249] for the majority of our AFQMC calculations. In this approach, differences between two quantities normally computed separately using independently generated auxiliary field configurations are instead computed using the same set of configurations. For sufficiently similar systems like many dipole-bound anions and their parent neutral molecules, this can result in a systematic cancellation of errors, which markedly reduces the variance associated with calculated observables. In our calculations, a set of randomly-seeded repeat simulations were initialized, and after an initial equilibration period, the mean and standard error of the cumulative averages were computed among the repeats. Convergence is attained when the mean is visually observed to plateau, and when the statistical error falls below a target threshold. The $\Delta\tau \to 0$ limit was estimated via linear extrapolations using simulations performed at $\Delta\tau = 0.01$ and $\Delta\tau = 0.005$ a.u.

2.4 Results

While DMC has been successfully employed to calculate cluster binding energies,[272,273] here we analyze its performance predicting the electron binding energies of molecular dipole-bound anions. As an initial test, we examined whether DMC could faithfully predict whether a given species binds an extra electron or not. Thus, in addition to considering several species known to form dipole-bound anions, we also consider SO, a molecule with a dipole moment of 1.55 D, which is below the threshold required for binding. Our results are summarized in Table 1. DMC calculations correctly predicted that all of the species studied, except SO, would form dipole-bound anions.

In Figure 3, the charge densities of neutral CH_3CN, the CH_3CN anion, and the SO molecule plus an extra electron are compared. While the DMC charge density for the neutral CH_3CN molecule is highly localized in the molecular region, that of the anion is far more diffuse and protrudes continuously out from the positive side of the dipole. The charge density plots of the other dipole-bound anions considered in this work manifest similar features and are reported in the Supplemental Information. In stark contrast, the charge density of SO plus an extra electron consists of two disjoint contributions – one associated with the neutral molecule and a second representative of an additional unbound electron positioned more than 50 Å from the molecule.

Interestingly, even though the fixed-node error in the energies of the neutral and anion are greater in magnitude than the electron binding energy, DMC calculations using single Slater determinant trial wave functions provide semi-qualitatively accurate EBEs of dipole-bound anions. Nevertheless, obtaining quantitatively accurate EBEs with the DMC approach employed here would be too computationally demanding to be practical. As presented in Table 1, achieving DMC statistical error bars smaller than the binding energies can require hundreds of thousands to millions of DMC iterations, even starting from a well-optimized variational wave function. For example, as depicted in Figure 4, statistical fluctuations in the energy of the CH_3CN anion simulated with 4650 walkers hover around 65000 cm^{-1}, meaning

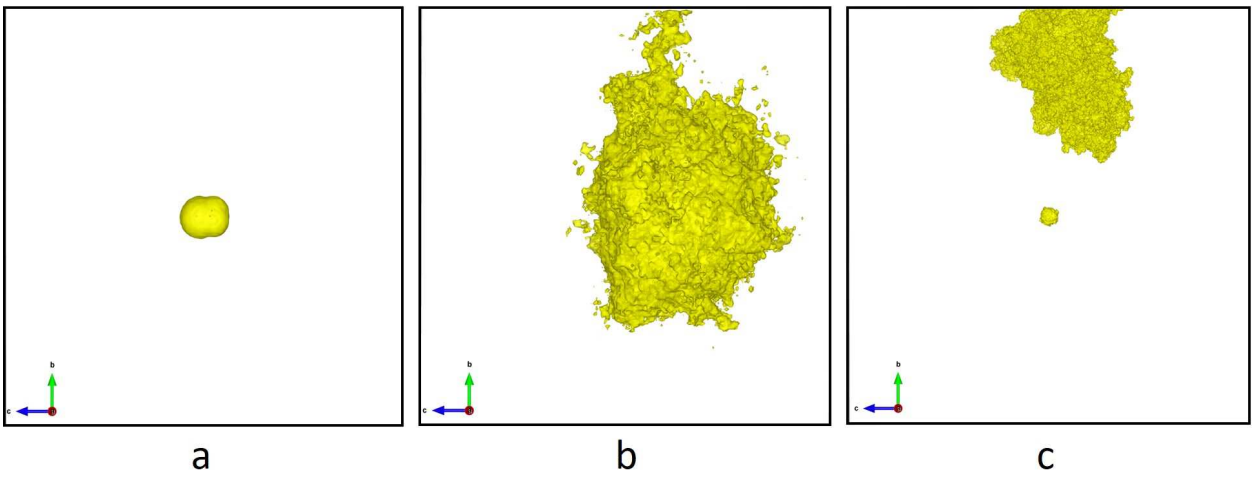


Figure 3: DMC charge densities of (a) neutral CH_3CN, (b) the CH_3CN anion, and (c) SO plus an extra electron. The isosurface values taken for each of these plots are 4×10^{-14} e/Å^3, 4×10^{-14} e/Å^3, and 1×10^{-20} e/Å^3, respectively. Molecules are placed in the center of the simulation box.

that over 500,000 samples must be taken to achieve on the order of 100 cm^{-1} error bars. Thus, while DMC EBEs that agree with experimental measurements and coupled cluster calculations can be gleaned from the noise, it is only with error bars that are still too large to make definitive statements and at great computational expense.

One stochastic technique capable of scaling to dipole-bound anions of large molecules with a substantial reduction in statistical noise is C-AFQMC. As shown in Table 1, C-AFQMC error bars are at least an order of magnitude smaller than DMC error bars using a similar number of samples, but using two orders of magnitude shorter projection time (2000 a.u. was used to obtain most DMC results; 20 a.u. was used to obtain the C-AFQMC results). The C-AFQMC results presented were moreover obtained starting from just HF wave functions. One might question whether the source of this improvement stems from the AFQMC algorithm

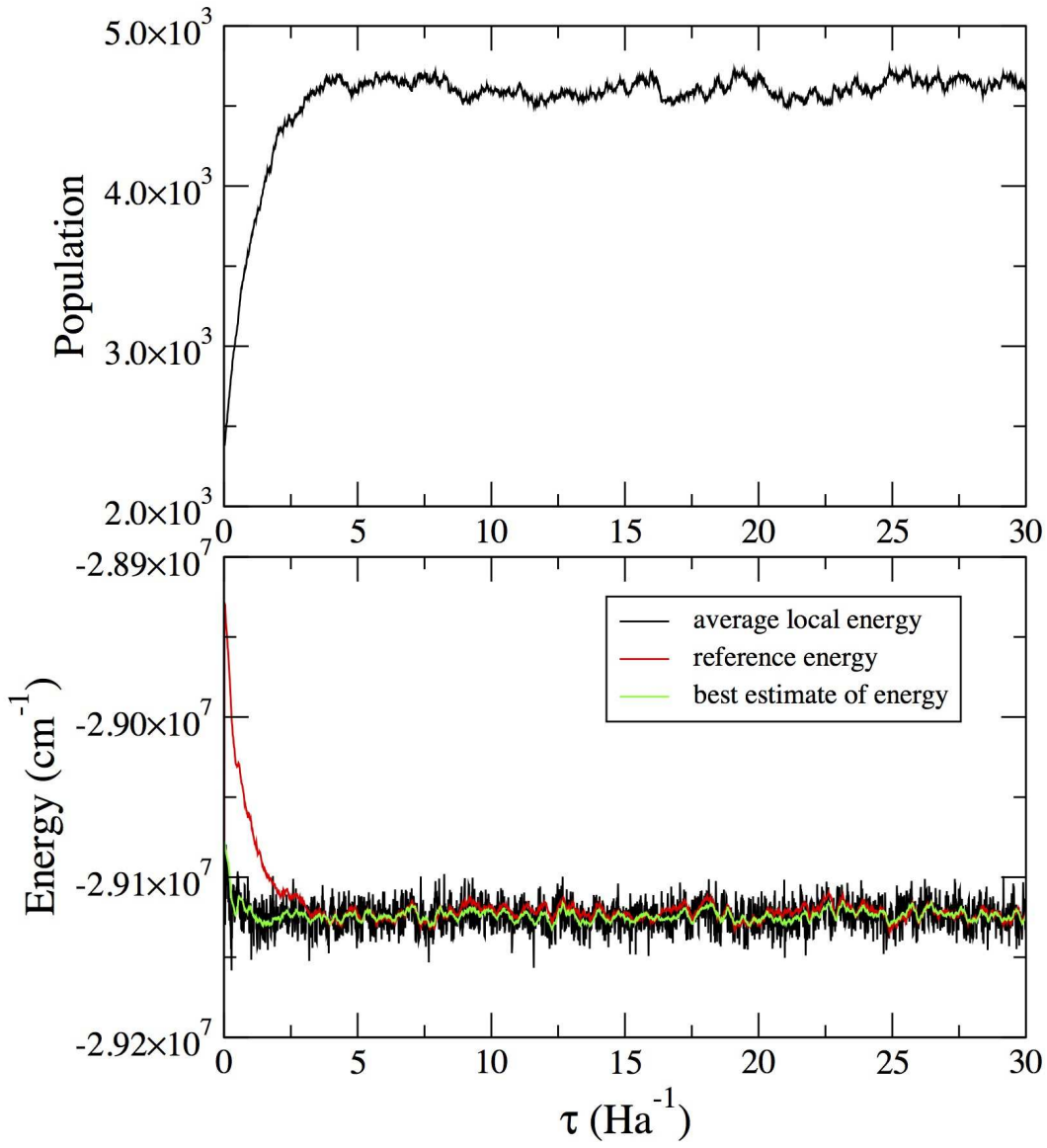


Figure 4: The time evolution of the DMC energy and walker population for the CH_3CN anion using $\Delta\tau = 0.01$ a.u. with 4650 walkers. Walkers were initialized with a UHF trial wave function expanded in terms of the aug-cc-pVDZ basis with a $7s7p$ set of diffuse gaussian-type orbitals. In the Figure, reference energy refers to E_T (see Supplemental Information), the average local energy refers to the local energy averaged over walkers at a given imaginary time, and best estimate of the energy refers to the energy averaged over all samples taken up to a certain imaginary time.

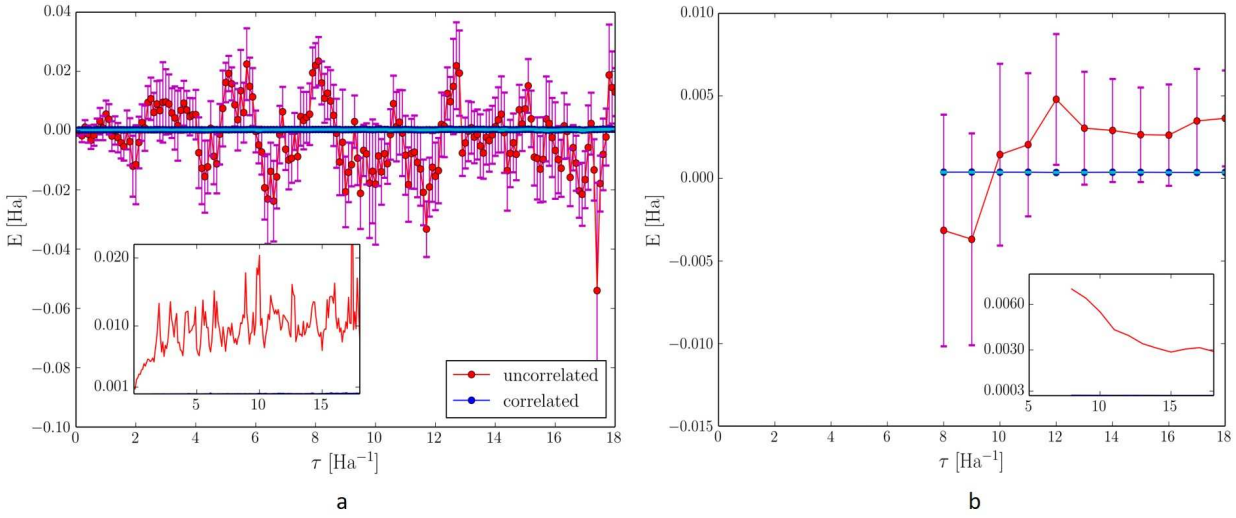


Figure 5: Comparison of energies from AFQMC calculations with and without correlated sampling for CH_3CN using the aug-cc-pvdz+$7s7p$ basis, an HF reference wave function, $\Delta\tau = 0.01$ a.u., 384 walkers per simulation, and 17 repeated simulations. (a) Mean values of the EBE (circles) among the repeats at each τ along the imaginary-time propagation; (b) Mean values of the cumulative averages taken for $\tau > 8$ a.u. The error bars give the standard errors, defined as the standard deviation times $\frac{1}{\sqrt{N_r}}$, where N_r is the number of simulation repeats, and are plotted in the insets for clarity.

or from the use of correlated sampling. Figure 5, which depicts the energy as a function of imaginary projection time for the CH_3CN molecule, demonstrates that the improvement may be attributed to *both* sources. Uncorrelated AFQMC simulations are accompanied by statistical fluctuations on the order of 10^4 cm^{-1}, which are smaller than the 10^5 cm^{-1} fluctuations associated with the EBE values from the DMC simulations (see Figure 4). C-AFQMC calculations are accompanied by almost imperceptible statistical fluctuations on the order of 22 cm^{-1}. Thus, AFQMC's sampling innovations yield meaningful gains above DMC, yet it is the use of correlated sampling that yields, *by far*, the largest improvements.

As shown in Table 1, C-AFQMC yields energies with sufficiently small error bars that meaningful comparisons can now be made against coupled cluster calculations and experiment. In general, the EBEs predicted by C-AFQMC are within error bars of both experimental and previous coupled cluster single, doubles, and perturbative triples (CCSD(T)) calculations. This demonstrates that phaseless approximation errors are mild and that for dipole-bound anions AFQMC is as accurate as CCSD(T). The C_3H_2 anion, however, stands as one cautionary tale. For the C_3H_2 anion, we located two different UHF solutions, neither of which proved suitable for trial functions in QMC calculations (see the Supplemental Information for more details). For this reason, we employed an ROHF trial wave function for the C_3H_2 anion instead. This example suggests that care must be taken when selecting trial wave functions for and ultimately simulating larger molecules that may have many competing low-lying states.

2.5 Conclusions

In this work, we have demonstrated that low-scaling QMC methods, and in particular, C-AFQMC, are capable of resolving the fine energy differences required to accurately predict electron binding energies of dipole-bound anions. Electron binding energies within wave numbers of previous coupled cluster and experimental results were obtained for HCN, CH_2CHCN, CH_3CN, C_3H_2, and $C_3H_2O_3$. Our results demonstrate that, while uncorrelated DMC and AFQMC methods can *qualitatively* describe dipole-bound species, only correlated

sampling techniques such as C-AFQMC are capable of achieving *quantitative* accuracy within computationally tractable amounts of time. The success of C-AFQMC in this work beckons for the further development of correlated DMC methods capable of resolving the molecular energy differences that lie at the heart of all chemical processes. These findings pave the way toward using stochastic methods to study the much larger polycyclic aromatic hydrocarbons (PAH)[274] and long-chain carbon anions[275] thought to contribute to the diffuse interstellar bands, as well as correlation-bound anions[14,276] and weakly bound clusters[277] whose size puts them beyond reach of most high accuracy methods.

2.6 Acknowledgements

H.H., J.S., and B.R. acknowledge Lai-Sheng Wang, Yuan Liu, G. Stephen Kocheril, Shiwei Zhang, and David Reichman for their ongoing support and insights. H.H. and B.R. acknowledge support from the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344, 15-ERD-013 and NSF grant DMR-1726213. S.U. and K.D.J. acknowledge the support of NSF grant CHE-1762337. This research was conducted using computational resources and services at the Center for Computation and Visualization, Brown University, the University of Pittsburgh's Center for Research Computing, and the Extreme Science and Engineering Discovery Environment (XSEDE).

3.0 The Role of High-Order Electron Correlation Effects in a Model System for Non-valence Correlation-bound Anions

The text and figures in this chapter have been reprinted from Upadhyay, S.; Dumi, A.; Shee, J.; Jordan, K. D. The Role of High-Order Electron Correlation Effects in a Model System for Non-Valence Correlation-Bound Anions. *J. Chem. Phys.* **2020**, *153*, 224118, DOI: 10.1063/5.0030942, with the permission of AIP Publishing. The author's contribution to the work included running the EOM calculations, generating the trial wavefunctions and running the VMC/DMC with the exception of the CIPSI trial wavefunctions and the coresponding QMC calculations. Additionally the author's contribution included the generation of figures and writing the manuscript. See Supplementary Materials for geometries, basis sets, additional simulation details, and instructions on generating radial orbital density plots. The data that support the findings of this study are openly available in a public, version-controlled repository at `https://github.com/shivupa/Water4_JCP_Special_Issue_Supplemental_Material`.

3.1 Summary

The diffusion Monte Carlo (DMC), auxiliary field quantum Monte Carlo (AFQMC), and equation-of-motion coupled cluster (EOM-CC) methods are used to calculate the electron binding energy (EBE) of the non-valence anion state of a model $(H_2O)_4$ cluster. Two geometries are considered, one at which the anion is unbound and the other at which it is bound in the Hartree-Fock (HF) approximation. It is demonstrated that DMC calculations can recover from the use of a HF trial wave function that has collapsed onto a discretized continuum solution, although larger electron binding energies are obtained when using a trial wave function for the anion that provides a more realistic description of the charge distribution, and, hence, of the nodal surface. For the geometry at which the cluster has a non-valence correlation-bound anion, both the inclusion of triples in the EOM-CC method

and the inclusion of supplemental diffuse d functions in the basis set are important. DMC calculations with suitable trial wave functions give EBE values in good agreement with our best estimate EOM-CC result. AFQMC using a trial wave function for the anion with a realistic electron density gives a value of the EBE nearly identical to the EOM-CC result when using the same basis set. For the geometry at which the anion is bound in the HF approximation, the inclusion of triple excitations in the EOM-CC calculations is much less important. The best estimate EOM-CC EBE value is in good agreement with the results of DMC calculations with appropriate trial wave functions.

3.2 Introduction

In recent years, there has been growing interest in a class of anions known as non-valence correlation-bound (NVCB) anions in which long-range correlation effects are crucial for the binding of the excess electron.[14,19,276,279–288] By definition, NVCB anions are unbound in the Hartree-Fock (HF) approximation. Due to their highly spatially extended charge distributions, large, flexible basis sets are required for the theoretical characterization of NVCB anions. However, with such basis sets, the wave function from Hartree-Fock (HF) calculations on the excess electron system collapses onto the neutral plus an electron in an orbital that can be viewed as a discretized representation of a continuum solution.[14] Methods that start from the HF wave function including second-order Møller-Plesset perturbation theory (MP2)[250] or coupled-cluster singles and doubles with perturbative triples (CCSD(T))[289] do not recover from this collapse onto the continuum, while methods such as orbital-optimized MP2 (OOMP2)[290] or Bruckner coupled-cluster[291] can overcome this problem.[14] The majority of calculations of NVCB anions have employed the equation-of-motion coupled-cluster singles and doubles (EOM-CCSD) method.[67] Among the NVCB anions studied computationally to date are C_{60}, C_6F_6, TCNE, $(NaCl)_2$, Xe_n clusters, large polyaromatic hydrocarbons, and certain $(H_2O)_n$ clusters.[14,19,276,279,280,285–288]

The EOM-CCSD method displays an $\mathcal{O}(N^6)$ scaling with system size, and higher order EOM-CC methods are even more computationally demanding. As a result, most of the

calculations of NVCB anions carried out to date have not been fully converged with respect to basis set or the level of excitations treated in the EOM procedure. We note, however, that by using domain-based local pair natural orbitals (DLPNO), electron affinity EOM-CCSD calculations have recently been carried out on systems described by up to 4,500 basis functions.[292]

In the present work, we apply two quantum Monte Carlo (QMC) methods to the problem of calculating the electron binding energy (EBE) of the non-valence anion of a model $(H_2O)_4$ cluster. The first approach considered is fixed-node diffusion Monte Carlo (DMC),[79,261,293,294] using various single Slater determinant (SD) and multideterminant (MD) trial wave functions. DMC is a real-space method, with the major sources of error resulting from the use of finite time steps and the fixed-node approximation. The finite time step error can be largely eliminated by running calculations at different time steps and then extrapolating to the zero time step limit. The fixed-node error results from imposition of a nodal surface via a trial wave function, which is necessary to ensure Fermionic behavior, and can be addressed by a variety of means including expanding the number of Slater determinants in the trial wave function or by applying the backflow transformation.[295] It is important to note that, by virtue of working in real space, fixed-node DMC energies are much less sensitive to the choice of the atomic basis set than methods such as EOM-CCSD that operate in a space of Slater determinants.

The second QMC approach considered is the auxiliary field QMC (AFQMC) method.[75,77,243,296–299] AFQMC calculations sample an over-complete space of nonorthogonal Slater determinants. The finite time step error can be mitigated as in DMC. The error that arises from constraining the phase of the wave function to zero can be systematically reduced by improving the trial wave function. Phaseless AFQMC is additionally subject to the limitations of the atomic basis set employed. DMC scales as $\sim \mathcal{O}(N^3)$ with system size, while AFQMC displays an $\sim \mathcal{O}(N^4)$ scaling in most implementations. One of the goals of these calculations is to determine whether DMC calculations can recover from the use of a trial wave function that has collapsed onto a discretized continuum orbital in the case of the excess electron. Additionally, we explore whether correlation effects that are missing in EOM-CCSD are important for electron binding.

3.3 Computational Methods

In our calculations, we employ a model $(H_2O)_4$ cluster that has been investigated in earlier studies by our group.[14,279] In this model, depicted in Figure 6, the monomers are arranged so that the net dipole moment is zero. If the distance R is varied, with all other geometrical parameters held fixed, the system can be tuned from a regime (large R) that the excess electron weakly binds in the HF approximation to one (small R) at which it is not bound in the HF approximation. i.e., at which it is NVCB in nature.

3.4 Methodology

3.4.1 EOM Coupled Cluster

The EOM methods considered in this study are EOM-MP2,[72] EOM-CCSD,[67] EOM-CCSD(T)(a)*,[301] and EOM-CCSDT,[302,303] listed in order of increasing sophistication in terms of treatment of correlation effects. In the EOM-MP2 and EOM-CCSD methods, the neutral molecule is treated at the MP2 and CCSD levels, respectively, and the amplitudes from these calculations are used to perform unitary transformation of the Hamiltonian. This "dressed" Hamiltonian is then used to carry out a 1-particle plus 2-particle-1-hole configuration interaction (CI) calculation on the anion. In the EOM-CCSDT method, the neutral species is first treated at the CCSDT level, and the transformed Hamiltonian is used to do CI calculation on the anion that includes up to 3-particle-2-hole configurations. The EOM-CCSD(T)(a)* method includes in an approximate manner both triple excitations in the ground state coupled cluster calculations and 3-particle-2-hole excitations in the treatment of the anion.[301]

The main basis set used for the EOM calculations reported in this study is aug-cc-pVTZ+7s7p, formed by supplementing the aug-cc-pVTZ Gaussian-type orbital (GTO) basis set[251,252] with a 7s7p set of diffuse functions centered at the middle of the cluster and similar to the set from Ref.The exponents of the supplemental functions start at 0.023622, with each successive exponent being smaller by a factor of 3.2. However, as seen from Table 2,

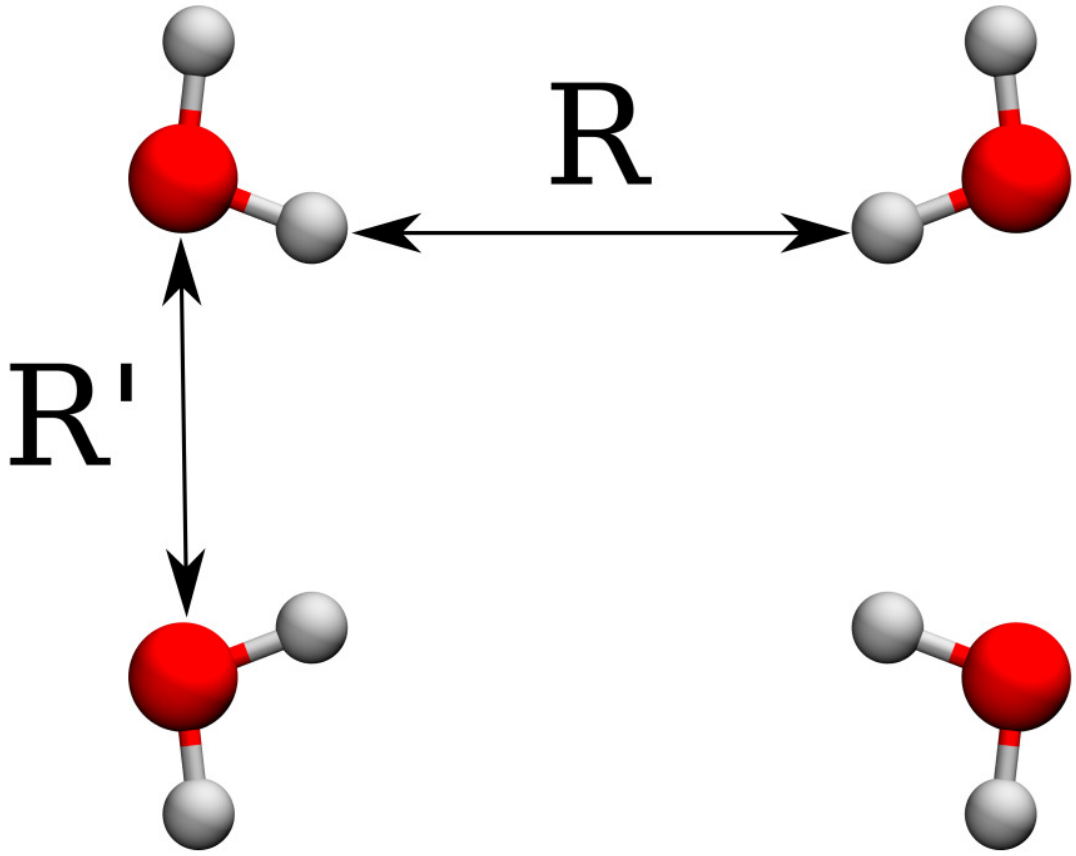


Figure 6: The model $(H_2O)_4$ system considered in this study. R' held fixed at 3.461 05 Å, and R is either 4 Å or 7 Å. Image generated using VMD.[300]

the supplemental 7s7p set of diffuse functions can be truncated to 3s1p without significantly impacting the EBE as calculated at the EOM-CCSD level. Moreover, as shown in Table 3, expanding the main basis set (i.e., the non-supplemented portion) from aug-cc-pVTZ to aug-cc-pVQZ[251,252] makes only a small impact on the EBE (4% at R = 4 Å) . In contrast, reducing the main basis set to aug-cc-pVDZ[251,252] leads to a 14% reduction in the EBE. (These results were obtained using the EOM-MP2 method, but as seen from comparison of the results in Tables 2 and 3, using the aug-cc-pVTZ+3s1p basis set in both cases, the EBEs from the calculations with the EOM-CCSD and EOM-MP2 methods agree to within 0.5 meV.) The smaller aug-cc-pVDZ+3s1p basis will be used in the EOM-CCSDT calculations, which would have been computationally prohibitive with aug-cc-pVTZ+7s7p or aug-cc-pVTZ+3s1p basis sets. Finally, EOM-CCSD(T)(a)* calculations were carried out with aug-cc-pVTZ+3s1p3d basis sets, where the exponents of the d functions match those of the s and p functions, to assess the importance of supplemental d functions on the EBEs. The EOM calculations utilized the frozen core approximation and were carried out using the CFOUR program.[304,305]

3.4.2 DMC

The DMC calculations were carried out using trial wave functions represented as products of one or more Slater determinants with a Jastrow factor with one-, two-, and three-body terms.[306–308] The parameters in the Jastrow factors were optimized using variational Monte Carlo (VMC), and the resulting trial wave functions were then employed in subsequent DMC calculations. Three types of SD trial wave functions were employed. These used HF orbitals, Becke-Lee-Yang-Parr (B3LYP) DFT orbitals,[309–312] and natural orbitals (NOs) from small restricted single plus double excitation configuration interaction (SDCI) calculations designed to bind the excess electron when it is not bound in the HF approximation. In addition, DMC calculation were carried out using MD trial wave functions, with the determinants being determined either from the restricted SDCI procedure or from configuration interaction using a perturbative selection made iteratively (CIPSI) calculations.[64] Details on these calculations are provided below.

Table 2: Dependence of the total energies and the EBE of the model $(H_2O)_4$ cluster at R = 4 Å on the supplemental diffuse basis functions. Results obtained using the EOM-CCSD method.

basis set	neutral (Ha)	anion (Ha)	EBE (meV)
aug-cc-pVTZ	-305.327947	-305.331344	92.4
aug-cc-pVTZ+1s	-305.327953	-305.332359	119.9
aug-cc-pVTZ+2s	-305.327957	-305.334226	170.6
aug-cc-pVTZ+3s	-305.327958	-305.334460	176.9
aug-cc-pVTZ+7s	-305.327958	-305.334462	177.0
aug-cc-pVTZ+7s1p	-305.327979	-305.334604	180.3
aug-cc-pVTZ+7s7p	-305.327987	-305.334622	180.6
aug-cc-pVTZ+3s1p	-305.327979	-305.334602	180.2

Table 3: Sensitivity of the EBE of the $(H_2O)_4$ model to the "core" basis set. Results obtained using the EOM-MP2 method.

	Neutral (Ha)	Anion (Ha)	EBE (meV)
	R = 4.0 Å		
aug-cc-pVDZ+3s1p	-305.0371957	-305.0428558	154.0
aug-cc-pVTZ+3s1p	-305.3092869	-305.3159306	180.8
aug-cc-pVQZ+3s1p	-305.4008845	-305.4078074	188.4
	R = 7.0 Å		
aug-cc-pVDZ+3s1p	-305.0383747	-305.0432259	132.0
aug-cc-pVTZ+3s1p	-305.3104923	-305.3157472	143.0
aug-cc-pVQZ+3s1p	-305.4021640	-305.4075716	147.1

To reduce the computational cost of the DMC calculations, the ccECP pseudopotentials[313,314] were employed together with GTO basis sets that we designate as cc-pVDZ / ccECP, aug-cc-pVDZ / ccECP, aug-cc-pVDZ / ccECP + 3s1p, and aug-cc-pVDZ / ccECP + 7s7p. The "core" cc-pVDZ / ccECP[313,314] basis set was designed for use with the ccECP pseudopotentials; the "aug" indicates that the diffuse aug functions from the aug-cc-pVDZ basis sets of Dunning and co-workers are included; and the 7s7p set of diffuse functions are those described above in the Section 3.4.1.[252] The T-moves scheme was used to control the localization error for nonlocal pseudopotentials.[315]

The double-zeta rather than the larger triple-zeta basis set was used as the core basis set due to the relative insensitivity of DMC calculations to the choice of the atomic basis set. For most of the DMC calculations a fixed population of 16,000 walkers and time steps of 0.001, 0.003, and 0.005 a.u. were employed, with the reported results obtained by linear extrapolation to zero time step. However, this population is much larger and the time steps much smaller than what is actually required to achieve well converged energies with minimized finite time step and fixed population errors. Indeed, DMC calculations using Hartree-Fock trial wave functions, larger time steps (specifically 0.05, 0.1, and 0.2 a.u.) and a smaller population of only 1,000 walkers produce an electron binding energy within error bars of that obtained using the smaller time steps and larger populations. Additionally, a DMC calculation with a B3LYP trial wave function with a time step of 0.05 is in agreement with the values obtained with the smaller time steps and larger populations suggesting that these parameters do not depend strongly on the choice of starting orbitals. In light of this, the 0.05 a.u. time step and smaller walker population were employed in the DMC calculations using CIPSI trial wave functions to mitigate the additional cost associated with the MD space. The VMC and DMC calculations were carried out using the QMCPACK code.[308,316] The orbitals for the SD-based trial wave functions and the restricted SDCI MD wave function were both generated using GAMESS,[258,259,317] whereas the CIPSI wave functions were generated using the Quantum Package 2.0 code.[318]

3.4.3 Restricted CI and CIPSI-generated Trial Wave Functions for DMC Calculations

The restricted SDCI procedure employed the HF wave function for the neutral molecule and a specially tailored SDCI wave function for the anion, which included all symmetry-allowed single and double excitations, with the latter restricted so that one of the electrons excited is from the orbital occupied by the excess electron in the HF wave function. This approach, when used with a flexible basis, gives a bound anion. NOs were generated from the SDCI wave function of the anion and were used in a SD trial wave function for subsequent DMC calculations. In addition, the SDCI wave function itself (expanded in terms of HF orbitals) was used in MD DMC calculations on the anion for R = 4 Å. In this case, a threshold of 0.001 on the magnitude of coefficients in the CI expansion was used in choosing the retained determinants. This resulted in a wave function with 1,392 Slater determinants.

By design, the restricted SDCI wave function does not allow for change of the correlation energy of the valence electrons due to the presence of the excess electron. This possibility is allowed for in the CIPSI MD trial wave functions. The CIPSI calculations were carried out using B3LYP orbitals rather than Hartree-Fock orbitals because the former avoids the problem of collapse onto a discretized continuum solution at R = 4 Å.[309–311] Since the CIPSI calculations have not approached the full configuration interaction limit as indicated by the second-order perturbative correction to the energy, a judicious choice of starting orbitals is required to construct a physically meaningful trial wave function. In order to generate compact wave functions for both the anion and the neutral, NOs were iteratively refined through successive CIPSI calculations, each beginning from a single determinant reference of natural orbitals from the previous iteration. For each NO-generating CIPSI calculation, approximately 100,000 determinants were retained and used to generate NOs for the next iteration, for a total of six NO generation cycles. With the determinant of resulting NOs as a reference, a final CIPSI calculation was carried out, stopping when at least 150,000 determinants were included in the variational space for the anion and at least 100,000 determinants for the neutral. The resulting determinant spaces were used as the DMC trial wave functions.

Both the restricted SDCI and the CIPSI calculations used to generate the trial wave functions for subsequent DMC calculations were carried out using the ccECP pseudopotentials. The aug-cc-pVDZ/ccECP+7s7p and aug-cc-pVDZ/ccECP+3s1p basis sets were used for the SDCI and CIPSI calculations, respectively.

3.4.4 AFQMC

AFQMC[75,77,243,296–299] utilizes the Hubbard-Stratonovich transformation[80] to represent the imaginary-time propagator as a multi-dimensional integral over auxiliary-fields. Ground-state properties are sampled from a random walk in the space of non-orthogonal Slater determinants subject to the phaseless constraint[313] introducing a bias which can be systematically reduced based on the quality of the nodal surface of the trial wave function employed. While sophisticated trial wave functions based on regularized orbital-optimized MP2 (κ-OOMP2)[319] or CASSCF[320–322] are required to obtain quantitative predictions for some biradicaloids and transition metals, high accuracy has been obtained, even for systems exhibiting non-trivial electron correlation such as dipole-bound anions,[224] with single-determinant trial wave functions consisting of HF or Kohn-Sham orbitals.[224,323]

In this work we perform calculations with a GPU implementation of AFQMC,[324] utilizing single-precision floating-point arithmetic and two-electron integrals decomposed via a modified Cholesky decomposition (10^{-5} cutoff).[325] These calculations made use of the aug-cc-pVTZ+7s7p basis set, a small imaginary-time step of 0.005 a.u, and correlated all electrons. For the neutral species and electrostatically bound anion (R = 7 Å), the Hartree-Fock wave function was used as the trial wave function. For the NVCB anionic species (R = 4 Å), a SD trial wave function comprised of natural orbitals from the restricted SDCI calculation as detailed in Section 3.4.3 (but now carried out without pseudopotentials) was used.

3.4.5 Radial orbital densities

To compare the description of the charge distribution of the excess electron as calculated using different theoretical methods, we generate radial electron density plots. This choice is motivated by the fact that the excess electron occupies an orbital belonging to the to-

tally symmetric representation. The radial electron densities are generated by numerically integrating over the angular components of the singly occupied molecular or natural orbital. First, Molden files are created from the output data from the various generating programs using cclib when supported.[326] With the Molden files as input, PySCF is used to generate the electron density on a uniform radial grid and 5810 point Lebedev-Laikov angular grid as tabulated in quadpy.[327–330] Finally, a numerical integration is performed over the angular components. An example of this workflow is presented in detail in the Supplementary Information.

3.5 Results

The EBEs obtained from the EOM and AFQMC calculations are summarized in Table 4, and the results from the various DMC calculations are summarized in Table 5. We consider first the results obtained for R = 4 Å, for which HF calculations do not bind the excess electron.

3.5.1 Results for R = 4 Å: the correlation bound region

From Table 4, it is seen that the EOM-CCSD/aug-cc-pVTZ+7s7p calculations give a value of the EBE of 181 meV for the $(H_2O)_4$ cluster model at R = 4 Å. This increases to 196 meV with the EOM-CCSD(T)(a)* method. The AFQMC calculations using the same basis set and for the anion a single determinant of NOs from the restricted SDCI calculation for the trial wave function produce an EBE value of 194 ± 10 meV, comparable to the EOM-CCSD(T)(a)* result. The EOM-CCSD(T)(a)* and EOM-CCSDT EBE values calculated with this basis set are nearly identical, demonstrating that the approximate treatment of triples in the former procedure introduces a negligible error in the EBE. The contribution of supplemental diffuse functions was checked using the EOM-CCSD(T)(a)* method and the aug-cc-pVTZ+3s1p2d basis set. These calculations reveal that the inclusion of the supplemental diffuse d functions leads to a ∼10 meV increase in the EBE. With the inclusion

of this correction, we obtain an estimated EOM-CCSDT EBE of 212 meV. It is expected that the inclusion of the supplemental d functions in the basis set used for the AFQMC calculations would lead to a similar increase in the EBE obtained using that method.

The restricted SDCI procedure, by itself, is not expected to give an accurate value of the EBE and is designed to generate appropriate trial wave functions for DMC or AFQMC calculations on the anion. In fact, the EBE resulting from the HF treatment of the neutral and the restricted SDCI treatment of the anion using the aug-cc-pVTZ+7s7p basis set is 345 meV, appreciably larger than the EOM and AFQMC values. This over-binding is due in part to the fact that the restricted SDCI wave function, like the HF wave function, overestimates the magnitude of the dipole moment of the water molecules, resulting in a too favorable electrostatic interaction. We also constructed a single determinant trial wave function for the anion using the natural orbitals of the restricted SDCI expansion. We note also that the single determinant of NOs generated from the restricted SDCI wave function and using the aug-cc-pVTZ+7s7p basis set places the anion 160 meV above the neutral when the latter is treated in the HF approximation. This is not surprising since this calculation neglects correlation effects other than those incorporated in the determination of the orbitals. What is important is that the approaches based on the restricted SDCI procedure provide a realistic description of the orbital occupied by the excess electron and avoid the collapse onto the discretized continuum as was observed with the HOMO in the HF calculations.

In light of the close agreement between the EOM-CCSD(T)(a)* and AFQMC values of the EBE of the $(H_2O)_4$ model at R = 4 Å, when using a comparable basis sets in the two approaches it is relevant to determine whether DMC calculations with sufficiently flexible trial wave functions give an EBE close to the AFQMC and EOM values consistent with these results. DMC calculations using HF trial wave functions together with the aug-cc-pVDZ/ccECP+7s7p basis set give an EBE of 183 ± 10 meV, appreciably smaller than the EOM-CCSD(T)(a)* and AFQMC values. Interestingly, essentially the same EBE is obtained from the DMC calculations using a Slater determinant of HF orbitals expanded in the aug-cc-pVDZ/ccECP basis set without the 7s7p supplemental set of diffuse functions. However, if the aug diffuse functions are also removed, the DMC calculations fail to bind the excess electron. We believe that this is a consequence of the fact that with the cc-pVDZ basis

Table 4: EBEs of the $(H_2O)_4$ model calculated using HF, EOM, and AFQMC methods and employing the aug-cc-pVTZ+7s7p basis set.

Method	EBE (meV)
R = 4.0 Å	
HF	-0.4
EOM-CCSD	180.6
EOM-CCSD(T)(a)*	195.8
EOM-CCSDT	197.5[1] (212.0)[2]
AFQMC SD/HF(N)//SD/NO SDCI(A)	194 ± 10
R = 7.0 Å	
HF	41.3
EOM-CCSD	140.2
EOM-CCSD(T)(a)*	141.7
EOM-CCSDT	143.3[1] (154.2)[2]
AFQMC SD/HF	181 ± 5

[1] This EOM-CCSDT/aug-cc-pVTZ+7s7p value was estimated by adding the difference of EBEs from the EOM-CCSD(T)(a)* and eom-ccsdt calculations with the aug-cc-pVDZ+3s1p basis set to the value from EOM-CCSD(T)(a)*/aug-cc-pVTZ+7s7p.

[2] The EOM-CCSDT/aug-cc-pVTZ+7s7p3d value was estimated by adding the difference between the EBEs calculated with the eom-ccsdpt(a)* with the aug-cc-pVTZ+3s1p and aug-cc-pVTZ+3s1p3d basis sets to the eom-ccsdt/aug-cc-pVTZ+7s7p estimated value in footnote [1] to assess the effect of incorporating diffuse d functions into the basis.

set there is a near zero probability of sampling regions of space at large distances from the molecule, which are important for describing the charge distribution of the excess electron.

A significantly larger value of the EBE is obtained from SD DMC calculations using B3LYP orbitals in place of HF orbitals. The resulting EBE of 212 ± 11 meV, within statistical

Table 5: EBEs of the $(H_2O)_4$ model calculated using the DMC method and various trial wave functions.[1]

wave function	basis set	EBE (meV)
	R = 4.0 Å	
SD/HF	aug-cc-pVDZ+7s7p	183 ± 10
SD/HF	aug-cc-pVDZ	176 ± 12
SD/HF	cc-pVDZ	-528 ± 25
SD/B3LYP	aug-cc-pVDZ+7s7p	212 ± 11
SD/HF(N)//SD/NO SDCI(A)	aug-cc-pVDZ+7s7p	205 ± 10
SD/HF(N)//MD/NO SDCI(A)	aug-cc-pVDZ+7s7p	202 ± 12
MD/CIPSI NO	aug-cc-pVDZ+3s1p	190 ± 9
	R = 7.0 Å	
SD/HF	aug-cc-pVDZ+7s7p	141 ± 14
SD/B3LYP	aug-cc-pVDZ+7s7p	164 ± 9
SD/HF(N)//SD/NO SDCI(A)	aug-cc-pVDZ+7s7p	160 ± 9
MD/CIPSI NO	aug-cc-pVDZ+3s1p	159 ± 8

[1] SD/X indicates that the trial wave function employed a single Slater determinant with X (either HF or B3LYP) orbitals. When different types of trial wave functions are used for the neutral (N) and anion (A) this is indicated by the double slash.

error, agrees with the EOM-CCSD(T)(a)* and AFQMC values. A similar value of the EBE is obtained from DMC calculations using a single determinant of HF orbitals for the neutral cluster and a single determinant of natural orbitals from the restricted SDCI procedure described in Section 3.4.3 for the anion. DMC calculations using a SD of HF orbitals for trial wave function of the neutral and a trial wave function for the anion retaining 1,392 of the most important determinants from the restricted SDCI calculation gives an EBE of 202 $\pm$ 12 meV, close to the values obtained using the single determinants B3LYP orbitals or of NOs from the SDCI calculation (for the anion). The DMC value of the EBE resulting from the anionic trial wave function using a SD of NOs from the restricted SDCI MD calculation results is 205 $\pm$ 10meV, similar to that from DMC calculations using as trial wave functions the MD restricted SDCI wave function for the anion and the HF wave function for the neutral.

Figure 7 compares the radial charge distributions of the singly occupied orbital from the HF and B3LYP calculations on the excess electron system as well as of the NOs associated with the excess electron from EOM-CCSD, restricted SDCI and CIPSI calculations. The collapse of the singly occupied orbital from the HF calculations onto a discretized continuum orbital is readily apparent. In contrast, the NOs from the EOM-CCSD and restricted SDCI calculations and the singly occupied orbital from the B3LYP calculation on the anion are more localized and are qualitatively similar to one another. These results are consistent with the nodal surface for the anion being significantly improved when using a SD trial wave function that has a physically reasonable charge distribution for the orbital occupied by the excess electron. Thus, although DMC calculations do recover from the collapse of the HF trial wave function onto a discretized continuum solution in the case of the anion, starting with such a trial function leads to a greater nodal surface error for the anion than for the neutral cluster. However, we also note that the radial distribution function of the singly occupied orbital from the B3LYP calculation on the anion has a spurious peak near 25 atomic units from the center of the cluster. This is likely a consequence of the self-interaction error in the B3LYP functional. The relevant NO extracted from the CIPSI calculations, which were carried using B3LYP orbitals, exhibits a similar shoulder.

Our final set of DMC calculations at R = 4 Å used MD trial wave functions determined

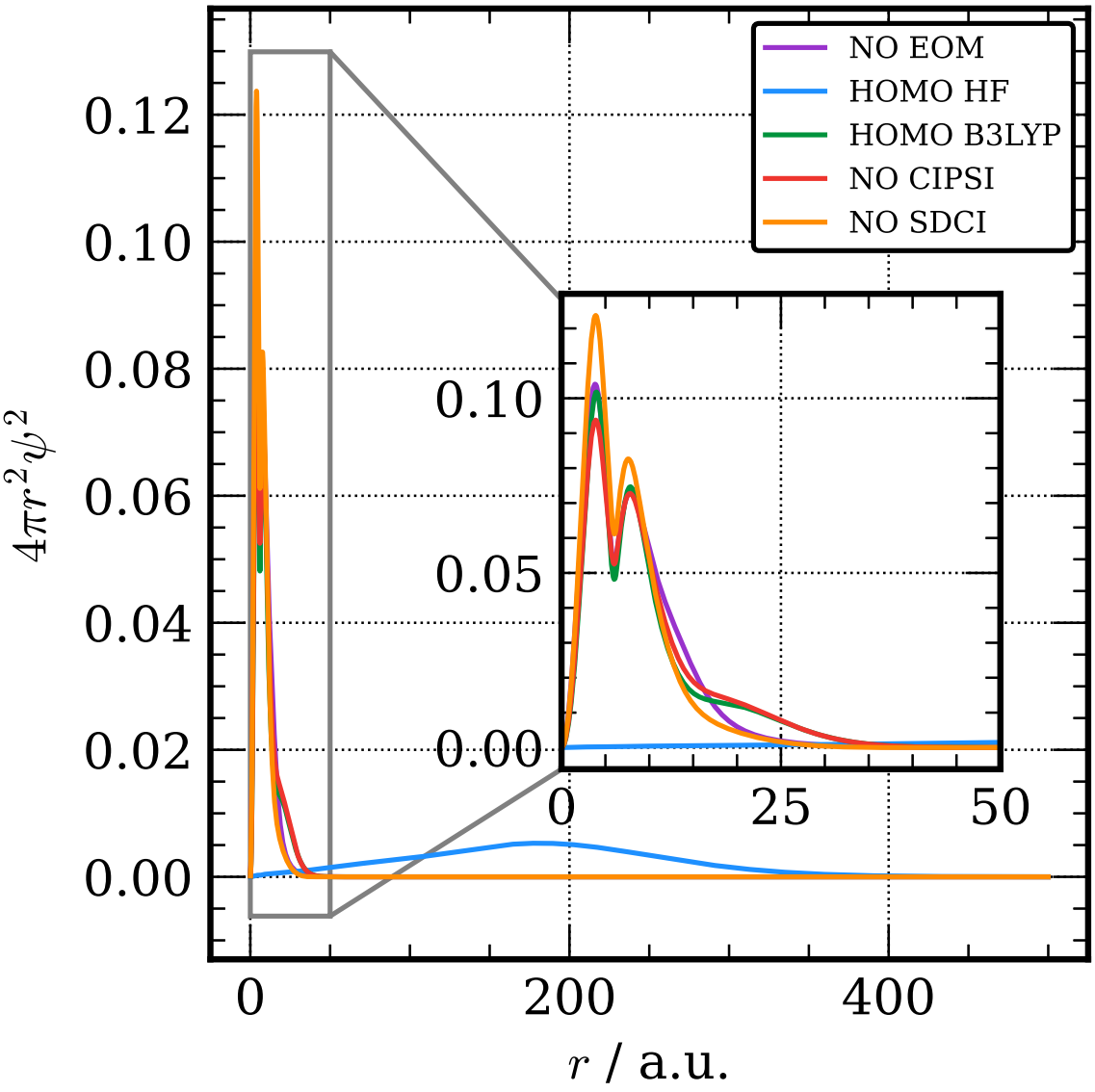


Figure 7: Radially integrated charge densities of the singly occupied orbitals from HF and B3LYP calculations and the singly occupied natural orbital from EOM-CCSD, SDCI, and CIPSI calculations of the model ($(H_2O)_4$) cluster anion at R = 4 Å. All plots generated using Matplotlib.[331]

.

from CIPSI calculations for the neutral and anionic clusters. The strategy used in performing the CIPSI calculations was presented in Section 3.4.3, where it was noted that these calculations, unlike those with the restricted SDCI wave functions, allow for the correlation between the valence electrons change due to the the presence of the excess electron. The DMC calculations using the CIPSI trial wave function resulted 190 ± 9 meV for R = 4 Å, slightly under-binding compared to the single determinant DMC value of the EBE obtained using B3LYP orbitals though in close agreement with the results of DMC calculations carried out with the restricted SDCI trial wave function.

3.5.2 Results for R = 7 Å: the electrostatically bound region

We now consider the results obtained for the $(H_2O)_4$ cluster model at R = 7 Å, for which HF calculations with the aug-cc-pVTZ+7s7p basis set bind the excess electron by 41 meV. In this case, the EOM-CCSD and EOM-CCSD(T)(a)* calculations give EBEs of 140 meV and 142 meV, respectively. Thus unlike the situation for R = 4 Å, the inclusion of triples in the EOM-CC procedure is relatively unimportant at R = 7 Å. The DMC calculations using SD HF trial wave functions give an EBE of 141 ± 14 meV, while the DMC calculations using as trial wave functions single determinants of B3LYP orbitals, single determinants generated using the restricted SDCI procedure, or MD trial wavefunctions generated using the CIPSI procedure give similar EBEs values ranging from 159 ± 8 to 164 ± 9 meV.

Since the anion is bound in the HF approximation at R = 7 Å, we also were able to calculate EBEs using separate, frozen-core coupled-cluster calculations for the neutral and anion with the following coupled-cluster methods: coupled-cluster singles, doubles, and a perturbative treatment of triples ΔCCSD(T),[289] coupled-cluster singles, doubles, and triples (ΔCCSDT),[332–335] and CCSDT with the perturbative treatment of quadruple excitations (ΔCCSDT(Q))[336] methods. The Δ indicates that the EBE is derived from the energy difference between the separate calculations on the neutral and anion. The ΔCCSDT and ΔCCSDT(Q) calculations were carried out with only the aug-cc-pVDZ+3s1p basis set. These calculations indicate that full treatment of the triples, and even approximate treatment of the quadruple excitation contributions, has less than a 1 meV effect on the EBE of the $(H_2O)_4$

cluster model at R = 7.0 Å. On the other hand, the inclusion of diffuse d function in the supplemental set of functions leads to a 12 meV increase in the EBE. With this correction we obtain an estimated EOM-CCSDT EBE of 154 meV, which is in good agreement with the DMC results using suitable trial wave functions.

The AFQMC calculations give an EBE of 181 ± 5 meV, significantly larger than the EOM-CC results or DMC values. This most likely reflects an inadequacy of the HF wave function used for the anion in the AFQMC calculations. Support for this interpretation is provided by examination of Figure 8, which shows the radial charge distribution of the excess electron for the $(H_2O)_4$ model at R = 7 Å. From this figure it is seen that although that the HF wave function has not collapsed onto the continuum as it did in the R = 4 Å cluster, it is still much more diffuse than that from calculations that include correlation effects. It is also seen from comparisons of Figures 7 and 8 that the charge distribution associated with the NO occupied by the excess electron in the EOM-CCSD calculations for the cluster with R = 7 Å, is more radially extended than that at R = 4 Å. Another noticeable difference between the charge density plots for R = 7 Å and 4 Å is the reduction of the long-range shoulder in the radial charge distribution of the HOMO from the B3LYP calculations on the anion and in the relevant NO from the CIPSI calculations on the anion carried out using B3LYP orbitals, suggesting that self-interaction errors are less problematical at R = 7 Å.

3.6 Conclusions

In this study we have applied various EOM-CC methods and two different quantum Monte Carlo methods to calculate the EBE of a model $(H_2O)_4$ cluster at two geometries, one at which the anion is bound in the HF approximation and the other at which it is not. Diffusion Monte Carlo calculations using single determinant trial functions based on Hartree-Fock orbitals are shown to bind the excess electron even when the initial wave function for the anion has collapsed onto the neutral plus discretized continuum orbital. However, such calculations significantly underestimate the EBE, whereas SD DMC calculations using trial wave functions for the anion with a more realistic charge distribution for the excess electron

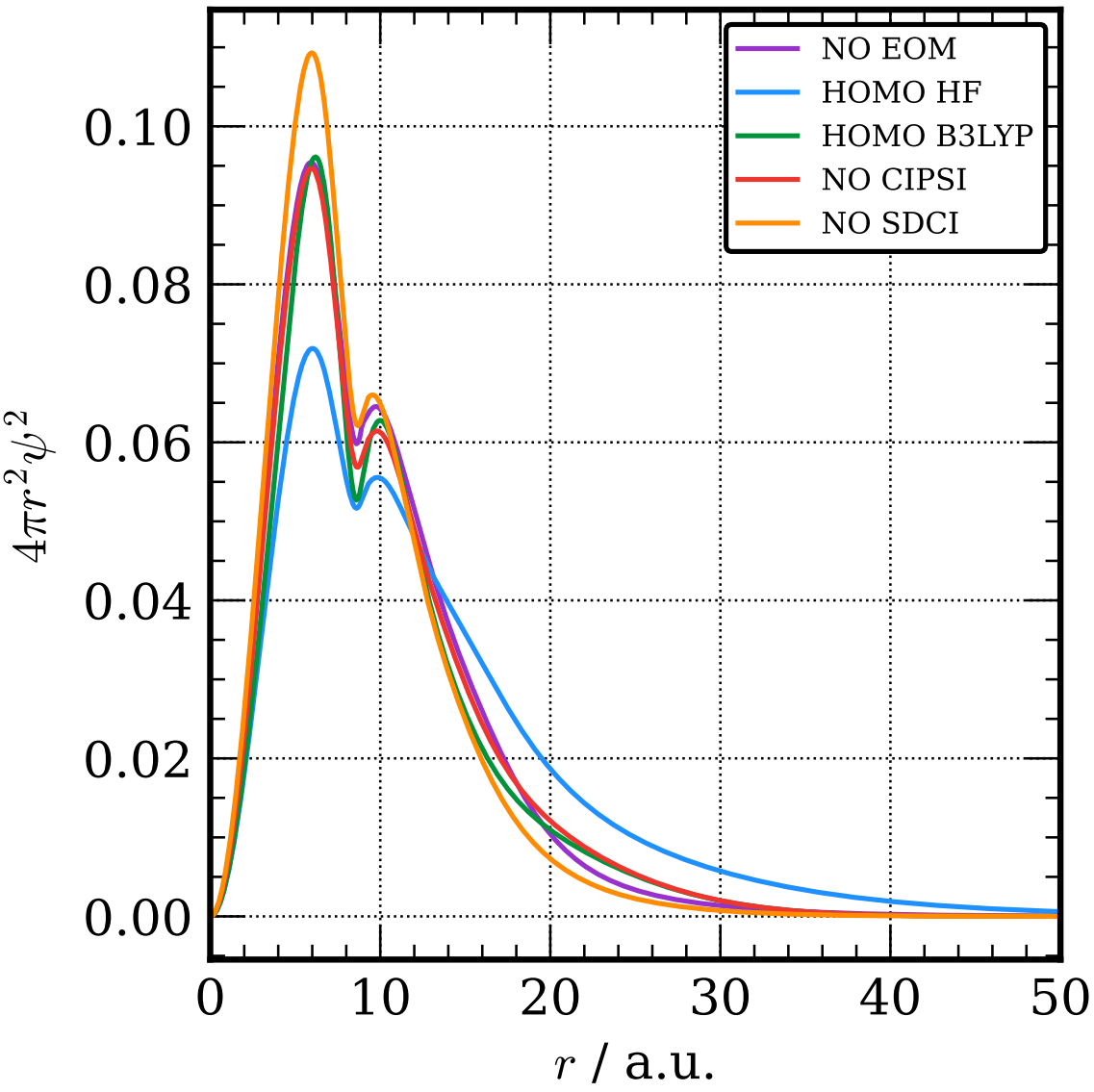


Figure 8: Radially integrated charge densities of the singly occupied orbitals from HF and B3LYP calculations and the singly occupied natural orbital from EOM-CCSD, restricted SDCI, and CIPSI calculations of the model ($(H_2O)_4$) cluster anion at R = 7 Å.

give larger EBE values that are in close agreement with our best estimate EOM-CCSDT values for both geometries considered.

For R = 4 Å, at which the anion is correlation bound, use of such trial wave functions accurately reflecting the physical charge density resulted in AFQMC-predicted EBE values in agreement with the EOM-CCSD(T)(a)* result (when using comparable basis sets). However, at R = 7 Å, AFQMC calculations with HF trial wave functions significantly overestimate the EBE compared to EOM-CC and DMC values, suggesting the need for an improved trial wave functions in this case. For the $(H_2O)_4$ model system, the restricted SDCI represents an economical way to create trial wave functions for QMC calculations on non-valence anions that are not bound in the Hartree-Fock approximation. However, it remains to be seen if this strategy will be as effective for systems in which the neutral species is more strongly correlated than the model $(H_2O)_4$ cluster.

Finally, we note that at R = 4 Å, for which the anion is NVCB in nature, the most frequently used method to characterize such anions, EOM-CCSD, underestimates the EBE by about 10% compared to the result of EOM-CCSDT calculations. Both DMC and AFQMC are viable alternatives to high order EOM methods, and while more computationally demanding for the $(H_2O)_4$ cluster, they demonstrate lower scaling with system size than EOM methods, making them attractive for the characterization of non-valence anions of much larger systems.

3.7 Acknowledgements

We acknowledge valuable discussions with Shiwei Zhang. This research was carried out with the support of grant CHE-1762337 from the National Science Foundation. SU acknowledges fellowship support from the Pittsburgh Quantum Institute. AD acknowledges funding support from National Science Foundation CHE-1807683. Computational resources for the EOM, DMC, and CIPSI calculations were provided by the Center for Research Computing, University of Pittsburgh. The AFQMC calculations used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported

by the Office of Science of the U.S. Department of Energy under contract no. DE-AC05-00OR22725.

4.0 The binding of atomic hydrogen on graphene from density functional theory and diffusion Monte Carlo calculations

The text and figures in this chapter have been reprinted from Dumi, A.; Upadhyay, S.; Bernasconi, L.; Shin, H.; Benali, A.; Jordan, K. D. The Binding of Atomic Hydrogen on Graphene from Density Functional Theory and Diffusion Monte Carlo Calculations. *J. Chem. Phys.* **2022**, *156*, 144702, DOI: 10.1063/5.0085982, with the permission of AIP Publishing. The author's contribution to the work included running the EOM calculations, generating the trial wavefunctions and running the VMC/DMC with the exception of the CIPSI trial wavefunctions and the coresponding QMC calculations. Additionally the author's contribution included the generation of figures and writing the manuscript. The Supplementary Material includes the total energies and error bars for the quantum Monte Carlo calculations, the total energies for the DFT calculations, and details of the convergence of the DFT total energies with respect to the k-point grid and kinetic energy cutoff of the plane wave basis, and a comparison of the density difference of DMC-PBE and DMC-HSE. The data that support the findings of this study are openly available on the Materials Database Facility at https://acdc.alcf.anl.gov/mdf/detail/dumi_dmc_hgraphene_v1.3, with the following DOI: 10.18126/s1wc-tya.

4.1 Summary

In this work density functional theory (DFT) and diffusion Monte Carlo (DMC) methods are used to calculate the binding energy of a H atom chemisorbed on the graphene surface. The DMC value of the binding energy is about 16% smaller in magnitude than the Perdew-Burke-Ernzerhof (PBE) result. The inclusion of exact exchange through the use of the Heyd–Scuseria–Ernzerhof (HSE) functional brings the DFT value of the binding energy closer in line with the DMC result. It is also found that there are significant differences in the charge distributions determined using PBE and DMC approaches.

4.2 Introduction

The unique electronic, optical, and transport properties of graphene make it an important system for a wide range of applications, many of which involve or are impacted by the adsorption of atoms or molecules. To bring these applications to fruition, a deeper understanding of the interaction of atoms and molecules with graphene is required, and, not surprisingly, this has been the subject of several experimental and theoretical studies.[338–350]

The adsorption of H atoms on graphene has been the subject of multiple theoretical studies.[340,347,349,350] It is known that there is both a weakly absorbed state in which barriers for diffusion are small and a much more strongly bound chemisorbed state,[351,352] which is the focus of this work. Chemisorbed H atoms open up the band gap and allow for tuning of electronic properties.[353] It has been demonstrated that even a single chemisorbed hydrogen atom causes an extended magnetic moment in the graphene sheet.[354,355] On the other hand, there is evidence that given the ready diffusion of H in the physisorbed state, the H atoms tend to pair up on the surface leading to non-magnetic species.[355] Finally, interest in the hydrogen/graphene system has also been motivated by the potential use of graphene and graphitic surfaces for hydrogen storage.[346] In spite of the interest in H chemisorbed on graphene, we are unaware of experimental values of the binding energy.

Most computational studies of adsorption of atoms and molecules on graphene have employed density functional theory (DFT), primarily due to its favorable scaling with system size, allowing for the treatment of larger periodic structures. However, a reliable theoretical description of interactions at the graphene surface has proven to be challenging for DFT.[338–340,356] In recent years considerable progress has been made in extending correlated wave function methods to periodic systems.[357–362] Among these methods, the diffusion Monte Carlo (DMC)[337] method, which is a real-space stochastic approach to solving the many-body Schrödinger equation is particularly attractive given its low scaling with the number of electrons and high parallelizability. DMC also has the advantages of being systematically improvable and its energy being much less sensitive to the basis set employed than methods that work in the space of Slater determinants. In DMC calculations, the atomic basis set is important only to the extent that it impacts the nodal surface. DMC has been

used to describe the adsorption of various species on graphene including O_2,[341] a water molecule,[342,363] and a platinum atom.[343] In a study of a physisorbed H atom on graphene, Ma et al. found that different DFT functionals gave binding energies ranging from 5 to 97 meV, while DMC calculations gave a value of only 5 ± 5 meV.[340] Various DFT calculations utilizing the Perdew-Burke-Ernzerhof (PBE)[364] and Perdew-Wang (PW91)[365] functionals predict the chemisorbed H atom species to be bound by 480 to 1,440 meV.[366–374] However, this large spread is primarily a result of some calculations employing small supercells resulting in an unphysical description of the low-coverage situation, too small a k-point grid, or small atom-localized basis sets that do not adequately describe the binding and introduce large basis set superposition error (BSSE). In the present work, we use the DMC method to calculate the binding energy of H to graphene in the chemisorbed state.

4.3 Methods

All calculations reported in this study used a 5x5x1 supercell of graphene, as it was large enough to make inconsequential the interaction between periodic images of the adsorbed hydrogen atom and to assure that there are essentially unperturbed C atoms between the buckled regions in adjacent images in the x and y directions. The geometries of graphene, both pristine and with a chemisorbed H atom, were provided by Kim et al.,[375] and were obtained using the PBE+D3 DFT method.[364,376] For all systems, a vacuum spacing of 16 Å was used.

4.3.1 Density functional theory calculations

The single particle orbitals used in the trial wave functions for variational Monte Carlo (VMC) and DMC calculations were calculated using the PBE functional with the correlation consistent electron core potential (ccECP)[313,314] pseudopotentials and a plane wave basis with an energy cutoff of 3,400 eV. Monkhorst-Pack k-point grid meshes[377] were employed with a 13.6 meV Marzari-Vanderbilt-DeVita-Payne cold smearing of the occupations.[378] The PBE

results were converged at a 6x6x1 k-point grid to 1 meV for graphene and graphene with an adsorbed hydrogen atom. The hydrogen atom trial was generated using a 1x1x1 k-point grid. Convergence studies can be found in Table S1 and S2 of the Supplementary Material.

In addition to the PBE calculations used to generate the trial wave functions for DMC, DFT calculations were carried out with the PBE0[379] and Heyd–Scuseria–Ernzerhof (HSE) functionals[380] to determine if inclusion of exact exchange proves important for the adsorption energy. Due to the inclusion of exact exchange, these calculations would be computationally prohibitive in a plane wave basis, particularly with the high energetic cutoff required by the ccECP pseudopotential. For this reason, they were carried out all-electron with the POB-TZVP Gaussian type orbital (GTO) basis set.[381] Due to the use of GTOs, these calculations suffer from basis set superposition error (BSSE), which we corrected using Grimme's geometry-dependent counterpoise correction scheme.[382,383] This correction resulted in a 113 meV reduction in the magnitude of the binding energy when using the PBE0 functional. For the PBE0 and HSE calculations, a 12x12x1 k-point grid was used to assure binding energies converged to within 2meV. Convergence data are supplied in Table S3 of the Supplementary Material.

The plane wave DFT calculations were carried out with the QUANTUM ESPRESSO version 6.3 code.[384–386] The Gaussian basis DFT calculations were carried out with CRYSTAL17,[387,388] save for the HSE calculation of the isolated hydrogen atom which was carried out using NWChem version 6.8[389] using the same basis as the calculations in CRYSTAL17.

4.3.2 Quantum Monte Carlo calculations

DMC is a projector quantum Monte Carlo (QMC) method, solving the Schrödinger equation in imaginary time $\tau = it$; any initial state $|\psi\rangle$, that is not orthogonal to the true ground state $|\phi_0\rangle$, will evolve to the ground state in the long time limit. When dealing with Fermionic particles, the DMC method requires the use of the fixed-node approximation[390] to maintain the antisymmetric property of the wave function. For efficient sampling and to reduce statistical fluctuations, we use a Slater-Jastrow trial wave function fixing the nodes through a Slater determinant comprised of single-particle orbitals, which, in this work,

are expanded in a B-spline basis. The Jastrow factor is a function that reduces the variance by explicitly describing dynamic correlation. The Jastrow factor contains terms for one-body (electron-ion), two-body (electron-electron) and three-body (electron-electron-ion) interactions. The one- and two-body terms were described with spline functions,[391] while the three-body terms were represented by polynomials.[307] 10 parameters were used for the one-body terms per atom type, and 10 parameters were employed per spin-channel for the two-body terms. The cutoffs on the one- and two-body terms were fixed to the Wigner-Seitz radius of the simulation cell. The three-body terms were comprised of 26 parameters per term with a cutoff of 10 Bohr. The parameters in the Jastrow factor were separately optimized for each geometry with the linear method[392] using VMC. To reduce the cost of the DMC calculations as well as to reduce the fluctuations near the ionic core regions, ccECP pseudopotentials were used to replace the core electrons.[313,314] The ccECP pseudopotentials were designed to be used with high-accuracy many-body methods. The non-local effects due to the pseudopotentials were addressed using the determinant-localization approximation along with the t-moves method (DLTM).[393,394] Finite size effects were addressed using twist averaging.[395] The twist angles were chosen to be the symmetry unique points of the 6x6x1 k-point grid shifted by half a grid step away from the gamma point in each direction.

The DMC calculations were performed using the branching scheme proposed by Zen et al. (ZSGMA)[396] with a population control target of 8,192 walkers and a time step of 0.005 a.u., which represented a balance between computational cost and finite timestep error in previous work.[341]

We define the binding energy as

$$E_b = E_{dgr+H} - (E_{gr} + E_H) \tag{31}$$

where E_{dgr+H} is the energy of the distorted graphene sheet with a chemisorbed atomic hydrogen, E_H is the energy of a hydrogen atom, and E_{gr} is the energy of a pristine graphene sheet. In the chemisorbed state, the hydrogen atom bonds directly over a carbon atom, causing this carbon to be pulled out of the sheet towards the hydrogen.[397,398] The adjacent carbons are also pulled in the direction of the hydrogen leading to a distorted graphene sheet.

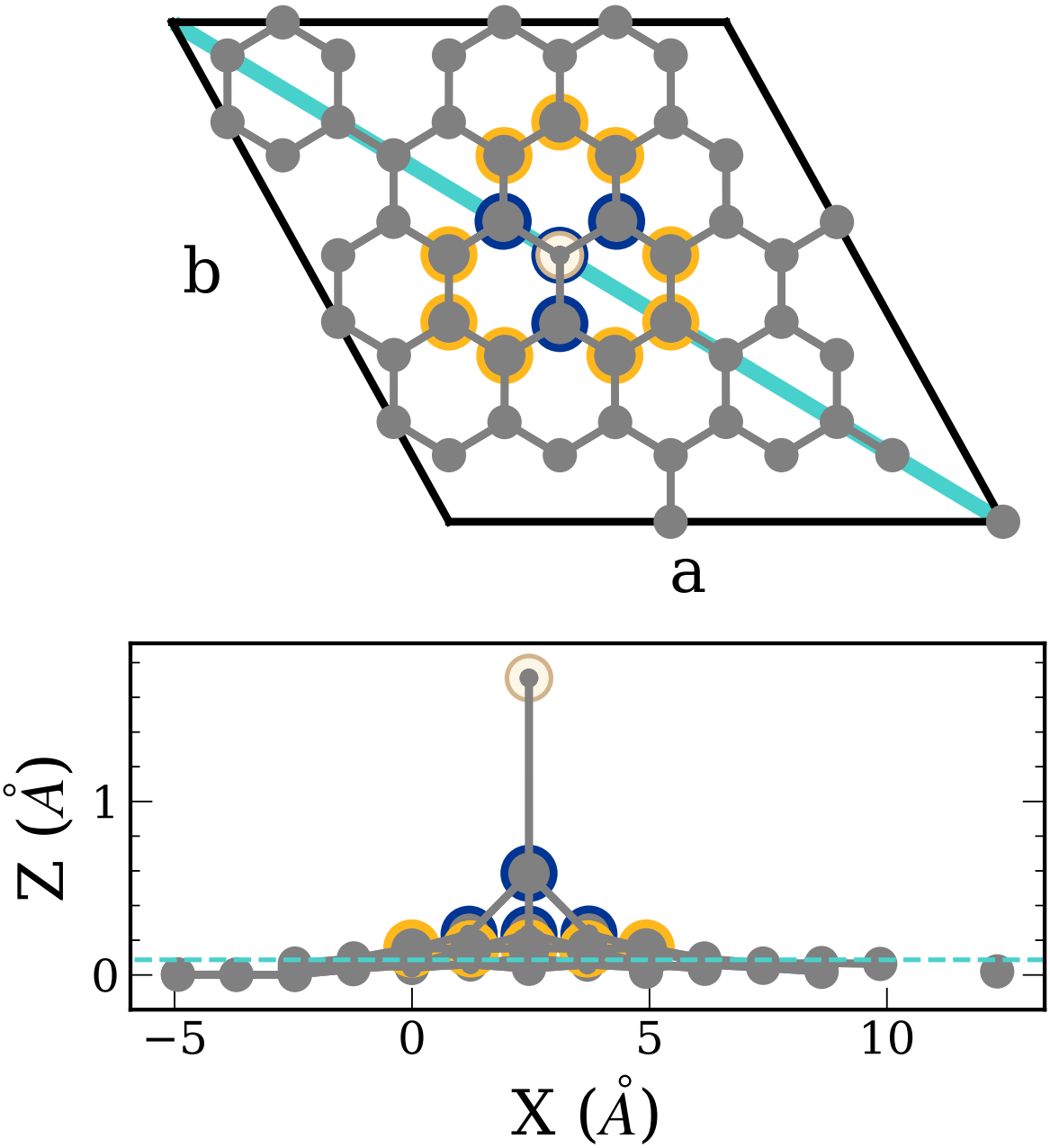


Figure 9: Perpendicular view of the simulation cell (top) and a parallel view obtained by projection onto the xz-plane (bottom). The carbon atoms are colored gray and the hydrogen atom is denoted as white. For the perpendicular view, the cyan line represents the slice of the cell used to visualized electron density differences. For the parallel view, the dotted cyan line represents the mean carbon z position. Blue outlined atoms are greater than one standard deviation away from the mean carbon z position, whereas yellow atoms are between 0.5-1.0 σ.

The QMC calculations were carried out using the QMCPACK code, with the workflow between QUANTUM ESPRESSO and QMCPACK managed by Nexus.[308,316,399] Figures 9 and 11 were rendered with matplotlib[331] and the density plots were generated using VESTA.[400]

4.4 Results & Discussion

4.4.1 Binding energy

Table 6: Binding energy (meV) of a hydrogen atom chemisorbed on graphene calculated with various DFT functionals and with DMC.

Method	Binding energy
This Work	
PBE	-821
PBE	-871
PBE0	-851 (-800)
HSE	-794 (-743)
DMC	-691 ± 19
Previous Work	
PW91	-810 to -830,[367] -870[368]
PBE	-790,[369] -840,[370] -980[371]

Calculation was done in the plane wave basis.

Calculation was done in the Gaussian basis set with corrections for BSSE. Values in parentheses include a correction for the basis set incompleteness as described in the text.

Table 6 contains a summary of the binding energies of a hydrogen atom chemisorbed on graphene from this work and selected values from previous publications using the PW91 and

PBE functionals. These literature values range from -790 to -980 meV. This wide spread of binding energies is caused by (1) the use in some studies of small supercells for which there are sizable interactions between the CH groups in adjacent cells, and (2) the use in some studies of small atom-centered basis sets without corrections for BSSE. Our calculations with the PBE functional in conjunction with a plane wave basis set give a binding energy of -821 meV. This should be contrasted with our -691 ± 19 meV DMC result. There are several possible sources for the difference between the PBE and DMC values of the binding energy. These include errors in the DFT calculations due to self interaction and planar graphene having more multiconfigurational character than H/graphene, with this being better described with DMC than with PBE. We note that the inclusion of the D3 dispersion correction with the PBE functional only changes the magnitude of the binding energy by 0.03 eV.

The PBE binding energy is 51 meV lower in magnitude in the plane wave than in the GTO basis set when the same k-point grid is used, and this value is used as a correction for the basis set incompleteness error for the results with other functionals in Table 6. The calculations in the GTO basis set give a slightly smaller in magnitude binding energy with PBE0 than with PBE. However, with HSE, we obtain a binding energy 77 meV smaller in magnitude than the PBE result. Applying the correction for the basis set incompleteness error, we obtain -800 meV for the PBE0 binding energy and -743 meV for the HSE binding energy, with the latter being in reasonable agreement with the DMC result of -691 meV. Although the 130 meV difference between the plane-wave PBE and DMC values of the binding energy may appear to be small, this energy difference, of that magnitude is consistent with an order of magnitude change in the hydrogen evolution current at room temperature on graphene electrodes.[375]

In order to better understand the origin of the difference in the PBE and HSE H-atom adsorption energies, we also carried out non-self-consistent calculations, using PBE densities to evaluate the HSE energies. These calculations gave a binding energy only 21 meV smaller in magnitude than obtained from the self-consistent HSE calculations. This demonstrates that the functional is more important than the density in establishing the binding energy. Detailed information can be found in Table S4 of the Supplementary Material.

Detailed results of the DMC calculations can be found in Tables S5-S7 in the Supple-

mentary Material.

4.4.2 Binding density

It is instructive to examine the change in the electron density associated with the binding of the H atom to the distorted graphene as determined from the PBE and DMC calculations. The density change is given by

$$\rho_b = \rho_{dgr+H} - (\rho_{dgr} + \rho_H), \tag{32}$$

where ρ_H is the charge density of the hydrogen atom, and ρ_{dgr+H} and ρ_{dgr} are the charge densities of the distorted graphene sheet with and without hydrogen, respectively. For the QMC density, the density was accumulated during the VMC and DMC calculations, the mixed estimator bias was found to be insignificant, and was thus not corrected.

The ρ_b density differences for both DMC and PBE are shown in Figure 10. The dark blue and gold regions represent a loss and gain of electron density, respectively. As expected, there is a shift in electron density from the carbon atom participating in the carbon-hydrogen bond as well as to the three adjacent carbon atoms. These qualitative changes in the density are consistent with previous theoretical and experimental studies.[397,398] The rehybridization from sp^2 to sp^3 of the carbon participating in the CH bond and the weakening of the π bonds due to the distortion of the graphene lead to the electron density shift. The change in the charge distribution is similar for PBE and DMC, with the most noticeable difference being a greater increase of density at remote C atoms in the DMC than in the PBE calculations.

4.4.3 Charge density differences between DMC and PBE

In this section, the difference between the DMC and PBE charge densities for distorted graphene with the adsorbed hydrogen atom as well as for pristine planar graphene without the adsorbed hydrogen atom are considered. The charge density difference for each system is calculated according to

$$\Delta\rho_{system} = \rho^{DMC}_{system} - \rho^{PBE}_{system}, \tag{33}$$

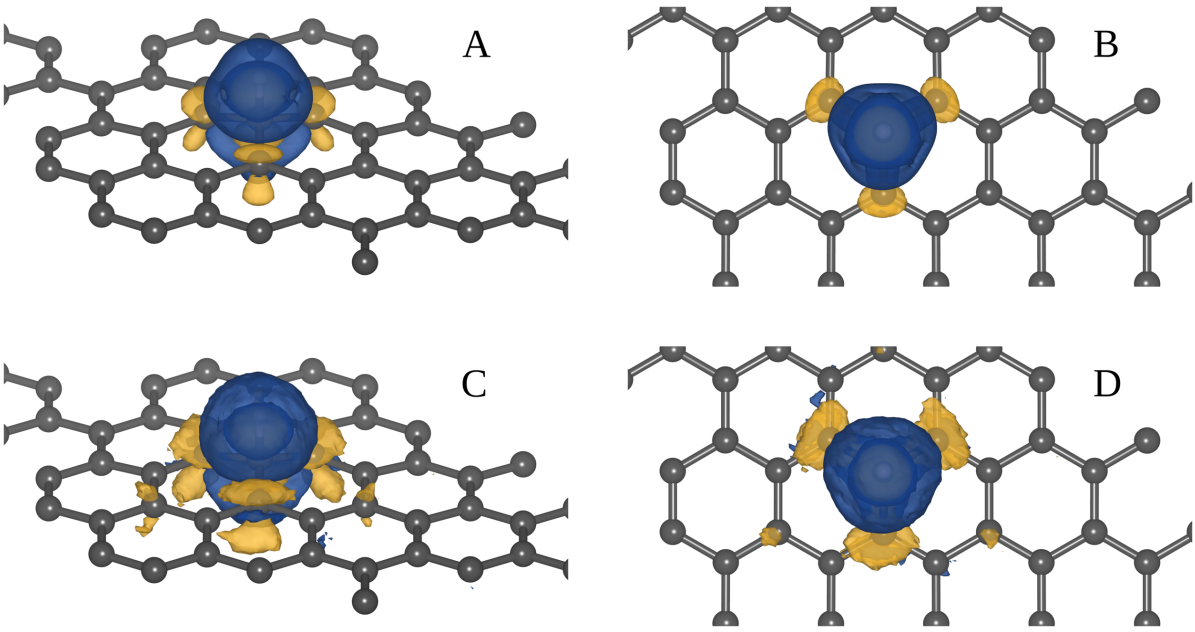


Figure 10: Change of the electron density due to the adsorption of the H atom to the distorted graphene sheet (Eq. 32). ρ_b from PBE calculations is shown from an oblique angle (A) and aligned along the c axis (B). ρ_b from DMC calculations (C) and (D) is shown from the same perspectives. Gold and blue represent a gain and loss of electron density, respectively. Note that there is a region of increased charge density at the C-H bond that is enveloped by a region of loss in the charge density. The binding density was visualized using an isovalue of 2.8×10^{-5} for DMC and 3.9×10^{-5} for PBE, in both cases capturing 95% of the differential charge density.

where ρ_{system}^{DMC} is the DMC charge density of a given system (either distorted graphene with the adsorbed hydrogen or pristine graphene) and ρ_{system}^{PBE} is the corresponding PBE charge density. $\Delta\rho_{gr}$ and $\Delta\rho_{dgr+H}$ are reported in Figure 11 along the 110 slice through the unit cell, which captures the carbon-hydrogen bond. From the top-down perspective in Figure 9, the 110 lattice plane bisects the cell diagonally through the longer of the two diagonals and is indicated by the solid cyan line. In Figure 11, blue represents areas where the PBE density is larger, while gold areas represent areas where the DMC density is larger. The DMC density, in comparison with the PBE density, has greater weight in the bonding region between atoms. We note that the HSE density displays similar differences as the PBE density. Figure S2 of the Supplementary Material includes a visualization of the DMC-HSE density difference. This is the case for both the planar graphene without hydrogen and the system with hydrogen chemisorbed to graphene. Even though there are significant differences between the PBE and DMC densities for both systems, the difference is similar in the two systems, consistent with it not introducing a large error in the PBE value of the binding energy.

4.5 Conclusions

Calculations of the binding energy of a hydrogen atom on a graphene sheet were carried out using various DFT methods and with DMC. The DMC calculations provide a benchmark value of the binding energy.

Our best estimate of the binding energy from DMC calculations is -691 ± 19 meV. The PBE result obtained with a plane-wave basis set gives a binding energy about 20% larger in magnitude than the DMC result. The global hybrid functional, PBE0, gives a binding energy close to that of PBE. In comparison, HSE, a range-separated hybrid functional, gives a smaller binding energy of -743 meV, after a correction applied for the basis set incompleteness error, and is much closer to the value from DMC calculations. Interestingly, there are significant differences in the DMC and PBE charge densities of both graphene and H/graphene.

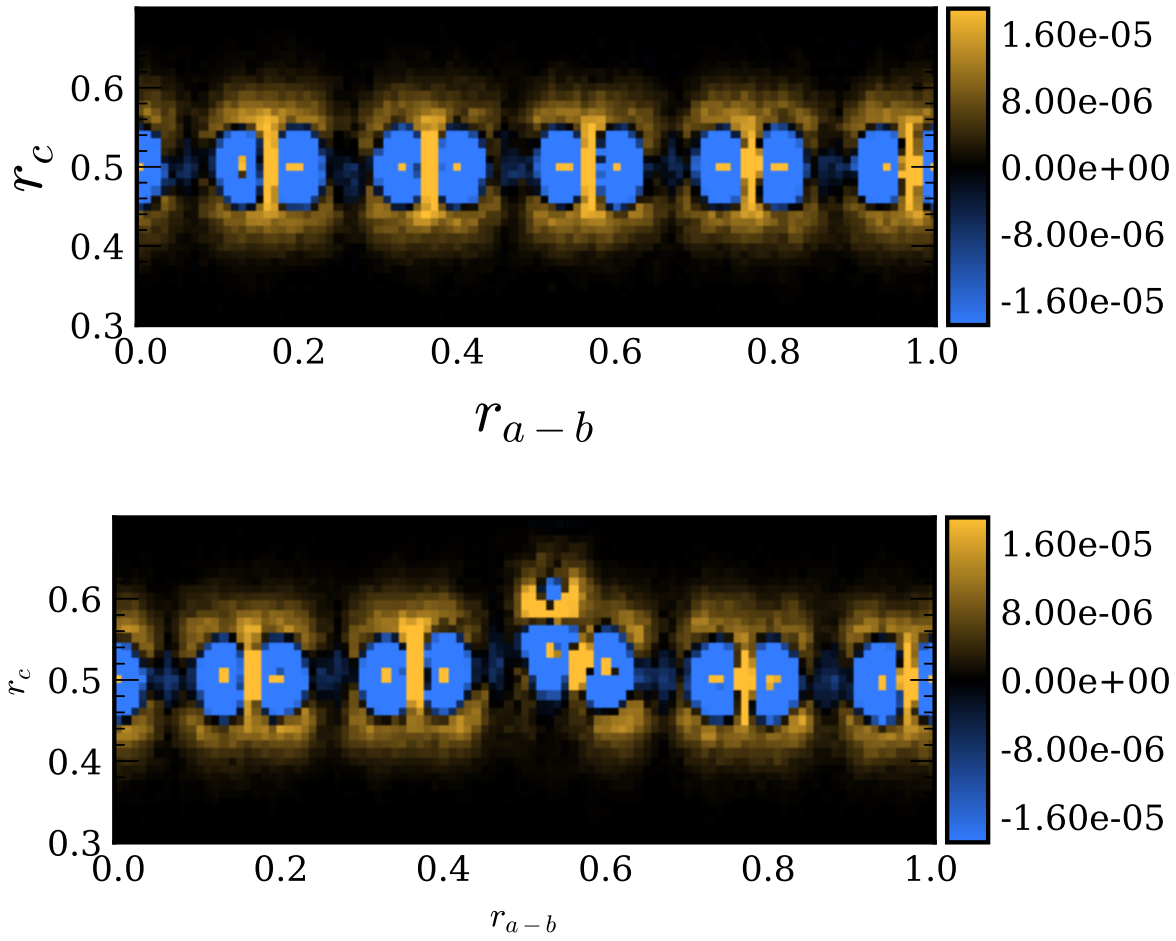


Figure 11: Visualization of the difference of PBE and DMC densities sliced along the 110 lattice plane of the unit cell for the graphene sheet, $\Delta\rho_{gr}$, (top) and H adsorbed onto graphene, $\Delta\rho_{dgr+H}$, (bottom). The abscissa represents traversing the 110 plane in fractional coordinates, while the ordinate represents traversing the c axis in fractional coordinates. Blue regions represent places where the PBE density is larger, while the gold color represents regions where the DMC density is larger.

4.6 Acknowledgements

We thank Dr. Dan Sorescu for helpful discussion and for sharing the coordinates of his calculations. A.B. and H.S. were 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. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. The DMC and plane wave DFT calculations used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under contract DE-AC02-06CH11357. The DFT calculation using Gaussian orbitals were carried out on computing resources in the University of Pittsburgh's Center for Research Computing. K.D.J. acknowledges NSF (CBET-2028826) for partial support of this work. S.U. was supported in part by the Pittsburgh Quantum Institute (PQI) Graduate Quantum Leader Award.

5.0 Capturing Correlation Effects for the Description of Positron Bound States

The diffusion Monte Carlo method is used to calculate the positron affinity of Be. The effect of electron-positron correlation on the nodal surface is explored. The best estimate of the positron affinity (PA) is in excellent agreement with the best estimate from literature.

5.1 Summary

The diffusion Monte Carlo (DMC) method is used to calculate the PA of Be. The effect of e^--e^+ correlation on the nodal surface is explored. Results from the literature are analyzed with respect to reference values for Be/e^+. The best estimate of the PA is in excellent agreement with the best estimate from literature.

5.2 Introduction

Positron chemistry can yield information about the electronic and vibrational properties of molecules.[†]Theoretical methods can yield insight into the nature of the interactions between matter and antimatter. Specifically, the properties of interest are the energy of a positron-matter bound state and the e^+-e^- annihilation rate. In this work, we will focus on the formation and energetics of the bound state.

Multicomponent methods are a powerful tool to describe matter-antimatter interactions. They are also conceptually similar to standard electronic structure methods. The origins of multicomponent methods can be traced to Thomas for the description of quantum nuclei,[82–85] and since then nuclear quantum effects have driven much of the development and application of multicomponent methods.[91–197] Soon after the first applications of multicomponent methods for nuclear quantum effects, a similar multicomponent method was used

to treat positron bound states.[86–90] Since then many positronic systems have been studied using multicomponent methods of varying accuracy and computational costs.[130,131,198–215]

The binding of a positron occurs in a diffuse nonvalence orbital, and are antimatter analogues of nonvalence anions.[†]In previous work on nonvalence anions, we have demonstrated the importance of e^--e^- correlation effects for such states.[278] Positron bound states are even more sensitive to e^--e^+ correlation effects as they have a favorable coulombic interaction and the lack of an exchange interaction between the positron and the electrons.

In this work, we aim to capture e^--e^+ correlation effects by developing accurate multicomponent wave function methods and then use these accurate correlated wave functions as trial wave functions for subsequent quantum Monte Carlo (QMC) calculations. A critical contribution of our approach is demonstrating the viability of a physically motivated configuration interaction ansatz, which incorporates the e^--e^+ dispersion interaction. This is motivated by a similar ansatz constructed for nonvalence correlation bound (NVCB) anions.[278]

5.3 Theory

5.3.1 Notation

Given the increased complexity of treating multiple types of interacting quantum particles, the notation for the following equations is first introduced. Generally, tensor quantities will be manipulated with the following labels,

$$ {}^{\gamma}_{\xi} A^{q}_{\mu\nu\cdots}, \tag{34} $$

where ξ is the species index, γ is the spin index, $\mu\nu\cdots$ are the tensor indices, and q are the qualifiers (e.g. "core" to distinguish a core contribution from the full tensor).

5.3.2 Mean Field

Similar to standard electronic structure calculations, multicompoonent calculations begin from a mean field description akin to the Hartree-Fock method for electrons. The wave function ansatz used is a product ansatz,

$$\psi_{\mathrm{MF}} = \prod_{\xi} {}^{\alpha}_{\xi}D \; {}^{\beta}_{\xi}D, \tag{35}$$

where the product runs over the number of quantum particle types and ${}^{\alpha}_{\xi}D$ and ${}^{\beta}_{\xi}D$ are the α and β determinants of particle ξ written in the Waller-Hartree double determinant representation. Note that if only a single positron is present this reduces to the electronic Slater determinant times a positronic orbital, and so this ansatz is a generalization of the representation used in earlier work.[86–90] Unlike previous multicomponent methods, we do not restrict ourselves to two interacting quantum particle types.

The use of this ansatz in a Hartree-Fock type framework results in a set of coupled Fock operators, which are solved self consistently. The elements of the Fock operator in an atomic orbital basis are given,

$${}^{\gamma}_{\xi}F_{\mu\nu} = {}_{\xi}H_{\mu\nu} + \sum_{\zeta \geq \xi} \sum_{\lambda,\sigma} {}_{\zeta}q \; {}_{\xi}q \left[\left[{}^{\gamma}_{\zeta}D_{\lambda\sigma} + {}^{\gamma'}_{\zeta}D_{\lambda\sigma} \right] \left({}_{\xi}\mu\nu \mid {}_{\zeta}\lambda\sigma \right) - \delta_{\xi\zeta} \left({}_{\xi}\mu\lambda \mid {}_{\zeta}\nu\sigma \right) \right], \tag{36}$$

where $\gamma \in \{ \alpha, \beta \}$ and $\gamma' \in \{ \alpha, \beta \} - \{ \gamma \}$ and $\mu\nu\lambda\sigma$ are a.o. indices.

If a particle is spin-restricted then only a single Fock operator is constructed, but the normalization of the density matrix is kept the same as the unrestricted case. For example, the restricted Fock operator for an electron-only calculation is given as,

$${}_{e^-}F_{\mu\nu} = {}_{e^-}H_{\mu\nu} + \sum_{\lambda,\sigma} \left[\left[{}_{e^-}D_{\lambda\sigma} + {}_{e^-}D_{\lambda\sigma} \right] \left(\mu\nu \mid \lambda\sigma \right) - \left(\mu\lambda \mid \nu\sigma \right) \right]. \tag{37}$$

Each quantum particle contributes to the total energy,

$${}_{\xi}E = \frac{1}{2} \left[\left({}^{\gamma}_{\xi}D + {}^{\gamma'}_{\xi}D \right) {}_{\xi}H + {}^{\gamma}_{\xi}D \; {}^{\gamma}_{\xi}F + {}^{\gamma'}_{\xi}D \; {}^{\gamma'}_{\xi}F \right], \tag{38}$$

with the total energy given as the sum of the contributions plus the nuclear repulsion energy,

$$E_T = \sum_{\xi} {}_{\xi}E + E_{\mathrm{nuc}}. \tag{39}$$

5.3.2.1 Convergence of the Self-Consistent Process

A standard multicomponent approach converges the density in a step-wise fashion, i.e. for an electron-positron calculation, converging the electron density followed by converging the positron density and repeating this cycle until overall self-consistency is achieved.[401] Recent work has found that iterating the densities simultaneously until self-consistency is achieved is more efficient.[401]

In our work, we iterate the densities simultaneously, but with an important distinction, we do not turn on interactions between quantum particle types until the non-interacting densities are converged. This has an obvious computational advantage as less terms need to be evaluated for the early Fock builds. However, this also has the advantage of more likely finding a Hartree-Fock minimum for problematic cases. If there are several low lying orbitals, HF may converge to a solution with an incorrect orbital occupied. In the multicompoent HF case, if the densities are iterated without first converging the independent densities, one is less likely to occupy the correct orbital finding a HF solution that is not a global minimum.

5.3.2.2 Implementation Details

The specific implementation chosen here is a direct SCF algorithm. The direct SCF algorithm is extremely efficient when combined with incremental Fock formation plus integral-based and density-based screening using the Cauchy-Schwarz inequality. Incremental Fock formation has been implemented, but given the highly coupled nature of multicomponent calculations, we have found screening must be handled with care. In further studies, we plan to implement density based screening to achieve a linear scaling Fock build algorithm. DIIS extrapolation is used to accelerate convergence. Symmetric orthogonalization and canonical orthogonalization are implemented with a preference for canonical orthogonalization as it handles linear dependencies in the a.o. basis.

5.3.3 Configuration Interaction

The multicomponent configuration interaction ansatz is formed from a linear combination of products of determinants,

$$\psi_{\mathrm{CI}} = \sum_i C_i \left[\prod_{\xi} {}^{\alpha}_{\xi}D \; {}^{\beta}_{\xi}D \right], \tag{40}$$

similar to the standard electronic structure CI method. The CI expansions used in this work are formed from excitations from a mean field solution. The excitation operators ${}^{\alpha}_{\xi}T^{a}_{i}$ and ${}^{\beta}_{\xi}T^{b}_{j}$ perform an excitation for quantum particle type xi and spin $\alpha(\beta)$ from hole $i(j)$ to particle $a(b)$. To make things explicit, we define our excitations in the following way

1. A maximum overall excitation level is specified.
2. For each particle type, a maximum excitation level is adopted for each spin along with a combined maximum excitation level.

This allows for highly flexible configuration interaction calculations. This is crucial for correlation bound positronic states as they are similar to nonvalence correlation bound anions, and previous work has shown that a flexible compact ansatz including only the doubles involving the excess electron to generate an excellent trial wave function for QMC. The analogous CI in the multicomponent case involves two quantum particle types, electrons and a positron, with single excitations and only the doubles with the positron, which is possible in this framework.

5.3.3.1 Diagonalization to solve the CI eigenvalue problem

The CI eigenvalue problem,

$$H_{\mathrm{CI}}C = EC \tag{41}$$

is solved using the Davidson diagonalization method.[402] For small CI spaces, the Hamiltonian matrix can be explicitly constructed, but for larger spaces a determinant-driven direct CI method is used.

The advantage of a direct CI approach is that the Hamiltonian matrix is never explicitly constructed. This represents a significant savings in memory usage and enables large CI

calculations for systems where the CI matrix would not fit in memory. Central to this method is the formation of the matrix-vector product in which the Hamiltonian is contracted with a trial vector,

$$\sigma = Hc. \tag{42}$$

In standard electronic structure theory, the formation of the σ vector is a well studied problem and many efficient algorithms have been developed.

The sigma vector for one quantum particle type, A, involves 3 terms,

$$\sigma = \sigma^{A\alpha A\alpha} + \sigma^{A\beta A\beta} + \sigma^{A\alpha A\beta}, \tag{43}$$

and similarly for two quantum particle types, A and B,

$$\sigma = \sigma^{A\alpha A\alpha}+\sigma^{B\alpha B\alpha}+\sigma^{A\beta A\beta}+\sigma^{B\beta B\beta}+\sigma^{A\alpha A\beta}+\sigma^{B\alpha B\beta}+\sigma^{A\alpha B\alpha}+\sigma^{A\beta B\beta}+\sigma^{A\alpha B\beta}+\sigma^{A\beta B\alpha}. \tag{44}$$

The general number of terms in a σ vector formation for a N quantum particle type system is

$$\binom{2N-1}{2}. \tag{45}$$

The single quantum particle sigma formations require different algorithms for the same spin ($\sigma^{A\alpha A\alpha}$ and $\sigma^{A\beta A\beta}$) and opposite spin ($\sigma^{A\alpha A\beta}$) cases. Similarly for the two particle sigma formation, there are two classes of contributions:

1. same particle and same spin contributions ($\sigma^{A\alpha A\alpha}$, $\sigma^{A\beta A\beta}$, $\sigma^{B\alpha B\alpha}$, $\sigma^{B\beta B\beta}$)
2. different particle or different spin contributions ($\sigma^{A\alpha A\beta}$, $\sigma^{B\alpha B\beta}$, $\sigma^{A\alpha B\alpha}$, $\sigma^{A\beta B\beta}$, $\sigma^{A\alpha B\beta}$, $\sigma^{A\beta B\alpha}$)

In the current iteration of the code, the matrix vector product is formed in a single step that scales as $\mathcal{O}(N_{\rm det}{}^{\gamma}_{\xi} N {}^{\gamma'}_{\xi} N$, where $N_{mathrmdet}$ is the total number of determinants and the other factors of N are the number of spin determinants for the particle spin pair.

Additionally, natural spin orbitals are constructed from the one particle reduced density matrices from the CI expansions in a state specific way. These natural orbitals can also be used as a single determinant trial wave function for subsequent QMC calculations.

CI calculations in a determinantal basis are not guaranteed to produce states that are spin eigenfunctions unlike CI calculations in a configuration state function basis. The standard solution to this is to implement a spin purification scheme in either a first or second order form. Both approaches are implemented to enforce spin symmetry for problematic cases.

5.3.3.2 Implementation details

The storage and processing of determinants is central to the configuration interaction code. We represent determinants using unsigned 64 bit integers for each particle and spin type. The integers are zero padded from the left until the entire 64 bit int is filled. These integers are stored in a list to represent a single spin determinant, ${}^{\gamma}_{\xi}D$. A list of unique determinants is stored for each particle type and spin type. The complete variational space is then stored as an unsorted map between a tuple of indices into the unique spin determinant lists and the index of the determinant made of those spin determinants in the variational space.

To speed up the Hamiltonian construction or sigma vector formation, we store a list of unique single and double excitations within each unique particle and spin type. For example, for a spin determinant ${}^{\gamma}_{\xi}D_i$, we store all connected singles in a two dimensional list in row i and all connected double excitations in another two dimensional list in row i. The storage of double excitations is normally not done as the storage requirement grows quickly. However for the same number of total determinants, the unique particle and spin determinant lists are shorter in a multicomponent calculation than an electron only calculation by nature of the product ansatz so we are able to accept the storage cost for the efficiency gain.

5.4 Results & Discussion

5.4.1 Be

We first consider a positron binding to a Be atom. Table 7 contains the positron affinity for Be calculated with several methods. Also tabulated are results from the literature using

Table 7: PAs of Be calculated using various methods. The methods and basis used as reported as "without e^+ method or basis // with e^+ method or basis". We also report if the $s^2 \rightarrow p^2$ near degeneracy is handled in the trial wave function.

method	PA (meV)	$s^2 \rightarrow p^2$
HF // HF [1]	-1.9	no
HF // rSDCI [1]	152.3	no
FCI // FCI [1]	10.5	yes
DMC[403]	100 ± 5	no
DMC[403]	33 ± 11	yes
SVM[404]	86	yes
DMC[212]	34 ± 8	yes
DMC[212]	54 ± 10	yes

[1] These calculations used the aug-cc-pVQZ basis set for the electrons and an 11s8p6d6f3g basis for the positron.

the DMC and SVM methods. As expected, the HF treatment fails to bind the positron. The restricted single double configuration interaction (rSDCI) approach, which includes all single excitations and only the double excitations that involve the positron, greatly increases the binding energy, in fact giving a value about 60% too large in magnitude. In the FCI results, the binding energy drops significantly. Since this is calculated using a Gaussian basis, we expect that this is still not fully resolving the electron-positron cusp.

The DMC results for the positron affinity of Be is presented in table 8. The positron affinities are presented for the zero time step extrapolated DMC energies with various trial wave functions. Using the Jastrow factor with a multi-Slater determinant trial function explicitly includes the electron-positron cusp. The three single determinant references yield similar PA. The single determinant energy versus time step for Be with a positron is presented in figure 12. For each of the single determinant wave functions, the extrapolated DMC energies are within error bars of each other. The HF nodal surface is expected to be the poorest as the positron is unbound in that case, which manifests as it has the largest time step dependence and the largest error bars in the same projection time. These results also show that the $s^2 \rightarrow p^2$ near degeneracy cannot be captured by a single Slater determinant as the FCI wave function captures this effect, but the natural orbitals from the FCI wave function do not.

The DMC results from the literature, tabulated in table 7, show the same marked decrease in the PA by including the $s^2 \rightarrow p^2$ excitations. With a more sophisticated treatment of electron-positron correlation, the PA nearly doubles as seen in table 7.[212] This indicates a strong impact of the correlation on the nodal surface.

To elucidate the importance of the nodal surface, we plot the DMC energy of the Be/e^+ system in figure 13. The energy of the DMC calculation with a HF trial wave function lies significantly far above the correlated trial wave functions. Including the dispersion effects with the rSDCI multi-determinant trial wave function, improves the DMC energy, but clearly a significant change in the nodal surface happens once the $s^2 \rightarrow p^2$ excitations are included. Including those excitations by using the FCI trial wave function, the DMC energy drops into excellent agreement with the SVM reference. The results from previous DMC studies lie higher in energy are included in the plot for reference.[212,403]

Table 8: PAs of Be calculated using the DMC method and various trial wave functions. The trial wave functions used are reported as "single or multi- determinant/ without e^+ wfn // SD or MD / with e^+ wfn".

trial wave function	PA (meV)
SD/HF//SD/HF	96 ± 5
SD/HF//SD/NO rSDCI	104 ± 4
SD/FCI NO//SD/FCI NO	95 ± 4
SD/HF//MD/rSDCI	151 ± 4
MD/FCI //MD/FCI	91 ± 4

We additionally want to comment on a topic with wider implications for the QMC field, time step extrapolations for energy differences. In the literature it has been found that for exitonic binding energies, one can use a very large time step and still resolve the correct energy difference even if the individual energies have an associated time step error.[405] In figure 14, we find that for diffuse states such as positron bound states, the time step dependence is still severe. For a sufficiently small time step however, one is within error bars without extrapolation. This leads us to recommend caution when resolving energy differences for systems with large relevant length scales, the diffuse positron in this case.

5.5 Conclusions

In this work, we exhaustively explore the binding of a positron to Be. We also compare our results to those from the literature. Our calculations with the most sophisticated trial wave functions give a positron binding energy for Be of 91±4 meV. This is in excellent agreement with the previous theoretical best estimate. Applications of the DMC method to Be clusters and molecules is underway.

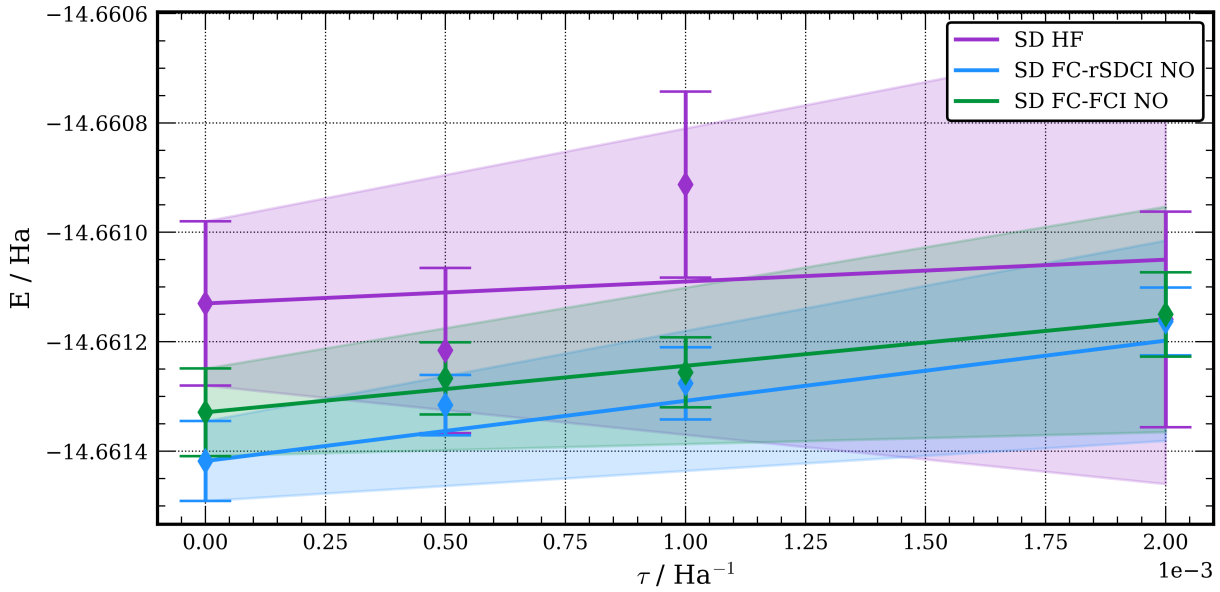


Figure 12: The DMC energy versus time step for Be with a bound positron using a trial wave function of a single determinant of Hartree-Fock orbitals, frozen core rSDCI natural orbitals, and frozen core FCI natural orbitals. Note at the largest timestep the DMC energies are overlapping. The shading represents the uncertainty in the linear fit to the DMC data. All plots generated using Matplotlib.[331]

.

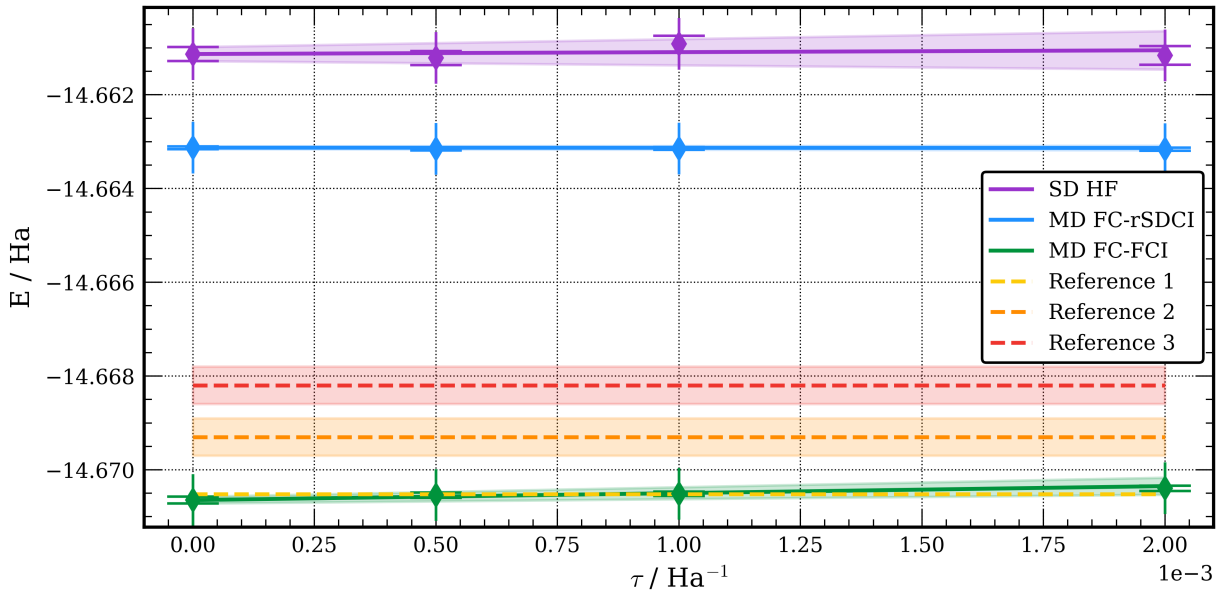


Figure 13: The DMC energy versus time step for Be with a bound positron with a Hartree-Fock single determinant trial wave function, frozen core rSDCI multi-determinant trial wave function, and frozen core FCI multi-determinant trial wave function. The reference energies are from (reference 1) SVM,[404] (reference 2) DMC with an explicitly correlated trial wave function for the positron,[212] and (reference 3) DMC with a positronic orbital obtained from a model potential calculation.[403] Reference values 2 and 3 are extrapolated to the zero time step limit and the shading represents the error bar on the extrapolated energy.

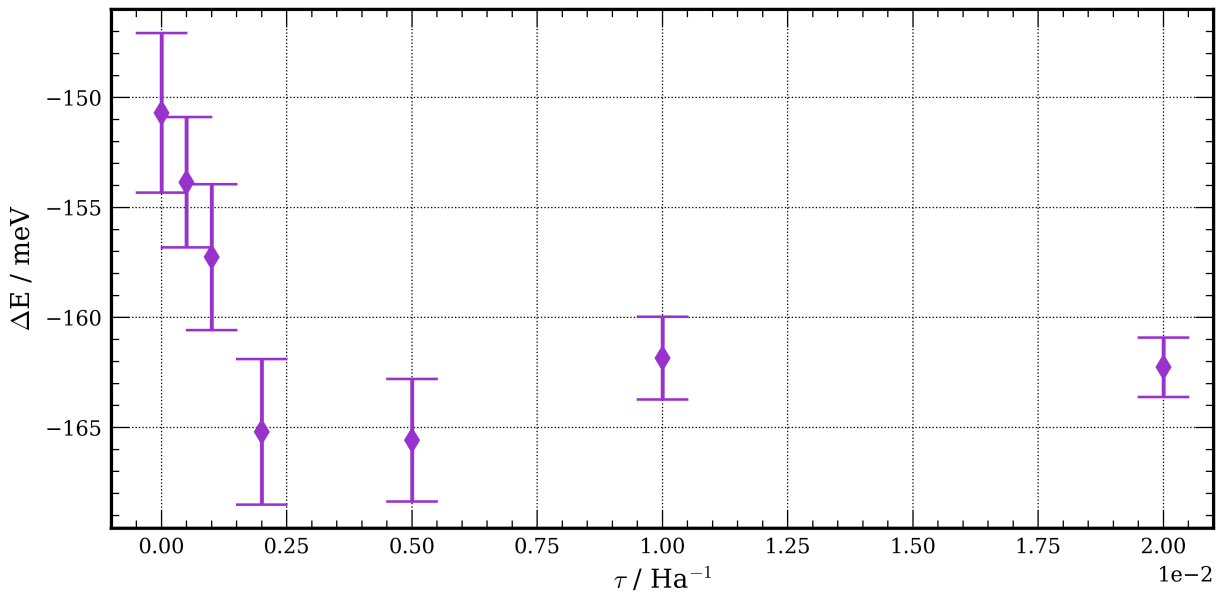


Figure 14: DMC positron affinity for Be calculated using an rSDCI multi-determinant trial wave function at several time steps sizes.

6.0 Conclusions

In this work, the importance of correlation effects in the binding of Fermions was explored. The first part of this work involved the study of e^--e^- correlation effects in the diffuse anionic states of nonvalence anions. First, the binding of an electron in an dipole field was studied using quantum Monte Carlo methods. It was found that quantum Monte Carlo methods are viable theoretical methods for the description of dipole bound anions. This motivated the study of a model $(H_2O)_4$ nonvalence correlation bound anion system. This model system exhibits a crossover between a correlation-bound anion and electrostatically-bound anion by tuning a single geometric parameter. In this model system, QMC methods were shown to provide quantitavely accurate electron binding energies in both the electrostatically-bound region and the correlation-bound region. In this work, a compact correlated ansatz was developed that yielded excellent nodal surfaces for QMC calculations. This work was then extended to positron binding to atoms and molecules. In a similar manner to the nonvalence correlation bound anion states, we developed a physically motivated CI ansatz for trial wavefunctions. This improved the nodal surfaces for subsequent DMC calculations. In the future, we envision that QMC will become the method of choice for the description of positron bound states. The favorable scaling with respect to the number of electrons and the ability to capture electron-positron correlation accurate will motivate this choice. The lessons learned and tools developed in the current work will serve as an important step in this direction. The chemisorption of a hydrogen atom on a graphene sheet was studied. This is a relatively weak interaction, and so the accurate description of electron correlation was crucial in this study.

Appendix Q-GPU: A Recipe of Optimizations for Quantum Circuit Simulation Using GPUs

The test and figures in this chapter have been adapted from Zhao, Y.; Guo, Y.; Yao, Y.; Dumi, A.; Mulvey, D. M.; Upadhyay, S.; Zhang, Y.; Jordan, K. D.; Yang, J.; Tang, X. In *2022 IEEE Int. Symp. High-Perform. Comput. Archit. HPCA*, 2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA), 2022, pp 726–740, DOI: 10.1109/HPCA53966.2022.00059 with the permission of IEEE. The author's contribution to the work is the preparation of the chemistry example and assisting in the benchmarking. Additionally the author assisted in the preparation of the manuscript.

A.1 Summary

In recent years, quantum computing has undergone significant developments and has established its supremacy in many application domains. Unfortunately, modern quantum computing is still positioned in the Noisy Intermediate-Scale Quantum (NISQ) era that is limited by the number of qubits, short qubit lifetime, and imperfect operations. While quantum hardware is accessible to the public through the cloud environment, a robust and efficient quantum circuit simulator is necessary to investigate the constraints and foster quantum computer development, such as quantum algorithm development and quantum device architecture exploration. In this paper, we observe that most of the publicly available quantum circuit simulators (e.g., QISKit from IBM) are not optimized and suffer from slow simulation and poor scalability. To this end, we systematically studied the deficiencies in modern quantum simulators and propose Q-GPU, a framework that leverages GPUs with comprehensive optimizations to allow efficient and scalable quantum circuit simulation (QCS). Specifically, Q-GPU features i) proactive state amplitude transfer, ii) zero state amplitudes pruning, iii) delayed qubit involvement, and iv) non-zero state compression. Experimental results across eight representative quantum circuits indicate that Q-GPU significantly improves the

simulation performance over the state-of-the-art GPU-based QCS by 2.53× on average. It also outperforms the most recent OpenMP CPU implementation, the Google Qsim-Cirq simulator, and the Microsoft QDK simulator.

A.2 Introduction

Quantum computing is a promising computing paradigm that has the potential to solve problems that cannot be handled by classical computers in a feasible amount of time.[407] In the past decade, there has been steady progress towards building a large quantum computer. The number of qubits in a real quantum machine has increased from 14 in 2011[408] to 76 in 2020.[409] IBM promises 1000 qubits quantum machine by the year 2023.[410] Despite this rapid progress, current quantum computing is still positioned in the Noisy Intermediate-Scale Quantum (NISQ) era where the public has very limited access to quantum machines. These machines are also constrained by the limited number of qubits, short lifetimes of qubits, and imperfect operations.[411] Thus, quantum circuit simulation (QCS) toolsets provide an essential platform to satisfy many needs, e.g., developing many different algorithms with a large number of qubits, validating and evaluating newly proposed quantum circuits, and design space exploration of future quantum machine architectures. Many companies, such as IBM, Google, Intel, and Microsoft have developed their quantum circuit simulators to provide precise end-end simulation.

In general, QCS is challenging as it is both compute-intensive and memory-intensive.[412,413] The reasons are: i) fully and accurately tracking the evolution of quantum system through classical simulation[414] requires storing all the quantum state amplitudes, which carries a memory cost that grows exponentially as the number of qubits in the simulated quantum circuit increases, and ii) applying a gate within a quantum circuit requires a traversal of all the stored state amplitudes, leading to exponentially scaling computational complexity. Modern GPUs have been used to fuel QCS in high-performance computing (HPC) platforms. Specifically, when applying a gate to a n-qubit quantum circuit, the 2^n state amplitudes are evenly divided into groups, and each group of amplitudes is updated independently in par-

allel by GPU threads. However, the promising parallelism of GPUs is diminished by the limited GPU on-board memory capacity. For example, simulating a quantum circuit with 34 qubits requires 256 GB of memory to store state amplitudes, which is beyond the memory capacity of any modern GPUs.

There exist several works optimizing QCS, including multi-GPU supported simulation,[415,416] OpenMP and MPI based CPU simulation,[417–419] and CPU-GPU collaborative simulation.[420] Most of these works focus on distributed simulation while failing to benefit from GPU execution due to the memory constraint. In particular, our characterization shows that the state-of-the-art GPU-based simulation[420] has low GPU utilization when the number of qubits in the quantum circuit is large. As a result, most state amplitudes are stored and updated on the CPU, failing to take advantage of the GPU parallelization. Moreover, the static and unbalanced allocation of state amplitudes introduces frequent amplitude exchange between CPU and GPU, which introduces additional data movement and synchronization overheads.

In this paper, we aim to provide a high-performance and scalable QCS using GPUs. We propose *Q-GPU*, a framework that significantly enhances the simulation performance for practical quantum circuits. The proposed framework leverages modern GPUs as the main execution engine and is featured with several end-to-end optimizations to fully take advantage of the rich computational parallelism on GPUs, while maintaining a minimum amount of data movement between the CPU and GPU. Specifically, our approach includes four optimizations. First, instead of statically assigning state amplitudes on GPU and CPU as done in prior works,[420] Q-GPU dynamically allocates groups of state amplitudes on the GPU and proactively exchanges the state amplitudes between CPU and GPU. Doing so maximizes the overlap of data transfer between CPU and GPU, thereby reducing the GPU idleness. Second, Q-GPU prunes zero state amplitudes to avoid unnecessary data movement between CPU and GPU. Third, we also propose compiler-assisted quantum gate reordering (complying with the gate dependencies) to enlarge the opportunity of pruning zero state amplitudes. Finally, we propose efficient GPU-supported lossless data compression to further reduce data transfer caused by non-zero amplitudes. This paper makes the following contributions:

- We use the popular IBM QISKit-Aer with its state-of-the-art CPU-GPU implementation,[421] and conduct an in-depth characterization of the simulation performance. We observe that the performance degrades significantly as the number of qubits increases due to the unbalanced amplitudes assignment, where most of the computation is done by the CPU.
- We implement a dynamic state amplitude assignment to allow the GPU to update all state amplitudes. However, such an implementation did not provide any performance improvements and even worsened compared to the CPU execution due to the massive and expensive data movement between CPU and GPU.
- We propose Q-GPU, a framework comprising end-to-end optimizations to mitigate the data movement overheads and unleash the CPU capability in QCS. Specifically, the proposed Q-GPU is featured with the following major optimizations: i) dynamic state amplitudes allocation and proactive data exchange between CPU and GPU, ii) dynamic zero state amplitude "pruning", iii) dependency-aware quantum gate reordering to enlarge the potential of zero amplitude pruning, and iv) GPU-supported efficient lossless compression for non-zero amplitudes.
- We evaluate the proposed Q-GPU framework using eight practical quantum circuits. Experimental results indicate that in all circuits tested, Q-GPU significantly improves the QCS performance and outperforms the baseline by 2.53× on average. We also compare Q-GPU with Google Qsim-Cirq[422] and Microsoft QDK,[423] and results show that Q-GPU approach outperforms Qsim-Cirq and QDK by 1.02× and 9.82×, respectively.

A.3 Background

A.3.1 Quantum Basics

Similar to the *bit* concept in classical computation, quantum computation is built upon the *quantum bit* or *qubit* for short.[424] A qubit is a two-level quantum system defined by two computational orthonormal basis states $|0\rangle$ and $|1\rangle$. A quantum state $|\psi\rangle$ can be expressed

by any linear combination of the basis states.

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle, \tag{46}$$

where a_0 and a_1 are complex numbers whose squares represent the probability amplitudes of basis states $|0\rangle$ and $|1\rangle$, respectively. Note that we have $|a_0|^2 + |a_1|^2 = 1$, meaning that after measurement, the read out of state $|\psi\rangle$ is either $|0\rangle$ or $|1\rangle$, with probabilities $|a_0|^2$ and $|a_1|^2$, respectively. The states of a quantum system are generally represented by *state vectors* as

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{47}$$

To be more general, for an n-qubit system, there are 2^n state amplitudes. Then, the quantum state $|\psi\rangle$ can be expressed as a linear combination

$$|\psi\rangle = a_{0\ldots00}|0\ldots00\rangle + a_{0\ldots01}|0\ldots01\rangle + \cdots + a_{1\ldots11}|1\ldots11\rangle. \tag{48}$$

Similarly, the state of a n-qubit system can also be represented by a state vector with 2^n dimensions as

$$|\psi\rangle = a_{0\ldots00}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_{0\ldots01}\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + a_{1\ldots11}\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} a_{0\ldots00} \\ a_{0\ldots01} \\ \vdots \\ a_{1\ldots11} \end{bmatrix}. \tag{49}$$

Quantum computation describes changes occurring in this state vector. A quantum computer is built upon a *quantum circuit* containing *quantum gates* (or quantum operations), and a quantum algorithm is described by a specific quantum circuit. In simple terms, quantum gates are represented by unitary operations that are applied on qubits to map one quantum state to another. A quantum gate that acts on k qubits is represented by a $2^k \times 2^k$ unitary matrix.

To illustrate how a quantum gate is applied to a state vector, let us consider a 2-qubit system with a Hadamard gate/operation operating on qubit 0. A Hadamard gate can be represented as

$$H \equiv \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{50}$$

Then the state vector of this 2-qubit system is updated through

$$\begin{bmatrix} a'_{00} \\ a'_{01} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a_{00} \\ a_{01} \end{bmatrix}, \tag{51}$$

$$\begin{bmatrix} a'_{10} \\ a'_{11} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a_{10} \\ a_{11} \end{bmatrix}. \tag{52}$$

For an n-qubit system, when a H gate is applied to qubit j the amplitudes are transformed as:[425]

$$\begin{bmatrix} a'_{\times\cdots\times 0_j \times\cdots\times} \\ a'_{\times\cdots\times 1_j \times\cdots\times} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a_{\times\cdots\times 0_j \times\cdots\times} \\ a_{\times\cdots\times 1_j \times\cdots\times} \end{bmatrix} \tag{53}$$

Therefore, the indices of every pair of amplitudes have either 0 or 1 in the jth bit, while all other bits remain the same[1]. Note that each pair of amplitudes can be updated in parallel.

A.3.2 Quantum Circuit Simulation (QCS)

The purpose of QCS is to mimic the dynamics of a quantum system,[414] and to reproduce the outcomes of a quantum circuit with high accuracy. There are several approaches to simulating a quantum circuit, each offering different advantages and drawbacks. We summarize the three most widely used approaches below.

- **Schrödinger style simulation:** Schrödinger simulation describes the evolution of a quantum system by tracking its quantum state. It tracks the transformations of the state vector according to Equation 53. Note that one can also track the density matrix $\rho = |\psi\rangle\langle\psi|$, which is useful when measurement is required during simulation.[414,415] In this work, we only consider quantum measurements at the end of circuits.
- **Stabilizer formalism:** Simulation based on the stabilizer formalism is efficient for a restricted class of quantum circuits.[414,424,426] Specifically, stabilizer circuits (a.k.a Clifford circuits) can be simulated in $O(poly(n))$ space and time costs. Rather than tracking the state vector, the quantum state is uniquely represented and tracked by its stabilizers,

[1] "$\times$" can be 0 or 1; the "$\times$" in the same position of $a_{\times\times\times\times\times\times\times 0}$ and $a_{\times\times\times\times\times\times\times 1}$ are the same.

which is essentially a group of operators derived from the Clifford group. A detailed description can be found in.[426]

- **Tensor network:** Tensor network simulators are useful when a single or few amplitudes of the full state vector are being updated as tensor networks.[427–430] For example, one type of tensor network that are extremely common are matrix product states (MPS). When applied to a single amplitude in Equation 48, the resulting state resembles a long string of matrix multiplications

$$|\psi\rangle = \sum_{j_0 \cdots j_{n-1} j_n} a_{j_0 \cdots j_{n-1} j_n} |j_0 \ldots j_{n-1} j_n\rangle \tag{54}$$

$$= \sum_{j_0 \cdots j_{n-1} j_n} Tr[A^{j_0} \ldots A^{j_{n-1}} A^{j_n}] |j_0 \ldots j_{n-1} j_n\rangle \tag{55}$$

The matrices A (rank-2 tensors) in Equation 55 can be thought of as a decomposition of the full coefficient tensor a. Despite the restriction of returning a limited number of amplitudes, tensor networks states are efficient as they compress the dimension of the problem from $O(2^n)$ to $O(nd^2)$ where d is the dimension of the individual tensors in Equation 55.

Among all these simulation methods, Schrödinger style simulation is widely used as the mainstream simulation method, and has been widely adopted in prior research works.[412,417–420,425,431–434] Also, industrial quantum circuit simulators such as IBM QISKit,[421] Google Qsim-Cirq[417,432] and Microsoft QDK[423] use full state vector simulations. In this work, we build Q-GPU based on IBM QISKit-Aer, a high-performance C++ simulation backend of QISKit, since it contains the state-of-the-art GPU support.

A.4 Characterization of QCS

A.4.1 Quantum Circuit Benchmarks

In this paper, we characterize the performance of QCS using a rich set of quantum circuits. Table 9 lists the circuit benchmarks.

Table 9: List of quantum circuit benchmarks.

Abbrv.	Application
`hchain`	Linear hydrogen atom chain[435]
`rqc`	Random quantum circuit[407]
`qaoa`	Quantum approximate optimization algorithm[436]
`gs`	Graph state[437,438]
`hlf`	Hidden linear function[439]
`qft`	Quantum Fourier transform[440]
`iqp`	Instantaneous quantum polynomial-time[441,442]
`qf`	Quadratic form[443]

- `hchain`: This circuit which describes a system of hydrogen atoms arranged linearly is a representative quantum chemistry application.[444–448] This circuit incorporates increased circuit depth and an early entanglement in terms of total operations.
- `rqc`: The random quantum circuit from Google[407,449] is used to represent the quantum supremacy compared to classical computers.
- `qaoa`: Quantum approximate optimization is a promising quantum algorithm in the NISQ era that produces approximate solutions for combinatorial optimization problems.[436]
- `gs`: This circuit is used to prepare graph states[450] that are multi-particle entangled states. Examples include many-body spin states of distributed quantum systems that are important in quantum error correction.[451]
- `hlf`: This benchmark circuit solves the 2D hidden linear function problem.[439]
- `qft`: The quantum Fourier transform circuit[440] is the quantum analog of the inverse discrete Fourier transform. It is an important function in Shor's algorithm.[452]
- `iqp`: The instantaneous quantum polynomial circuit provides evidence that sampling the output probability distribution of a quantum circuit is difficult when using classical

approaches.[441,442]

- `qf`: This circuit implements a quadratic form on binary variables encoded in qubit registers. It is used to solve the quadratic unconstrained binary optimization problems.[443]

A.4.2 Baseline QCS

In this paper, we use the popular IBM QISKit-Aer simulator. We consider the state-of-the-art GPU-supported simulation[420] in QISKit-Aer as our baseline. We run all simulations on a server with dual 10-core Intel Xeon Silver 4114 CPUs at 2.2 GHz, 384 GB of memory, and an NVIDIA P100 GPU with 16 GB of memory connected through PCI-e[2]. We use CUDA v10 and Nvprof[453] to conduct our characterization. The simulation in QISKit-Aer has three key steps: 1) state vector partitioning, 2) static state amplitudes allocation, and 3) on-demand amplitudes exchange.

Step 1: State vector partitioning: QISKit-Aer first partitions the state vectors into "chunks". Chunk is the granularity used in the simulator to update the state vector. For illustrative purposes, let us assume we have a 7-qubit circuit, i.e., that there are in total 2^7 different state amplitudes from $a_{0000000}$ to $a_{1111111}$. All the states are stored in a vector (i.e., the state vector), and this state vector is partitioned into chunks. For example, assuming we divide the state vector into 8 chunks, each chunk contains 16 state amplitudes as shown in Figure 15. The three most significant bits are used to index the chunks, and the remaining bits are as offsets within a chunk.

Step 2: Static chunk allocation: After partitioning, these chunks are allocated into GPU memory based on the GPU memory availability. As illustrated in Figure 15, if a GPU can only store 3 chunks, the remaining 5 chunks will be stored in the host CPU memory. For example, when 64 GB memory is needed to simulate 32 qubits, the first 16 GB is allocated in GPU memory (in P100 GPU with 16 GB memory) and the remaining 48 GB is in the CPU memory.

Step 3: Reactive chunk exchange: During circuit simulation, a chunk exchange between

[2]We used the P100 GPU as the HPC platform to test and evaluated the proposed Q-GPU. It is important to emphasize that our approach is not bound to P100. Our proposed Q-GPU is applicable to any other GPU architectures.

the GPU and the CPU arises when the requested state amplitudes are not locally available on the GPU. In QISKit-Aer, the chunk exchange between the CPU and the GPU is triggered on-demand. That is, when both the chunks on the CPU and the GPU are involved in one state-update calculation, the corresponding CPU chunks are transferred to GPU for updating. After the operation, the updated chunks are transferred back to the CPU. Note that, the amount of data exchange in the following scenarios is dependent on the qubits in the specific gate simulation.

- **Case 1: All the indices of the qubits involved in the current gate are smaller than the chunk size:**. For example, a gate on qubit 0 requires amplitudes $a_{\times\times\times\times\times\times\times 0}$ and $a_{\times\times\times\times\times\times\times 1}$ (see Equation 53). In this case, each chunk can be updated independently without requiring extra data movement.
- **Case 2: Some indices of qubits involved in the current gate are outside the chunk boundary:** In this scenario, let us assume there is a gate that operates on q_6, thereby the required pairs of amplitudes are $a_{\times 0\times\times\times\times\times}$ and $a_{\times 1\times\times\times\times\times}$. However, as depicted in Figure 15, none of the chunks contains a pair of required amplitudes, i.e., the computation for updating amplitudes involves more than one chunk. Specifically, to update the pairs of amplitudes, we need ($chunk_0$, $chunk_2$), ($chunk_1$, $chunk_3$), ..., and ($chunk_5$, $chunk_7$). However, ($chunk_1$, $chunk_3$) involves one chunk on the GPU and one chunk on the CPU. In this scenario, data exchange is required. In the baseline QISKit-Aer simulation, the requested chunks are always copied from CPU to GPU. That is, in the example above, the CPU copies $chunk_3$ to GPU. After the $chunk_3$ is updated together with $chunk_1$, it is copied back to the CPU memory.

Note that, as the GPU memory capacity is much less compared to the CPU host memory, a large number of chunks are statically allocated on CPU memory when the number of qubits is large. For instance, on the P100 GPU with 16 GB memory, we observe from experiments that when simulating a circuit that has 34 qubits, the state vector is divided into 8192 chunks, 496 chunks are allocated on GPU, while the remaining 7696 chunks are all on CPU. Therefore, one can expect that most of the time, the CPU does the state amplitude update without benefiting from the GPU acceleration.

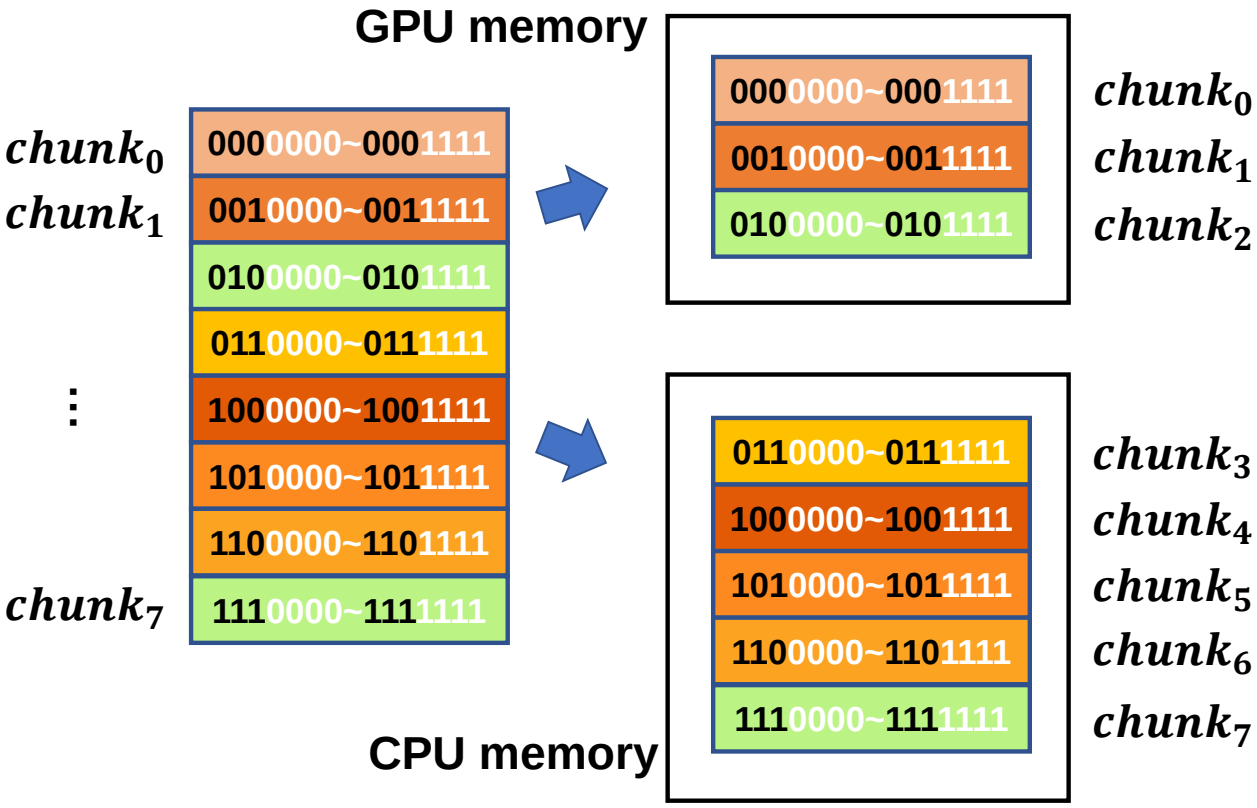


Figure 15: Example of baseline execution where the state vector is statically partitioned and allocated on CPU and GPU.

A.4.3 Characterization and Observations

In this section, we quantify the simulation performance of the baseline QISKit-Aer. We first study the scalability when the number of qubits increases. We observe that, if there are less than 30 qubits in the circuit, the baseline GPU simulates much faster than compared CPU-based simulation (e.g. 9.67× speedup for 29-qubit circuits on average), since the entire state vector fits in the P100 GPU memory and there is no need for data exchange and synchronization. However, the baseline GPU performance significantly drops when the number of qubits is larger than 30. It becomes even worse than running on the CPU alone when the number of qubits reaches 32. In particular, we observe a factor of 1.8× slowdown

for `qft_33`[3] as an example.

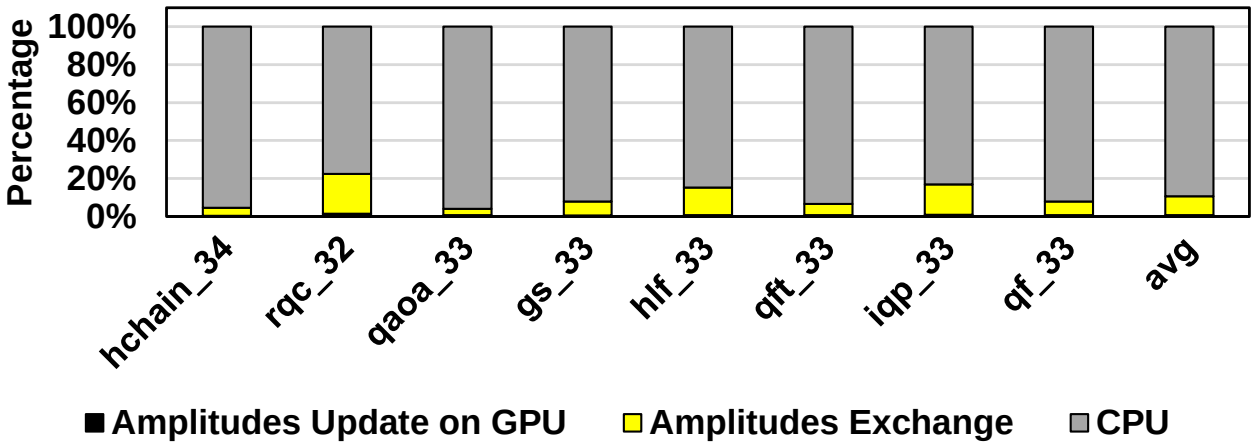


Figure 16: Baseline execution time breakdown.

To investigate the reason for this slowdown, we show the breakdown of the execution time in Figure 16. One can observe that, on average, 89.34% of the execution is spent on the CPU, indicating that the GPUs are not properly used in the baseline execution for large number qubit circuits. Moreover, the overheads involve amplitude exchange and synchronization occupies 9.91% of the average execution time, and the computation time of GPU only occupies 0.71% of total time on average. *In other words, most of the computation is performed by the CPU and the GPU is idle due to the static state chunk allocation in the baseline GPU execution.* In Figure 20, ① depicts the execution timeline of the baseline.

A.4.4 Will a Naive Optimization Work?

To improve the GPU utilization during simulation, an intuitive optimization would dynamically allocate the chunks and transfer the chunks to GPU for updates. In this section, we investigate whether the naive implementation works well or not.

We implemented the dynamic state vector chunk allocation in QISKit-Aer. Figure 17 depicts the execution time of the naive optimization normalized to the baseline execution. Surprisingly, none of the quantum circuits we studied show improvements when using dynamic allocation. To further investigate the reason, we break down the execution time and

[3]In this paper, we use n in the circuit name (e.g., `circ_n`) to represent a circuit with n qubits.

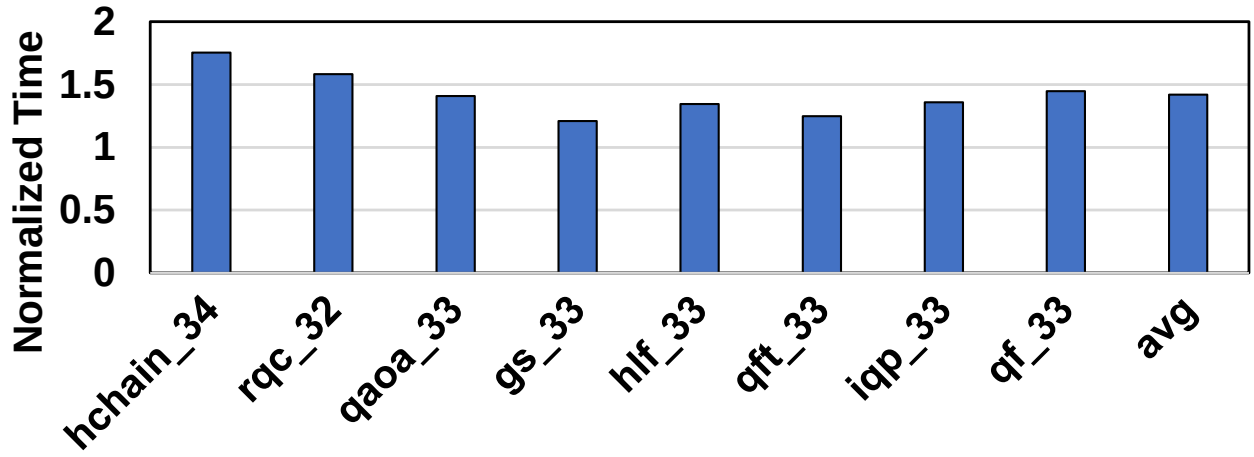


Figure 17: Normalized execution time of naive approach.

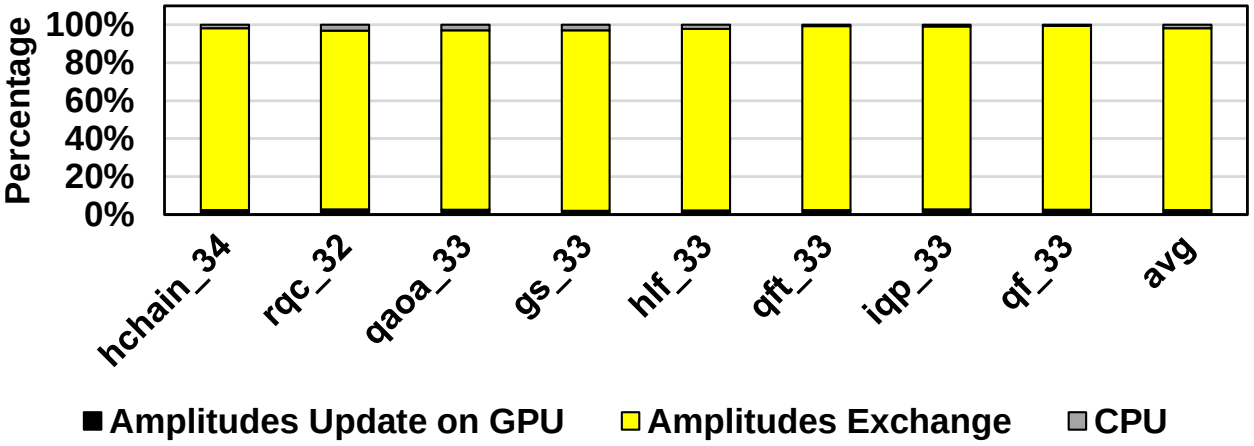


Figure 18: Execution time breakdown of naive optimization.

show the results in Figure 18. As can be seen from the figure, while CPU execution time significantly reduces and the data movement dominates, indicating that the GPU is waiting for data most of the time during execution. Therefore, naive dynamic allocation alone does not work to deliver good QCS performance. More sophisticated end-to-end optimizations are required to systematically improve the QCS performance and scalability.

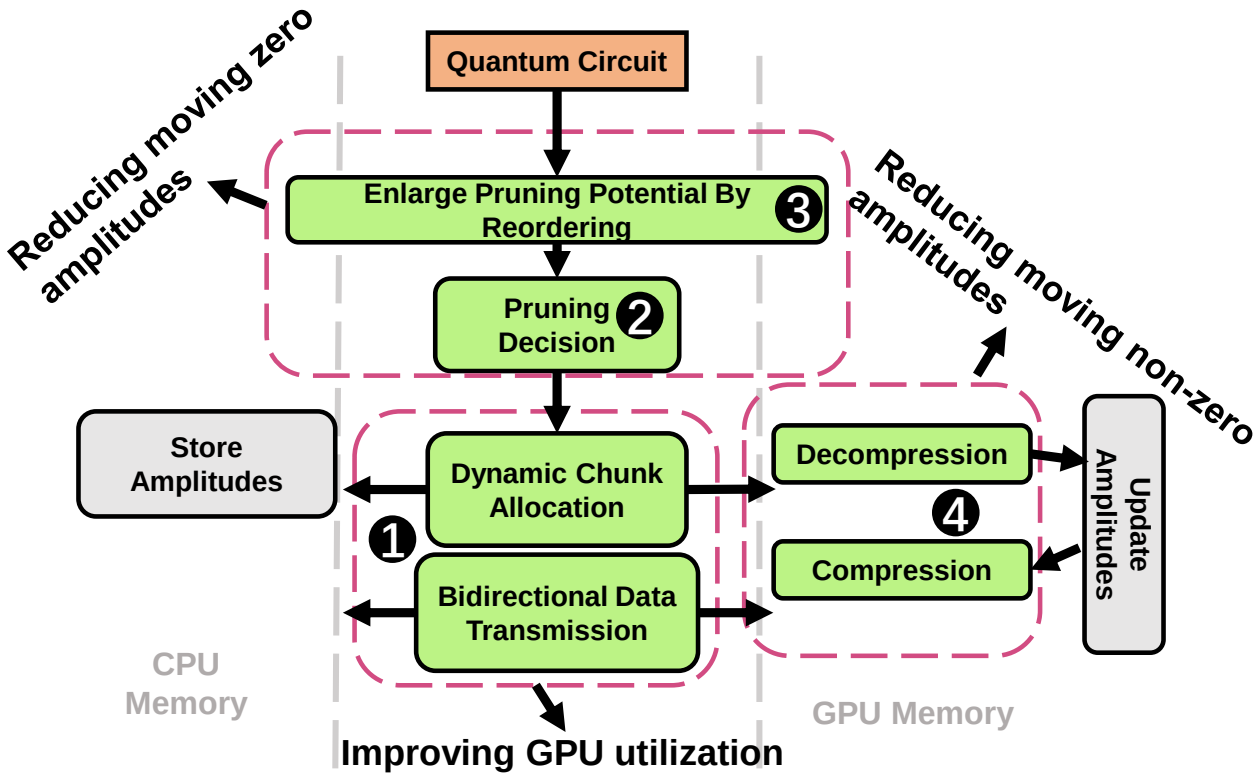


Figure 19: High level overview of Q-GPU.

A.5 Q-GPU

In this paper, we propose *Q-GPU*, a framework that features several end-to-end optimizations. Figure 19 depicts the high-level overview of Q-GPU. (❶) Q-GPU performs proactive state amplitude transfer to fully utilize the bi-directional data transfer bandwidth between CPU and GPU (Section A.5.1). (❷) Before copying state amplitudes to GPU, Q-GPU performs dynamic redundancy elimination that prunes zero state amplitudes to avoid unnecessary data movements (Section A.5.2). (❸) Q-GPU features a compiler-assisted, dependency-aware quantum gate reordering to enlarge the potential of pruning (i.e., the number of zero amplitudes). (❹) Q-GPU implements a GPU-supported, lossless amplitude compression to further reduce the data transfer caused by non-zero state amplitudes with minimal runtime overheads (Section A.5.4).

A.5.1 Proactive State Amplitudes Transfer

In the naive execution, one reason behind the poor GPU utilization is the sequential state amplitude transfer between CPU and GPU. Specifically, when the GPU finishes updating all local chunks, those chunks are first copied back to CPU memory before the CPU can transfer the next batch of un-updated chunks to the GPU. This restriction is reasonable in the scenarios when particular chunks are involved in consecutive updates since the chunks being copied from the GPU's memory cannot be overwritten during the copying. In other words, data movements are synchronized to avoid data conflicts. However, if the subsequent chunks from the CPU are not copied to the same memory locations on the GPU where current chunks are stored, such data conflict does not exist. As a result, one can transfer the chunks simultaneously from the CPU to the GPU and from the GPU to the CPU.

In our work, Q-GPU leverages CUDA streams to enable concurrent and bi-directional chunk copy to fully utilize the available bandwidth between the CPU and GPU. To avoid potential data conflict, Q-GPU implements two CUDA streams and partitions the GPU memory into two halves. One stream is responsible for the first half partition that acts as a buffer holding the chunks the GPU is currently updating. The other stream is responsible for the second half partition that acts as a buffer for "prefetching" the next chunks for the GPU to update. The two memory partitions work as "circular buffers" to feed the GPU with the required chunks. These two streams can potentially overlap and execute concurrently.

Figure 20 illustrates the timeline of the baseline and each of our optimizations. The proposed proactive state amplitude transfer (Ⓘⓘⓘ) achieves Ⓐ cycles savings compared with the baseline (Ⓘ). We also show that the naive approach (Ⓘⓘ) performs worse than the baseline.

A.5.2 Pruning Zero State Amplitudes

While overlapping improves the bandwidth utilization, the total amount of amplitudes that are transferred remains unchanged. To reduce the data movement, we observe that there exist a considerable amount of zero state amplitudes that do not need to be updated during simulation. Thus, those zero state amplitudes can be pruned before transferring the chunks.

Source of zero amplitudes: Let us assume there are n qubits, the initial states are

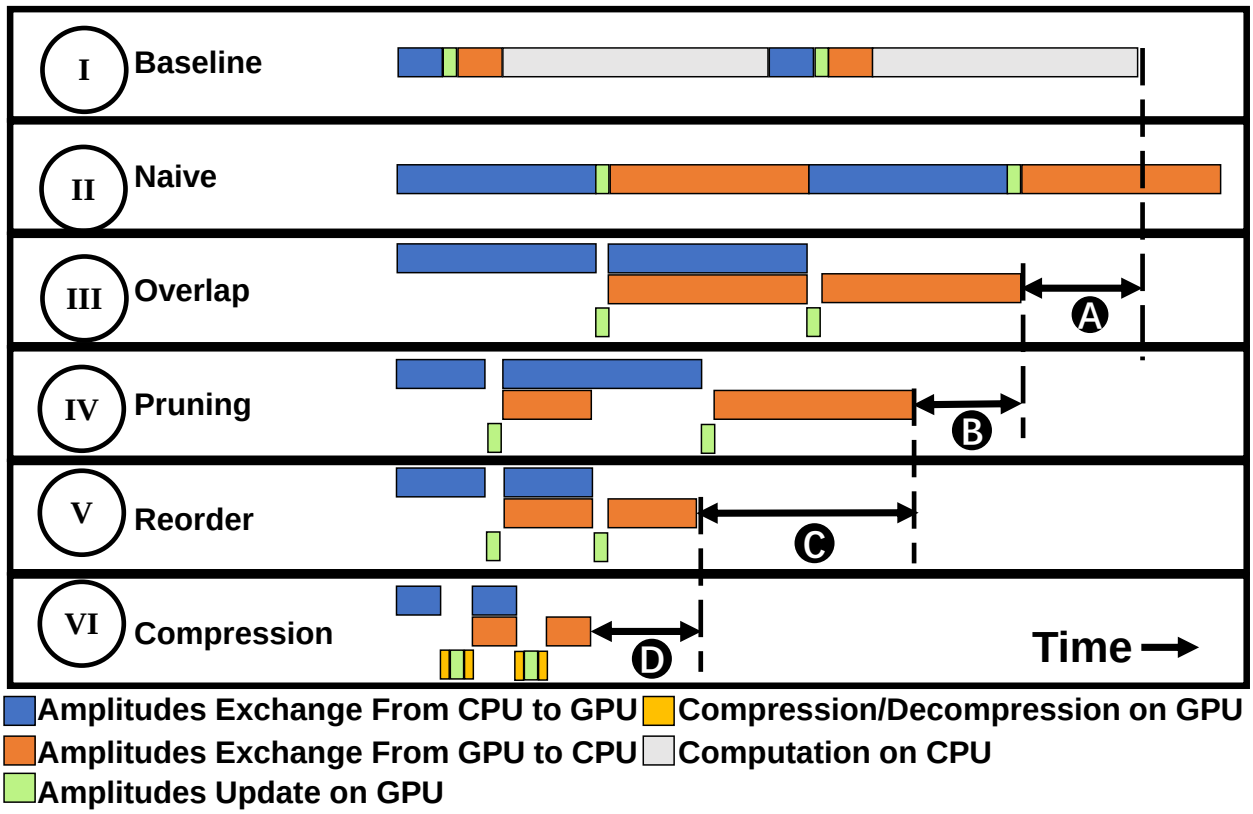
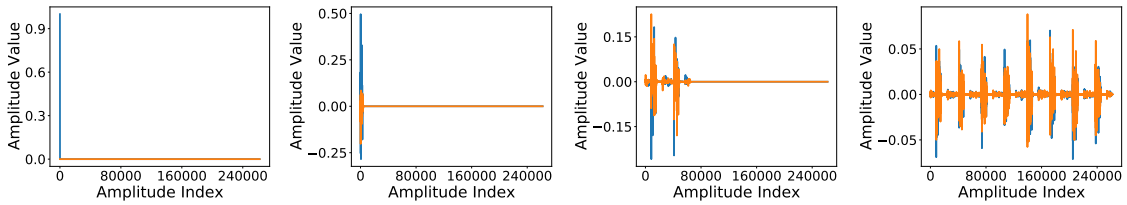


Figure 20: Time-line graph showing the benefits of each optimization in Q-GPU.

usually set as $|0\rangle^{\otimes n}$ in the general QCS, indicating that all qubits have zero probability of being measured as $|1\rangle$. Hence, all state amplitudes are zeros, except for $a_{0_1 0_2 \ldots 0_n}$ which is 1. As the state of a particular qubit is unchanged until an operation is being applied on it, its state remains $|0\rangle$ until that operation happens. For instance, if a particular qubit q_k is $|0\rangle$, all the state amplitudes $a_{\times \cdots \times 1_k \times \cdots \times}$ are zeros since q_k has zero probability to be measured as $|1\rangle$. In general, if m of n-qubits are not involved, amplitudes $a_{\times 0_{k_1} \times 0_{k_2} \cdots \times 0_{k_m} \times \times}$ are possible to be non-zero values, whereas the remaining amplitudes are guaranteed to be zero values, i.e. $2^n - 2^{n-m}$ amplitudes are zero values. Therefore, even if only one qubit is not involved, then half of the state amplitudes are zeros.

Pruning potential: To investigate the potential of pruning, Table 10 lists the number of total operations and the number of operations before all qubits are involved. For circuits like `iqp`, we can expect a significant reduction of data movement after pruning since many qubits

are not involved until the end of execution. However, for `qft` and `qf`, all qubits are involved at the beginning of execution, diminishing the potential of pruning benefits. We also use `hchain_18` as an example and plot the distribution of state amplitudes after each operation (i.e., quantum gate) being applied in a quantum circuit. Figure 21 shows the state amplitude distribution after 0, 30, 60 and 90 operations. One can observe that a large portion of state amplitudes are zeros at the beginning of the simulation. During simulation, the amplitudes are gradually updated to non-zero values since more qubits are involved.



(a) (b) (c) (d)

Figure 21: State amplitudes distribution of `hchain_18`, after 0, 30, 60 and 90 operations from left to right. Blue and orange lines denote real and imaginary parts of an amplitude respectively.

In general, let us assume we have an operation involving m states, if all of the states are zero, these m states remain zeros after applying any operation. As a result, we do not need to transfer the zero state amplitudes to the GPU as their values will not change. Therefore, one can reduce the data movement between CPU and GPU by pruning the zero state amplitudes. One intuitive approach is to check each state value by traversing all states. However, a more efficient approach can be adopted, as we illustrate below.

Pruning Mechanism: In the proposed Q-GPU, we use bits in a binary string as flags to indicate whether a qubit has been involved after a set of gate operations (denoted as *involvement* in Algorithm 1). Initially, all the bits in *involvement* are set to 0. When q_k is involved, the kth bit in *involvement* is set to 1. Recall that the state vector is partitioned into chunks, the index of a chunk, i.e., $iChunk$, determines whether a chunk will be transferred or not. To compare $iChunk$ with flag bits in *involvement*, we define $iChunk'$ as the left-shifted $iChunk$ to align with *involvements*. When $iChunk'$ is larger than *involvement*, it indicates

Table 10: The number of total operations and the number of operations before all qubits are involved for all circuits with 34 qubits.

Circuit	Total Operations	Operations Before Completely Involved
`hchain`	1786	272
`rqc`	124	54
`qaoa`	754	19
`gs`	37	16
`hlf`	48	16
`qft`	184	13
`iqp`	146	132
`qf`	222	16

that at least one bit of $iChunk'$ is 1 and the corresponding flag bit in *involvement* is 0. In this situation, the corresponding qubit (i.e., indexed by this flag bit) has not been involved by any operation. As such, we skip the remaining chunks and stop the iteration (line 5). On the other hand, if $iChunk'$ is smaller than or equal to *involvement*, the redundancy within a chunk is determined by $iChunk'$ & *involvement* (line 8). For a qubit whose corresponding bit in $iChunk'$ is 1, if it has already been involved by previous operations, its corresponding bit in *involvement* is also 1. Therefore, for all the qubits that is 1 in $iChunk'$, if all of them have already been involved by previous operations, $iChunk'$ & *involvement* results in $iChunk'$ itself. Otherwise, all the state amplitudes within this chunk are zeros, and we can prune this chunk. Moreover, the *chunkSize* here is dynamically determined rather than a statically fixed value, which enhances the benefit of the above-discussed strategy. Specifically, we select *chunkSize* by finding the least non-zero bit of *involvement*. This is useful, especially at the beginning of the simulation where many state amplitudes are zeros. For instance, assuming we have an 8-qubit circuit and the *involvement* flag is 00000011 at the early execution stage, the *chunkSize* is dynamically set to 2, which has fewer zeros within a chunk compared to

Algorithm 1: Pruning zero state amplitudes.

Variable list: N Total chunks number in CPU,
$involvement$ Flag indicating which qubits are involved

```
1  /* Determine chunkSize by locating the least non-zero bit of
       involvement                                                    */
2  chunkSize, N = getChunkSize(involvement)
3  for iChunk ← 0 to N − 1 do
4      iChunk' = iChunk << chunkSize
5      if iChunk' > involvement then
6          break
7      if iChunk'&involvement ≠ iChunk' then
8          continue
9      /* Amplitudes update                                          */
10     ...
11 updateInvolvement(involvement)
```

a larger chunk. The *involvement* flag bits are updated according to the qubits involved in each operation (line 14). In Figure 20, the proposed pruning mechanism (Ⓘⓥ) further saves (Ⓑ) cycle over Ⓘⓘⓘ.

A.5.3 Reordering to Delay Qubit Involvement

In order to enlarge the potential of pruning, such that more state amplitudes are zeros during simulation, we propose compiler-assisted, dependency-aware quantum operation reordering to delay the involvement of qubits. Specifically, when applying a gate, we choose the one that incurs the minimum number of additional qubits to be involved with those qubits that have been already involved by previous operations. For example, Figure 22a shows the `gs_5` circuit in the original execution order. The first five gates are H gates, where each gate applies to an individual qubit. As a result, once these gates have been applied, all

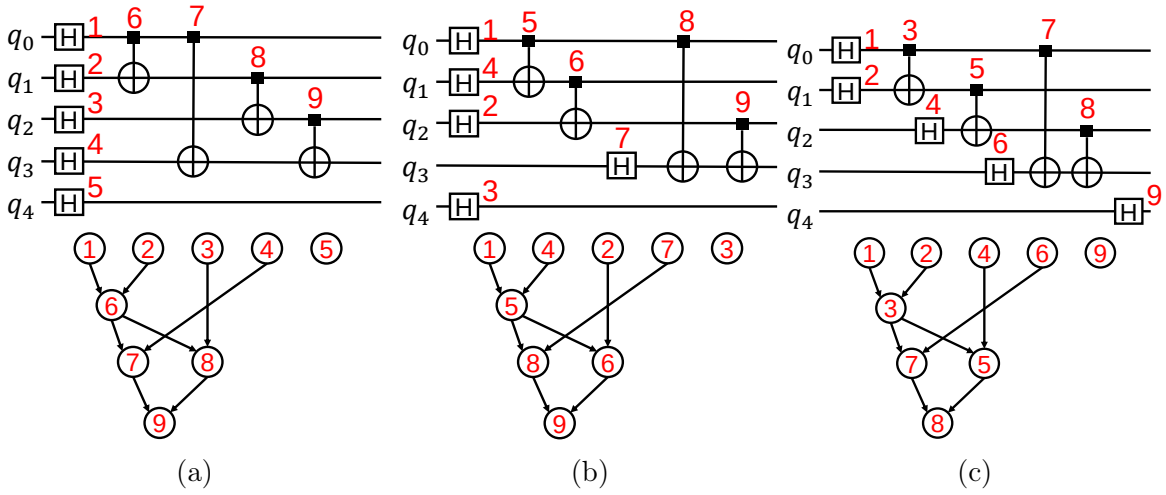


Figure 22: A walk-through example to illustrate the reordering benefits using `gs_5`. The red number denotes the operation orders before and after reordering.

the five qubits are involved. The next operation is a CNOT gate applied to qubits q_0 and q_1 ($CNOT_6$). All the state amplitudes are likely non-zero because the qubits are involved by the H gates. Therefore, applying this CNOT gate requires updating all the non-zero amplitudes in the state vector, leading to moving and traversing the entire state vector on the GPU. However, the $CNOT_6$ can be executed before some of the H gates without violating the circuit semantics. This gate reordering allows more zero state amplitudes (fewer data movements) when simulating the $CNOT_6$ gate. It is also important to emphasize that any reordering must ensure that the gate dependencies are presented. For instance, $CNOT_6$ and $CNOT_7$ cannot be reordered due to the dependency on q_0.

To this end, we propose a compiler-assisted optimization to reorder the gate sequence with the goal of delaying the qubit involvement. Specifically, gates that are applied on different qubits in a quantum circuit can be executed independently in any order and the execution sequence of these independent gates does not affect the final simulation result.[412,454,455] This provides us the opportunity to reorder the independent gates, we use a directed acyclic graph (DAG) to represent the gate dependency in a circuit. Based on the DAG, we reorder the

independent gates such that the simulation sequence involves the minimum number of new qubits when simulating each gate. Specifically, we investigate two heuristic strategies: 1) greedy reordering, and 2) forward-looking reordering.

Greedy reordering: greedy reordering traverses the DAG in topological order and greedily selects the gate (i.e., node in the DAG) that introduces the minimum number of new qubits to the list of updated qubits. The details of this method are illustrated in Algorithm 2. First, gates without predecessors in the DAG can be executed at the first steps and are put into *exeList*. Second, we traverse the gates in *exeList* and find the one that introduces the minimum number of newly involved qubits (lines 13 to 19). Then, we remove this gate from *exeList* and append it to the list of re-ordered gates. Third, we traverse the descendants of this gate and if a descendant does not have any predecessors other than this current gate, it will be added to *exeList* (lines 22 to 27). The second and the third steps are repeated until *exeList* is empty. In the rest of this section, we use Figure 22a as the example to illustrate how we perform reordering. At first, the *exeList* is [g_1, g_2, g_3, g_4, g_5]. Since each of these five gates involves one new qubit, we randomly select one gate among them to start simulation. In this example, g_1 is selected as the starting gate. After traversing all its descendants, no new gates can be added into *exeList*. Next, the *exeList* becomes [g_2, g_3, g_4, g_5]. In the next three steps, we randomly select g_3, g_5 and g_2 since no new gates can be executed and all gates in *exeList* have equal priority. Then the *exeList* becomes [g_4, g_6]. At this time, *involvedQubits* is [q_0, q_1, q_2, q_4]. Therefore, g_4 involves one new qubit (q_3), whereas g_6 will not introduce any new qubits since it acts on q_0 and q_1 that are already in the involved list. Therefore, we will greedily select g_4 to execute since it involves the least new qubits. One can follow these reordering steps to reach the new ordering shown in Figure 22b. As a result, the number of involved qubits at each step is $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 5 \rightarrow 5 \rightarrow 5$. Since the baseline is $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 5 \rightarrow 5 \rightarrow 5 \rightarrow 5$, the final involvement is delayed by two steps. However, a better solution for reordering is to select g_2 and g_6 in the second and the third step, since applying these two gates only adds one qubit to *exeList*, while applying g_3 and g_5 adds two. Thus greedy reordering may misses the optimal choice.

Forward-looking reordering: To address the deficiency in greedy-reordering, we propose Forward-looking reordering that looks ahead of all the equal-priority gate candidates before

Algorithm 2: Quantum operation reorder.

Input : DAG A DAG representing circuit dependencies.

Output: $gatesList$List of gates after reordering,

```
1  gatesList = [ ]
2  exeList = [ ]
3  /* First we build DAG and push gates without predecessors to an
       execution list                                                */
4  for g in DAG do
5      if g.numPredecessors() == 0 then
6          exeList.append(g)
7  /* Then we traverse DAG in topological order and greedily decides the
       execution order of the gates                                  */
8  while exeList ≠ ∅ do
9      nextGate = NULL
10     minCost = 0
11     for g in exeList do
12         cost = g.getCost()
13         if cost < minCost then
14             minCost = cost
15             nextGate = g
16     exeList.erase(nextGate)
17     gatesList.append(nextGate)
18     for g in nextGate.descendants() do
19         g.numPredecessors() = g.numPredecessors() − 1
20         if g.numPredecessors() == 0 then
21             exeList.append(g)
```

making a decision. We implemented a *cost* counter to determine the priority of the gates in *exeList*. In *greedy reordering*, the *cost* is simply computed by counting new involved qubits (line 3-8 in Algorithm 3). The *cost* in forward-looking reordering is computed using Algorithm 3. Note that, *exeList* and *involvedQubits* are just copies of the original ones, thus their original values are not changed. In forward-looking reordering, the *cost* of selecting a gate in *exeList* consists of two components: *costCurrent* and *costLookAhead* (line 1). The *costCurrent* is the same with the *cost* used in greedy reordering. Let us still use the example in Figure 22a to illustrate Algorithm 3. Initially, the *exeList* is also [g_1, g_2, g_3, g_4, g_5]. We take g_1 as an example to explain the computation of *costLookAhead*. First, we assume g_1 has already been executed. Then, the *costCurrent* is 1 and *involvedQubits* becomes [q_0] (lines 3-8). Since no descendants of g_1 can be executed, the *exeList* becomes [g_2, g_3, g_4, g_5] (lines 9-14). Then, we traverse the *exeList*. For each gate in *exeList*, we compute the cost of selecting this gate by counting the new involved qubits (lines 18-21) and selecting the least cost as *costLookAhead*. Now, executing any gate in *exeList* will involve one new qubit, thus *costLookAhead* is computed as 1 (lines 16-26). Similarly, one can find that all gates at the first step have equal priority. For the purpose of illustration, we assume g_1 is randomly selected. Then the *exeList* becomes [g_2, g_3, g_4, g_5]. Although all gates still have equal *costCurrent*, we can find that g_2 has the least *costLookAhead*. The reason is that, when we assume executing g_2 and look ahead from g_2, we find that executing g_6 introduces no new qubits. In contrast, look ahead after executing other gates will introduce new qubits. Finally, we get the result of forward-looking reorder as shown in Figure 22c. Clearly, the *involvement* at each step become $1 \rightarrow 2 \rightarrow 2 \rightarrow 3 \rightarrow 3 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 5$. Compared with greedy reordering, we further delay the final involvement by two steps.

Reorder effectiveness: To assess the performance of the reordering algorithms discussed above, we implement them to reorder the original operation sequences for all benchmark circuits that have 22 qubits and plot the *involvement* (Algorithm 1 in Section A.5.2) after each gate has been applied. For the purpose of illustration, we depict the results of three representative benchmark circuits in Figure 23. For each order, i.e original order, greedy-reorder, and forward-looking reorder, the "speed" of reaching the maximum *involvement* indicates the pruning potential. We observe that, forward-looking reordering results in the

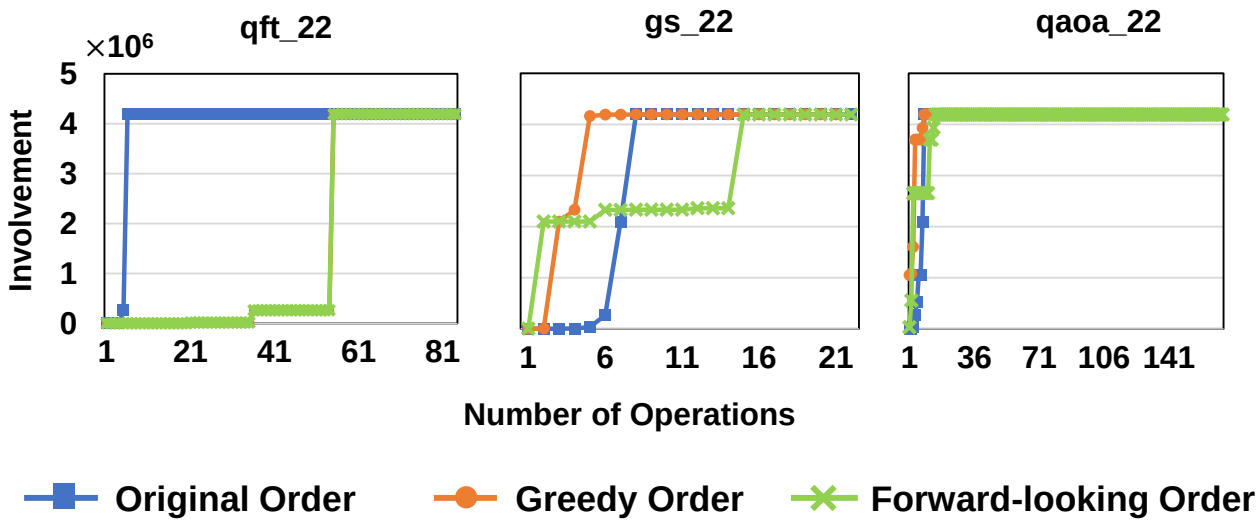


Figure 23: Qubit *Involvement* during simulation in three representative circuits.

largest pruning potential, while greedy reordering only works for `qft_22` and even results less pruning potential than baseline for `gs_22`. Particularly. for `gs_22` and `qft_22`, forward-looking reordering effectively delays the involvement of qubits. Thus, we can expect the pruning potentials of these circuits to be enlarged by forward looking reordering. However, for `qaoa_22`, none of the reordering algorithms work due to the prevalent dependencies among the gates. Refering back to Figure 20, when reordering (Ⓥ) is employed, we can prune more chunks, which saves additional Ⓒ cycles compared to (IV).

A.5.4 Non-zero State Compression

Compressibility: While pruning removes the zero state amplitudes, those non-zero amplitudes still cause data movement overheads especially for circuits that do not have large pruning potentials (e.g., `qaoa` in Figure 23). Targeting reducing the data movement caused by non-zero state amplitudes, we investigate the potential compressibility and propose a GPU-supported efficient lossless data compression in Q-GPU. Specifically, we observe that many non-zero entries within a state vector, after each operation, have similar amplitude values. In other words, there is a significant "spatial" similarly among consecutive state am-

plitudes in the state vector. To demonstrate the compressibility, we use `qaoa_20` and `iqp_20` as examples and show the residuals by subtracting the consecutive state amplitudes. As one can observe from Figure 24, for `qaoa_20`, most of the residuals are zero or very close to zero, indicating a potential for residual-based compression. However, `iqp` will be less compressible due to more diverse distribution.

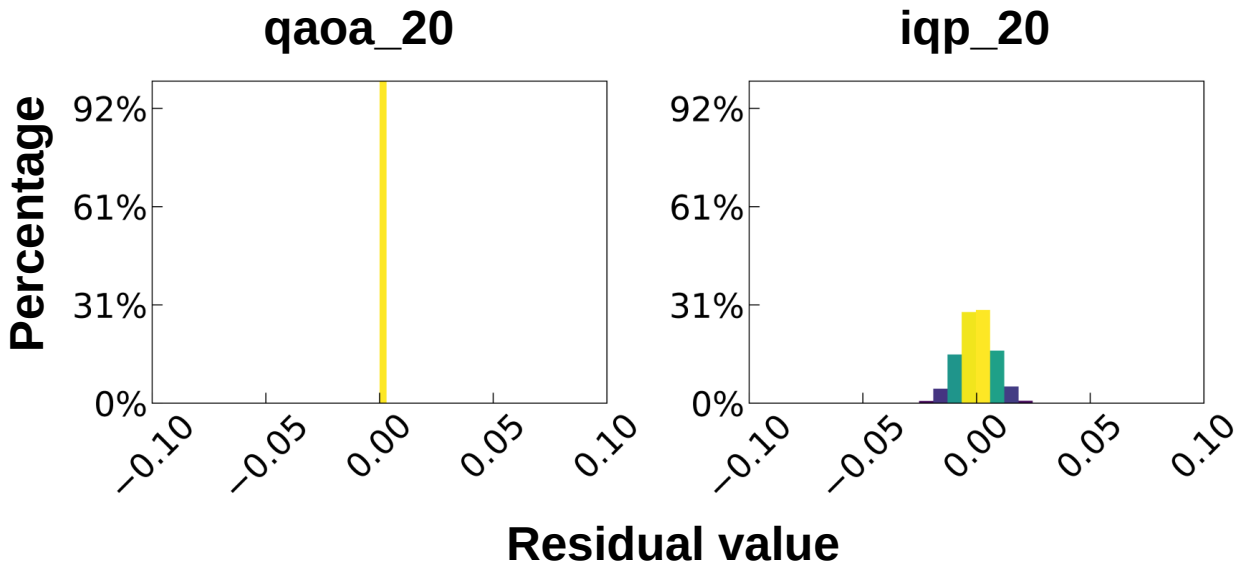


Figure 24: Residual distributions for `qaoa_20` and `iqp_20`.

Compression Strategy: We use the GFC algorithm[456] in Q-GPU. We implement the GFC as GPU kernels to perform the compression in parallel, thereby reducing the compression and decompression overheads. Specifically, the amplitudes on the GPU are partitioned into micro-chunks with a size of 32 amplitudes. Each GPU warp iteratively compresses/decompresses in parallel. Figure 25 (on the top) shows the compressed format. For the 32 values of a micro-chunk, we first store a 4-bit prefix for each of them, where one bit is used to record the sign of the residual and another three bits are a count of leading zero bytes of the residual. Figure 25 also illustrates the GPU support of compression and decompression in Q-GPU. The compression is performed on the GPU after updating the chunk before copying it to the CPU. All of the chunks are equally divided into "segments". We empirically choose the segment size to match the GPU parallelism such that the GPU is properly utilized during compression. The compressed segments are transferred to the CPU

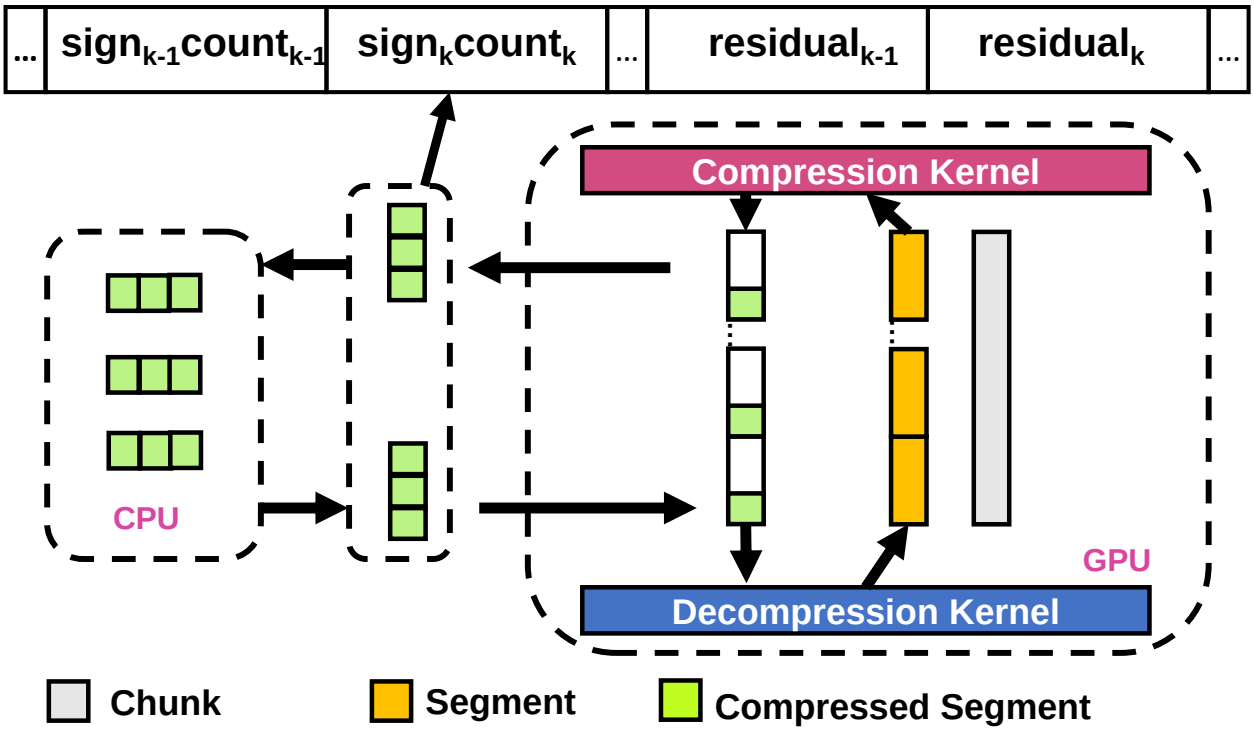


Figure 25: Overview of compression in Q-GPU.

instead of the original state chunks. The CPU keeps the compressed segments and copies the compressed segments to the GPUs upon request. Once the chunks are copied to the GPU, the amplitudes are decompressed, updated, and then compressed. As can be seen from Figure 20, compression (Ⓥ) saves Ⓓ cycles over Ⓥ and introduces negligible overhead. Later, in section A.6, we quantify the overheads incurred by the compression and decompression procedures.

A.6 Experimental Evaluation

In this section, we evaluate Q-GPU using the eight circuits in Table 9. We implement Q-GPU by substantially extending IBM QISKit-Aer. The evaluation is conducted on the same CPU-GPU platform used for characterization. For all experiments, the default opti-

mizations in QISKit-Aer are turned on in both baseline and Q-GPU evaluation. To show the effectiveness of each optimization, we test *six different versions* of executions for all quantum circuit benchmarks:

- *Baseline:* This version is the implementation with state-of-the-art GPU support[421] in QISKit-Aer that supports GPU acceleration. As illustrated in Section A.4.2, state amplitudes are statically allocated on the GPU and CPU in this version.
- *Naive:* This version is the intuitive implementation discussed in Section A.4.4, which dynamically allocates state amplitudes to GPU. The performance of this version is dominated by expensive data movements.
- *Overlap:* This version implements the first optimization – proactive state amplitude transfer – in Q-GPU. This version is built upon the *Naive* version and its details are discussed in Section A.5.1.
- *Pruning:* This version adds the proposed pruning mechanism (Section A.5.2) to *Overlap*. By skipping the data movement of zero state amplitudes, the amount of data movement is reduced.
- *Reorder:* In this version, we implement *forward-looking reorder* algorithm (Section A.5.3) to enlarge the potential for pruning. This reordering is performed by a simple compiler pass integrated in the Q-GPU.
- *Compression/Q-GPU:* In this version, all optimizations are employed with compression. We also call it *Q-GPU*. Compression (Section A.5.4) is added on top of *Reorder*. This version achieved the best performance.

A.6.1 Overall Performance

Figure 26 shows the overall performance and scalability among the six versions for all eight quantum circuits. The y-axis in the figure denotes the normalized execution time to the *Baseline* version. From the figure, one can make the following observations. First, by adding the proposed optimization in Q-GPU, our approach significantly reduces the execution time of QCS across all the circuits. Specifically, *Overlap*, *Pruning*, *Reorder*, and *Compression/Q-GPU* see a 24.96%, 44.54%, 56.78%, and 71.66% execution time reduction

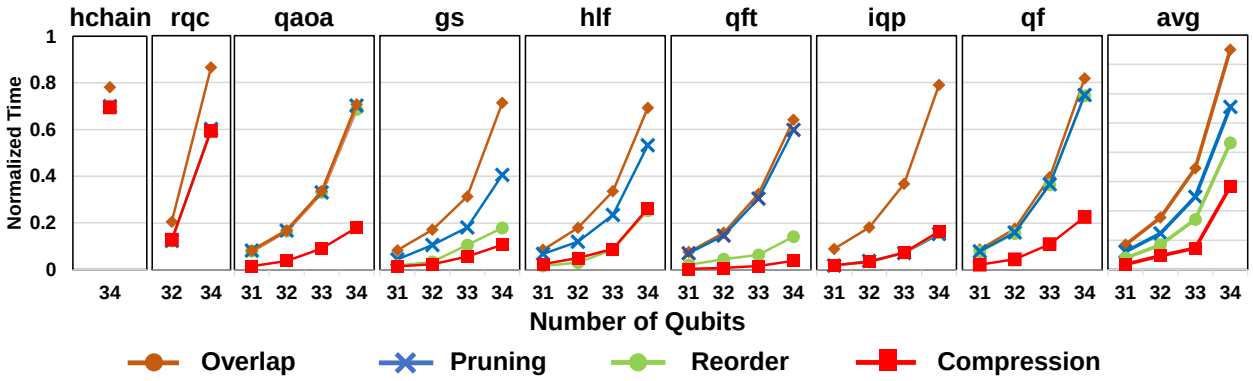


Figure 26: Normalized simulation time for circuits with different number of qubits (the lower the better).

over the baseline execution for the largest number of qubits that can run on our platform. Second, the scalability of QCS performances is significantly improved by "breaking" the memory capacity in Q-GPU. The average achieved performance outperforms baseline by 2.53× for 34 qubits. Although we only simulate up to 34 qubits due to the CPU memory limitation (384 GB) in our system (Section A.4.3), one can infer from the trend that our optimizations are scalable to larger sized circuits. Third, Q-GPU has different accelerations for different circuits. Specifically, for `gs`, `qft`, `qaoa` and `iqp`, higher execution time reduction is observed, whereas for `hchain` and `rqc`, less speedup is observed. This is because, for `hchain` and `rqc`, reordering cannot enlarge the pruning potential because of dependent gates. Their amplitude residuals also have disperse distribution (similar to `iqp` in Figure 24). Thus, either *Reorder* or *Compression* improves little for these two benchmarks. Finally, for different circuits, a certain version may not have the same acceleration effects. For example, *Overlap* version generates a similar execution time reduction in all circuits tested. However, for *Pruning*, *Reorder* and *Compression*, the runtime reduction is different between different circuits. For example, *Pruning* and *Reorder* improve little for `qaoa` and `qf` because these two circuits do not have much potential of pruning the zero amplitudes. That is, their qubits get involved quickly with dependent operations. However, `qaoa` achieves significant benefits by compression as the great potential of compressibility. (discussed in Section A.5.4).

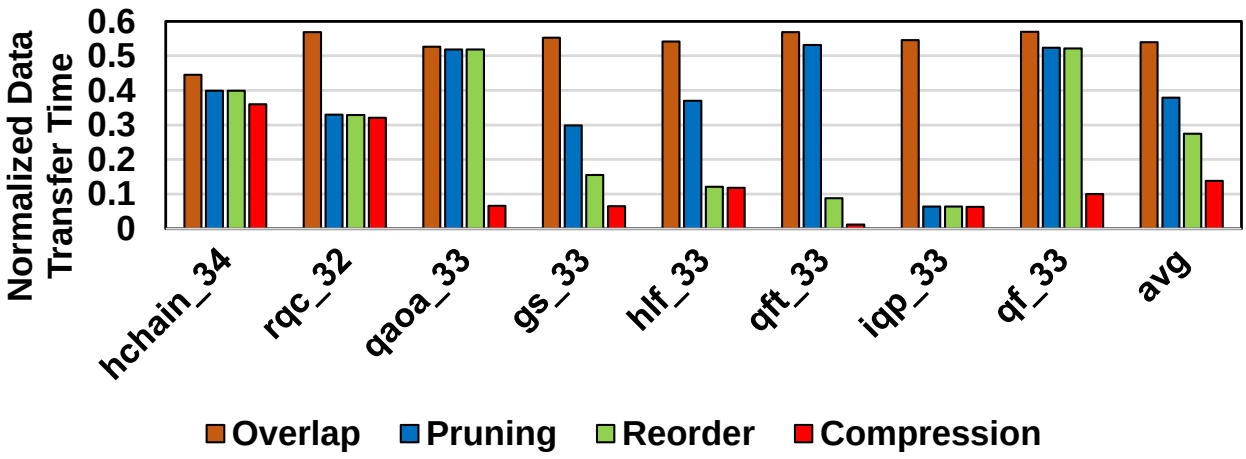


Figure 27: Normalized data transfer time (lower the better).

To further understand the execution reduction, Figure 27 plots, for each version, the exposed data movement time. In this figure, the y-axis represents the data movement time normalized to the *Naive* version. Clearly, one can observe a step-wise data movement reduction in the versions with our optimizations. First, *Overlap* uniformly reduces the data transfer time by an average of 46.14%. Note that, the savings generated in *Overlap* are independent of circuit types, that is the reason behind execution time reduction in Figure 26. For *Pruning* and *Reorder*, the reduction of data movement time varies in different circuits. This is because the number of zero state amplitudes and the potential of pruning heavily rely on the circuit type. For example, `qaoa`, `qft`, and `qf` get all qubits involved at early stage of simulation. Hence, pruning is less effective for these circuits compared to others. Also, as discussed in Section A.5.3, *Reorder* has little effects on `hchain`, `rqc`, `qaoa`, and `qf` due to dependent operations in these circuits. Therefore, *Reorder* delivers similar data transfer time reduction with *Pruning* for these circuits. However, for those circuits with less dependent operations, *Reorder* significantly reduces their data movement time by enlarging the pruning potential. For circuits like `qaoa`, `gs`, `qft` and `qf`, *Compression* effectively reduces the data movement by leveraging the spatial similarity discussed in Section A.5.4. In a nutshell, for all circuit benchmarks tested, the reductions of data transfer time are the main reason behind the execution time reduction in Figure 26.

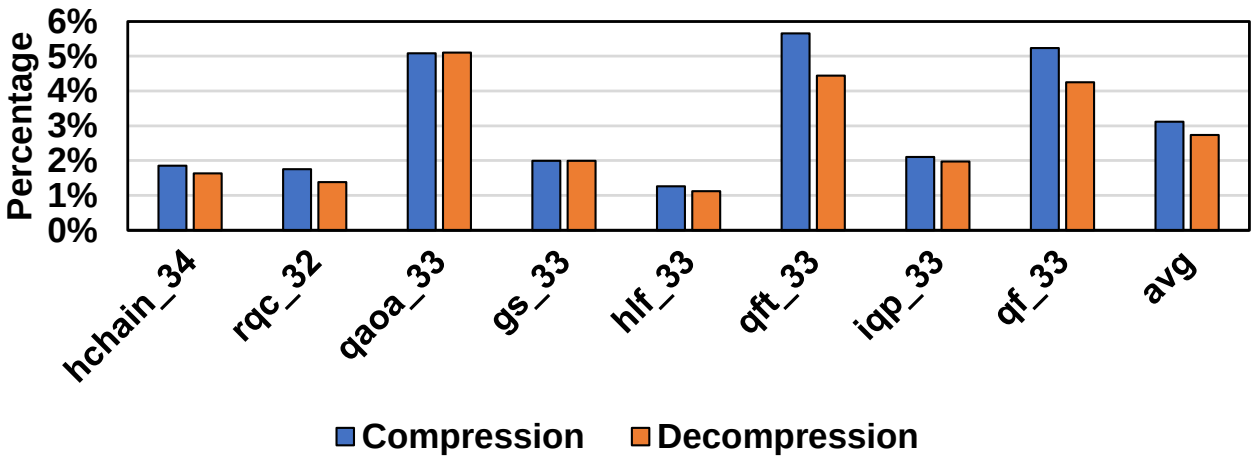


Figure 28: Compression and decompression overheads.

We also quantify the computation time of compression and decompression in Figure 28. Overall, the compression and decompression overhead is 3.12% and 2.74% of the GPU execution time. Potentially one may further optimize the compression and decompression by overlapping them on GPU, but we found the overhead is negligible compared to the significant reduction in execution time that we achieved. We also want to emphasize that the execution times reported in Figure 26 have all the sources of overhead included.

A.6.2 Comparison with OpenMP

Many publicly available quantum simulators and existing works employ OpenMP to parallelize the QCS on CPUs.[412,419,432] We compare Q-GPU with these OpenMP implementations. Specifically, we chose the OpenMP implementation in the most recent QISKit simulator and plot the results in Figure 29. We also compared our approach with other simulators in the next section, where the OpenMP is used by default in the simulators. On average, across eight circuits, Q-GPU outperforms the OpenMP QISKit by 1.79×. Particularly, Q-GPU achieved 12.79× speedup in `qft`. For `gs`, `iqp` and `qf`, Q-GPU achieves more than 2× speedup. However, for `hchain` and `rqc`, Q-GPU performs worse than OpenMP. This is because the pruning potential and the compressibility are low in both circuits where

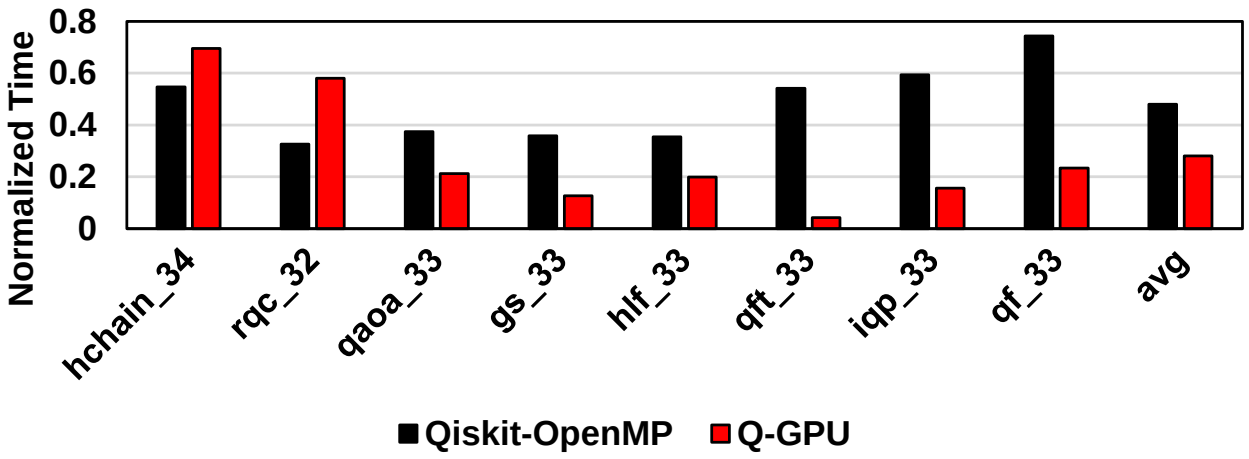


Figure 29: Comparison with OpenMP.

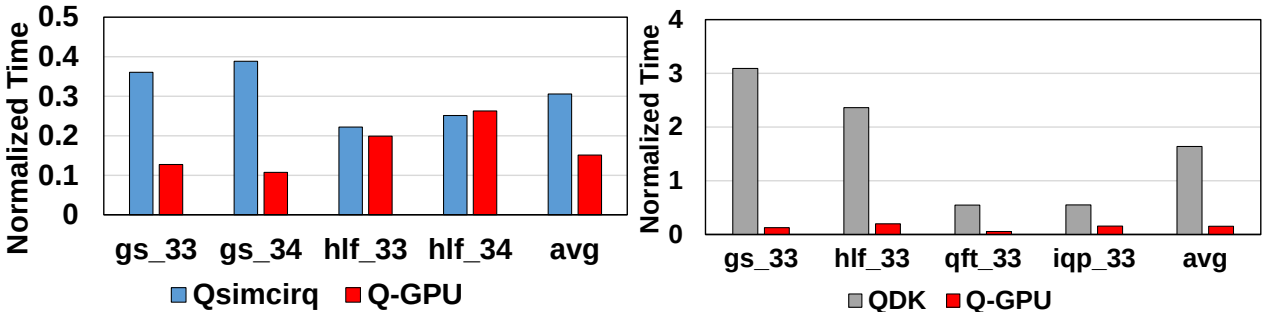
Q-GPU is less effective.

A.6.3 Comparison with Other Simulators

We compare Q-GPU with other simulators, including Google Qsim-Cirq v0.8.0 plus Cirq v0.9.2[422] and Microsoft QDK v0.15.[423] In our experiments, we run these simulators on the same CPU (Section A.4.2). Note that, both Qsim-Cirq and QDK are OpenMP enabled and we observe that they used all available threads during execution on the CPU. We report the results in Figure 30.

It is important to note that, to enable the simulation of the same circuits on Qsim-Cirq, we need to first transform our circuit benchmarks into OpenQASM codes.[457] Then, we need to import the OpenQASM codes to Qsim-Cirq for execution. Unfortunately, not all the transformed circuits can be simulated on Qsim-Cirq due to the lack of support for particular gates (i.e., the "cp" gate cannot be recognized by Qsim-Cirq). As a result, we can only run `gs` and `hlf` successfully. This motivates our future research on uniform support of Quantum programming models. Figure 30a shows the normalized speedup of the proposed Q-GPU compared to Qsim-Cirq. Q-GPU outperforms the Google Qsim-Cirq by 1.02× on average.

To run the same quantum circuit on Microsoft QDK v0.15, we have to further convert



(a) Comparison with Qsim-Cirq

(b) Comparison with QDK

Figure 30: Comparisons of Q-GPU to the simulator from Microsoft QDK v0.15 and Google Qsim-Cirq v0.8.0.

.

the OpenQASM codes to “qsharp", i.e., the quantum language used in Microsoft. The conversion only succeeded for `qft`, `iqp`, `hlf`, and `gs`. The normalized simulation time is plotted in Figure 30b. On average, Q-GPU performs 9.82× better than Microsoft QDK.

A.7 Related Works

To the best of our knowledge, Q-GPU is the first work that systematically optimizes quantum circuit simulation on a GPUs. We summarize the related prior efforts below.

Prior works have focused on QCS optimizations on different platforms, from readily available devices to cloud environments.[412,417–419,428,431,432,458] Thomas et al.[417] simulated 45-qubits circuit using 8,192 nodes. They optimized single node performance by using automatic code generation and optimization of compute kernels. Edwin et al.[419] claimed to simulate more than 49 qubits by partitioning quantum circuits to “subcircuits” and delay their entanglements. In,[418] the authors proposed lossy data compression to reduce the memory requirement of simulating large-scale quantum circuits. Aneeqa et al.[412] focused on fully exploiting single CPU performance for simulating a large number of qubits. The developed algorithm aims

to reorder circuits such that more gates can be simulated in parallel. Compared with all these efforts, Q-GPU takes advantage of GPUs while managing the data movement between CPU and GPU. First, we identify the source of zero state amplitudes in QCS, and propose a pruning mechanism to safely reduce unnecessary computation on these states, which saves not only computation but also data movement. Unlike prior works using reordering to aggregate gates,[412,455,459] we propose reordering algorithms to enlarge the pruning potential. Moreover, Q-GPU is the first framework that leverages the GPU to implement a lossless compression that does not affect accuracy of QCS. Finally, it is important to emphasize that Q-GPU is complementary to existing cloud-based quantum simulation frameworks, and can be integrated within these frameworks for further QCS improvements.

There are also several works that utilize GPUs to accelerate QCS.[415,416,420,434,460–462] Most of these works have limited capability in simulating large quantum circuits due to the limited memory capacity of GPUs. Ang et al.[415] proposed a multi-GPU centric QCS framework that tracks the density matrix. However, their framework cannot simulate a large number of qubits since it is limited by the aggregated memory capacity of multi-GPUs. For a single-node, they can only simulate up to 14 qubits on an NVIDIA V100 GPU. Jun et al.[420] proposed a CPU-GPU co-simulation method that enables simulation using a GPU even when the required memory exceeds the GPU memory capacity. Their method is also integrated into the IBM QISKit and is used as the baseline in this paper. In summary, compared to prior work, Q-GPU breaks the GPU memory capacity limitation, i.e., it is able to simulate 34 qubits which require 256 GB memory on a 16 GB memory GPU, and fully takes advantage of GPU parallelization. The fundamental design innovation behind this is to dynamically and proactively transfer the state amplitudes through end-to-end optimizations to minimize the data movement overheads caused by state amplitudes transfer.

A.8 Concluding Remarks

In this paper, we propose Q-GPU, a framework tailored with GPU optimizations to effectively improve the quantum circuit simulation performance for quantum circuits with a large

number of qubits. The Q-GPU is able to deliver scalable simulation performance based on the four internal end-to-end optimizations, including i) proactive state amplitudes transfer, ii) zero state amplitudes pruning, iii) delayed qubit involvement, and iv) lossless non-zero state compression. Experimental results across eight representative quantum circuits indicate that Q-GPU achieves 2.53× average execution time reduction on a single GPU. It also outperforms the most recent OpenMP CPU implementation and other publicly available quantum simulators.

Algorithm 3: Cost calculation in forward-looking reordering.

Input : g Gates from $exeList$,

$exeList$ List of gates that are executable,

$involvedQubits$ Set of qubits which have already been acted on.

Output: $cost$ Potential involved qubits after executing g.

```
1  costCurrent = 0, costLookAhead = 0
2  /* First we compute additional qubits that will be acted on by
       executing current gate                                        */
3  for q in g.qubits() do
4      if q not in involvedQubits then
5          costCurrent = costCurrent + 1
6          involvedQubits.insert(q)
7  exeList.erase(g)
8  for g' in g.descendants() do
9      if g'.numPredecessors() == 1 then
10         exeList.push(g')
11 /* Then we traverse current exeList and compute the cost of selecting
       a gate that involve least additional qubits                   */
12 for g'' in exeList do
13     curCostLookAhead = 0
14     for q' in g''.qubits() do
15         if q' not in involvedQubits then
16             curCostLookAhead = curCostLookAhead + 1
17     if curCostLookAhead < costLookAhead then
18         costLookAhead = curCostLookAhead
19 cost = costCurrent + costLookAhead
20 return cost
```

Bibliography

[1] Jordan, K. D.; Luken, W. Theoretical Study of the Binding of an Electron to a Molecular Dipole: $LiCl^-$. *J. Chem. Phys.* **1976**, *64*, 2760–2766, DOI: 10.1063/1.432599 (pages 1, 20).

[2] Gutowski, M.; Skurski, P.; Boldyrev, A. I.; Simons, J.; Jordan, K. D. Contribution of Electron Correlation to the Stability of Dipole-Bound Anionic States. *Phys. Rev. A* **1996**, *54*, 1906–1909, DOI: 10.1103/PhysRevA.54.1906 (pages 1, 20, 21).

[3] Barnett, R. N.; Landman, U.; Cleveland, C. L.; Jortner, J. Electron Localization in Water Clusters. II. Surface and Internal States. *J. Chem. Phys.* **1988**, *88*, 4429–4447, DOI: 10.1063/1.453801 (pages 1, 20).

[4] Fermi, E.; Teller, E. The Capture of Negative Mesotrons in Matter. *Phys. Rev.* **1947**, *72*, 399–408, DOI: 10.1103/PhysRev.72.399 (pages 1, 20).

[5] Turner, J. E.; Anderson, V. E.; Fox, K. Ground-State Energy Eigenvalues and Eigenfunctions for an Electron in an Electric-Dipole Field. *Phys. Rev.* **1968**, *174*, 81–89, DOI: 10.1103/PhysRev.174.81 (pages 1, 20).

[6] Crawford, O. H. Bound States of a Charged Particle in a Dipole Field. *Proc. Phys. Soc.* **1967**, *91*, 279–284, DOI: 10.1088/0370-1328/91/2/303 (pages 1, 20).

[7] Garrett, W. R. Critical Binding of an Electron to a Non-Stationary Electric Dipole. *Chem. Phys. Lett.* **1970**, *5*, 393–397, DOI: 10.1016/0009-2614(70)80045-8 (pages 1, 20).

[8] Garrett, W. R. Critical Binding of an Electron to a Rotationally Excited Dipolar System. *Phys. Rev. A* **1971**, *3*, 961–972, DOI: 10.1103/PhysRevA.3.961 (pages 1, 20).

[9] Desfrançois, C.; Abdoul-Carime, H.; Khelifa, N.; Schermann, J. P. From $\frac{1}{r}$ to $\frac{1}{r^2}$ Potentials: Electron Exchange between Rydberg Atoms and Polar Molecules. *Phys. Rev. Lett.* **1994**, *73*, 2436–2439, DOI: 10.1103/PhysRevLett.73.2436 (pages 1, 20, 22).

[10] Jordan, K. D.; Wang, F. Theory of Dipole-Bound Anions. *Annu. Rev. Phys. Chem.* **2003**, *54*, 367–396, DOI: 10.1146/annurev.physchem.54.011002.103851 (pages 1, 20).

[11] Gutowski, M.; Skurski, P. Dispersion Stabilization of Solvated Electrons and Dipole-Bound Anions. *J. Phys. Chem. B* **1997**, *101*, 9143–9146, DOI: 10.1021/jp972600z (page 2).

[12] Gutowski, M.; Jordan, K. D.; Skurski, P. Electronic Structure of Dipole-Bound Anions. *J. Phys. Chem. A* **1998**, *102*, 2624–2633, DOI: 10.1021/jp980123u (page 2).

[13] London, F. The General Theory of Molecular Forces. *Trans. Faraday Soc.* **1937**, *33*, 8b–26, DOI: 10.1039/TF937330008B (page 2).

[14] Voora, V. K.; Kairalapova, A.; Sommerfeld, T.; Jordan, K. D. Theoretical Approaches for Treating Non-Valence Correlation-Bound Anions. *J. Chem. Phys.* **2017**, *147*, 214114, DOI: 10.1063/1.4991497 (pages 2, 20, 28, 30, 32).

[15] Hendricks, J. H.; Lyapustina, S. A.; de Clercq, H. L.; Bowen, K. H. The Dipole Bound-to-Covalent Anion Transformation in Uracil. *J. Chem. Phys.* **1998**, *108*, 8–11, DOI: 10.1063/1.475360 (pages 2, 20).

[16] Compton, R. N.; Carman, H. S.; Desfrançois, C.; Abdoul-Carime, H.; Schermann, J. P.; Hendricks, J. H.; Lyapustina, S. A.; Bowen, K. H. On the Binding of Electrons to Nitromethane: Dipole and Valence Bound Anions. *J. Chem. Phys.* **1996**, *105*, 3472–3478, DOI: 10.1063/1.472993 (pages 2, 20).

[17] Desfrançois, C.; Abdoul-Carime, H.; Schermann, J. P. Electron Attachment to Isolated Nucleic Acid Bases. *J. Chem. Phys.* **1996**, *104*, 7792–7794, DOI: 10.1063/1.471484 (pages 2, 20).

[18] Desfrançois, C.; Periquet, V.; Bouteiller, Y.; Schermann, J. P. Valence and Dipole Binding of Electrons to Uracil. *J. Phys. Chem. A* **1998**, *102*, 1274–1278, DOI: 10.1021/jp9728417 (pages 2, 20).

[19] Voora, V. K.; Jordan, K. D. Nonvalence Correlation-Bound Anion State of C_6F_6: Doorway to Low-Energy Electron Capture. *J. Phys. Chem. A* **2014**, *118*, 7201–7205, DOI: 10.1021/jp408386f (pages 2, 30).

[20] Dirac, P. A. M. The Quantum Theory of the Electron. *Proc. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character* **1997**, *117*, 610–624, DOI: 10.1098/rspa.1928.0023 (page 5).

[21] Anderson, C. D. The Apparent Existence of Easily Deflectable Positives. *Science* **1932**, *76*, 238–239, DOI: 10.1126/science.76.1967.238 (page 5).

[22] Muehllehner, G.; Karp, J. S. Positron Emission Tomography. *Phys. Med. Biol.* **2006**, *51*, R117, DOI: 10.1088/0031-9155/51/13/R08 (page 5).

[23] Vaquero, J. J.; Kinahan, P. Positron Emission Tomography: Current Challenges and Opportunities for Technological Advances in Clinical and Preclinical Imaging Systems. *Annu. Rev. Biomed. Eng.* **2015**, *17*, 385–414, DOI: 10.1146/annurev-bioeng-071114-040723 (page 5).

[24] Siegel, R. W. Positron Annihilation Spectroscopy. *Annu. Rev. Mater. Sci.* **1980**, *10*, 393–425, DOI: 10.1146/annurev.ms.10.080180.002141 (page 5).

[25] Selim, F. A. Positron Annihilation Spectroscopy of Defects in Nuclear and Irradiated Materials- a Review. *Materials Characterization* **2021**, *174*, 110952, DOI: 10.1016/j.matchar.2021.110952 (page 5).

[26] Hori, M.; Aghai-Khozani, H.; Sótér, A.; Barna, D.; Dax, A.; Hayano, R.; Kobayashi, T.; Murakami, Y.; Todoroki, K.; Yamada, H.; Horváth, D.; Venturelli, L. Buffer-Gas Cooling of Antiprotonic Helium to 1.5 to 1.7 K, and Antiproton-to–Electron Mass Ratio. *Science* **2016**, *354*, 610–614, DOI: 10.1126/science.aaf6702 (page 5).

[27] Smorra, C.; Sellner, S.; Borchert, M. J.; Harrington, J. A.; Higuchi, T.; Nagahama, H.; Tanaka, T.; Mooser, A.; Schneider, G.; Bohman, M.; Blaum, K.; Matsuda, Y.; Ospelkaus, C.; Quint, W.; Walz, J.; Yamazaki, Y.; Ulmer, S. A Parts-per-Billion Measurement of the Antiproton Magnetic Moment. *Nature* **2017**, *550*, 371–374, DOI: 10.1038/nature24048 (page 5).

[28] DiSciacca, J.; Marshall, M.; Marable, K.; Gabrielse, G. Resolving an Individual One-Proton Spin Flip to Determine a Proton Spin State. *Phys. Rev. Lett.* **2013**, *110*, 140406, DOI: 10.1103/PhysRevLett.110.140406 (page 5).

[29] Ulmer, S.; Smorra, C.; Mooser, A.; Franke, K.; Nagahama, H.; Schneider, G.; Higuchi, T.; Van Gorp, S.; Blaum, K.; Matsuda, Y.; Quint, W.; Walz, J.; Yamazaki, Y. High-Precision Comparison of the Antiproton-to-Proton Charge-to-Mass Ratio. *Nature* **2015**, *524*, 196–199, DOI: 10.1038/nature14861 (page 5).

[30] Gabrielse, G.; Khabbaz, A.; Hall, D. S.; Heimann, C.; Kalinowsky, H.; Jhe, W. Precision Mass Spectroscopy of the Antiproton and Proton Using Simultaneously Trapped Particles. *Phys. Rev. Lett.* **1999**, *82*, 3198–3201, DOI: 10.1103/PhysRevLett.82.3198 (page 5).

[31] Gribakin, G. F.; Young, J. A.; Surko, C. M. Positron-Molecule Interactions: Resonant Attachment, Annihilation, and Bound States. *Rev. Mod. Phys.* **2010**, *82*, 2557–2607, DOI: 10.1103/RevModPhys.82.2557 (pages 5, 6).

[32] Osmon, P. E. Positron Lifetime Spectra in Molecular Gases. *Phys. Rev.* **1965**, *140*, A8–A11, DOI: 10.1103/PhysRev.140.A8 (page 5).

[33] Osmon, P. E. Positron Lifetime Spectra in the Noble Gases. *Phys. Rev.* **1965**, *138*, B216–B218, DOI: 10.1103/PhysRev.138.B216 (page 5).

[34] Charlton, M.; Werf, D. P. van der; Lewis, R. J.; Watkeys, P. R.; Kerrigan, S. J. Three-Body Effects in Positron Annihilation on Molecules. *J. Phys. B: At. Mol. Opt. Phys.* **2006**, *39*, L329, DOI: 10.1088/0953-4075/39/17/L03 (page 5).

[35] Tao, S. J. Resonance Annihilation of Positrons in Chlorine and Argon. *Phys. Rev. Lett.* **1965**, *14*, 935–936, DOI: 10.1103/PhysRevLett.14.935 (page 5).

[36] Young, J. A.; Gribakin, G. F.; Lee, C. M. R.; Surko, C. M. Role of Combination Vibrations in Resonant Positron Annihilation. *Phys. Rev. A* **2008**, *77*, 060702, DOI: 10.1103/PhysRevA.77.060702 (page 5).

[37] Young, J. A.; Surko, C. M. Role of Binding Energy in Feshbach-Resonant Positron-Molecule Annihilation. *Phys. Rev. Lett.* **2007**, *99*, 133201, DOI: 10.1103/PhysRevLett.99.133201 (pages 5, 6).

[38] Gilbert, S. J.; Barnes, L. D.; Sullivan, J. P.; Surko, C. M. Vibrational-Resonance Enhancement of Positron Annihilation in Molecules. *Phys. Rev. Lett.* **2002**, *88*, 043201, DOI: 10.1103/PhysRevLett.88.043201 (page 6).

[39] Barnes, L. D.; Gilbert, S. J.; Surko, C. M. Energy-Resolved Positron Annihilation for Molecules. *Phys. Rev. A* **2003**, *67*, 032706, DOI: 10.1103/PhysRevA.67.032706 (page 6).

[40] Coulson, C. A.; Robinson, P. D. Wave Functions for the Hydrogen Atom in Spheroidal Coordinates I: The Derivation and Properties of the Functions. *Proc. Phys. Soc.* **1958**, *71*, 815, DOI: 10.1088/0370-1328/71/5/312 (page 7).

[41] Jones, R. O. Density Functional Theory: Its Origins, Rise to Prominence, and Future. *Rev. Mod. Phys.* **2015**, *87*, 897–923, DOI: 10.1103/RevModPhys.87.897 (page 10).

[42] Burke, K.; Wagner, L. O. DFT in a Nutshell. *Int. J. Quantum Chem.* **2013**, *113*, 96–101, DOI: 10.1002/qua.24259 (page 10).

[43] Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871, DOI: 10.1103/PhysRev.136.B864 (page 10).

[44] Mermin, N. D. Thermal Properties of the Inhomogeneous Electron Gas. *Phys. Rev.* **1965**, *137*, A1441–A1443, DOI: 10.1103/PhysRev.137.A1441 (page 11).

[45] Vignale, G.; Rasolt, M. Density-Functional Theory in Strong Magnetic Fields. *Phys. Rev. Lett.* **1987**, *59*, 2360–2363, DOI: 10.1103/PhysRevLett.59.2360 (page 11).

[46] Levy, M. Universal Variational Functionals of Electron Densities, First-Order Density Matrices, and Natural Spin-Orbitals and Solution of the v-Representability Problem. *Proc. Natl. Acad. Sci.* **1979**, *76*, 6062–6065, DOI: 10.1073/pnas.76.12.6062 (page 11).

[47] Miller, K. J.; Ruedenberg, K. Electron Correlation and Augmented Separated-Pair Expansion in Berylliumlike Atomic Systems. *J. Chem. Phys.* **1968**, *48*, 3450–3464, DOI: 10.1063/1.1669638 (page 11).

[48] Sims, J. S.; Hagstrom, S. Combined Configuration-Interaction Hylleraas-Type Wave-Function Study of the Ground State of the Beryllium Atom. *Phys. Rev. A* **1971**, *4*, 908–916, DOI: 10.1103/PhysRevA.4.908 (page 11).

[49] Bunge, C. F. Accurate Determination of the Total Electronic Energy of the Be Ground State. *Phys. Rev. A* **1976**, *14*, 1965–1978, DOI: 10.1103/PhysRevA.14.1965 (page 11).

[50] Komasa, J.; Rychlewski, J.; Jankowski, K. Benchmark Energy Calculations on Be-like Atoms. *Phys. Rev. A* **2002**, *65*, 042507, DOI: 10.1103/PhysRevA.65.042507 (page 11).

[51] Valderrama, E.; Ludeña, E. V.; Hinze, J. Assessment of Dynamical and Nondynamical Correlation Energy Components for the Beryllium-Atom Isoelectronic Sequence. *J. Chem. Phys.* **1999**, *110*, 2343–2353, DOI: 10.1063/1.477970 (page 11).

[52] Kato, T. On the Eigenfunctions of Many-Particle Systems in Quantum Mechanics. *Commun. Pure Appl. Math.* **1957**, *10*, 151–177, DOI: 10.1002/cpa.3160100201 (page 11).

[53] Allen, T. L.; Shull, H. Electron Pairs in the Beryllium Atom. *J. Phys. Chem.* **1962**, *66*, 2281–2283, DOI: 10.1021/j100818a001 (page 13).

[54] Smith, D. W.; Fogel, S. J. Natural Orbitals and Geminals of the Beryllium Atom. *J. Chem. Phys.* **1965**, *43*, S91–S96, DOI: 10.1063/1.1701519 (page 13).

[55] Weinhold, F.; Wilson, E. B. Reduced Density Matrices of Atoms and Molecules. I. The 2 Matrix of Double-Occupancy, Configuration-Interaction Wavefunctions for Singlet States. *J. Chem. Phys.* **1967**, *46*, 2752–2758, DOI: 10.1063/1.1841109 (page 13).

[56] Couty, M.; Hall, M. B. Generalized Molecular Orbital Theory II. *J. Phys. Chem. A* **1997**, *101*, 6936–6944, DOI: 10.1021/jp9639531 (page 13).

[57] Vu, N.; Mitxelena, I.; DePrince, A. E. An Adiabatic Connection for Doubly-Occupied Configuration Interaction Wave Functions. *J. Chem. Phys.* **2019**, *151*, 244121, DOI: 10.1063/1.5130660 (page 13).

[58] Henderson, T. M.; Bulik, I. W.; Stein, T.; Scuseria, G. E. Seniority-Based Coupled Cluster Theory. *J. Chem. Phys.* **2014**, *141*, 244104, DOI: 10.1063/1.4904384 (page 13).

[59] Limacher, P. A.; Ayers, P. W.; Johnson, P. A.; De Baerdemacker, S.; Van Neck, D.; Bultinck, P. A New Mean-Field Method Suitable for Strongly Correlated Electrons: Computationally Facile Antisymmetric Products of Nonorthogonal Geminals. *J. Chem. Theory Comput.* **2013**, *9*, 1394–1401, DOI: 10.1021/ct300902c (page 13).

[60] Limacher, P. A.; Kim, T. D.; Ayers, P. W.; Johnson, P. A.; De Baerdemacker, S.; Van Neck, D.; Bultinck, P. The Influence of Orbital Rotation on the Energy of Closed-Shell Wavefunctions. *Mol. Phys.* **2014**, *112*, 853–862, DOI: 10.1080/00268976.2013.874600 (page 13).

[61] Tecmer, P.; Boguslawski, K.; Johnson, P. A.; Limacher, P. A.; Chan, M.; Verstraelen, T.; Ayers, P. W. Assessing the Accuracy of New Geminal-Based Approaches. *J. Phys. Chem. A* **2014**, *118*, 9058–9068, DOI: 10.1021/jp502127v (page 13).

[62] Stein, T.; Henderson, T. M.; Scuseria, G. E. Seniority Zero Pair Coupled Cluster Doubles Theory. *J. Chem. Phys.* **2014**, *140*, 214113, DOI: 10.1063/1.4880819 (page 13).

[63] Kossoski, F.; Damour, Y.; Loos, P.-F. Hierarchy Configuration Interaction: Combining Seniority Number and Excitation Degree. *J. Phys. Chem. Lett.* **2022**, *13*, 4342–4349, DOI: 10.1021/acs.jpclett.2c00730 (page 13).

[64] 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 13, 34).

[65] Bartlett, R. J. Coupled-Cluster Theory and Its Equation-of-Motion Extensions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 126–138, DOI: 10.1002/wcms.76 (page 14).

[66] Bartlett, R. J. Many-Body Perturbation Theory and Coupled Cluster Theory for Electron Correlation in Molecules. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359–401, DOI: 10.1146/annurev.pc.32.100181.002043 (page 14).

[67] Comeau, D. C.; Bartlett, R. J. The Equation-of-Motion Coupled-Cluster Method. Applications to Open- and Closed-Shell Reference States. *Chem. Phys. Lett.* **1993**, *207*, 414–423, DOI: 10.1016/0009-2614(93)89023-B (pages 14, 30, 32).

[68] Stanton, J. F.; Bartlett, R. J. The Equation of Motion Coupled-cluster Method. A Systematic Biorthogonal Approach to Molecular Excitation Energies, Transition Probabilities, and Excited State Properties. *J. Chem. Phys.* **1993**, *98*, 7029–7039, DOI: 10.1063/1.464746 (page 14).

[69] 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 (pages 14, 20).

[70] Sinha, D.; Mukhopadhyay, S. K.; Chaudhuri, R.; Mukherjee, D. The Eigenvalue-Independent Partitioning Technique in Fock Space: An Alternative Route to Open-Shell Coupled-Cluster Theory for Incomplete Model Spaces. *Chem. Phys. Lett.* **1989**, *154*, 544–549, DOI: 10.1016/0009-2614(89)87149-0 (page 14).

[71] Stanton, J. F.; Gauss, J. Analytic Energy Derivatives for Ionized States Described by the Equation-of-motion Coupled Cluster Method. *J. Chem. Phys.* **1994**, *101*, 8938–8944, DOI: 10.1063/1.468022 (page 14).

[72] 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 (pages 14, 32).

[73] Trotter, H. F. On the Product of Semi-Groups of Operators. *Proc. Amer. Math. Soc.* **1959**, *10*, 545–551, DOI: 10.1090/S0002-9939-1959-0108732-6 (page 15).

[74] Suzuki, M. Generalized Trotter's Formula and Systematic Approximants of Exponential Operators and Inner Derivations with Applications to Many-Body Problems. *Commun.Math. Phys.* **1976**, *51*, 183–190, DOI: 10.1007/BF01609348 (page 15).

[75] Zhang, S.; Carlson, J.; Gubernatis, J. E. Constrained Path Monte Carlo Method for Fermion Ground States. *Phys. Rev. B* **1997**, *55*, 7464–7477, DOI: 10.1103/PhysRevB.55.7464 (pages 15, 18, 31, 38).

[76] Zhang, S.; Carlson, J.; Gubernatis, J. E. Constrained Path Quantum Monte Carlo Method for Fermion Ground States. *Phys. Rev. Lett.* **1995**, *74*, 3652–3655, DOI: 10.1103/PhysRevLett.74.3652 (pages 15, 18).

[77] Zhang, S.; Krakauer, H. Quantum Monte Carlo Method Using Phase-Free Random Walks with Slater Determinants. *Phys. Rev. Lett.* **2003**, *90*, 136401, DOI: 10.1103/PhysRevLett.90.136401 (pages 15, 18, 31, 38).

[78] Toulouse, J.; Assaraf, R.; Umrigar, C. J., Chapter Fifteen - Introduction to the Variational and Diffusion Monte Carlo Methods In *Advances in Quantum Chemistry*, Hoggan, P. E., Ozdogan, T., Eds.; Electron Correlation in Molecules – Ab Initio

Beyond Gaussian Quantum Chemistry, Vol. 73; Academic Press: 2016, pp 285–314, DOI: 10.1016/bs.aiq.2015.07.003 (pages 16, 21).

[79] 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 16, 20, 21, 31).

[80] Hubbard, J. Calculation of Partition Functions. *Phys. Rev. Lett.* **1959**, *3*, 77–78, DOI: 10.1103/PhysRevLett.3.77 (pages 17, 38).

[81] Stratonovich, R. L. On a Method of Calculating Quantum Distribution Functions. *Sov. Phys. Dokl.* **1957**, *2*, 416 (page 17).

[82] Thomas, I. L. Protonic Structure of Molecules. I. Ammonia Molecules. *Phys. Rev.* **1969**, *185*, 90–94, DOI: 10.1103/PhysRev.185.90 (pages 18, 63).

[83] Thomas, I. L. The Protonic Structure of Methane, Ammonia, Water, and Hydrogen Fluoride. *Chem. Phys. Lett.* **1969**, *3*, 705–706, DOI: 10.1016/0009-2614(69)87015-6 (pages 18, 63).

[84] Thomas, I. L.; Joy, H. W. Protonic Structure of Molecules. II. Methodology, Center-of-Mass Transformation, and the Structure of Methane, Ammonia, and Water. *Phys. Rev. A* **1970**, *2*, 1200–1208, DOI: 10.1103/PhysRevA.2.1200 (pages 18, 63).

[85] Thomas, I. L. "Vibrational" and "Rotational" Energy Levels as Protonic Structure in Molecules. *Phys. Rev. A* **1971**, *3*, 565–567, DOI: 10.1103/PhysRevA.3.565 (pages 18, 63).

[86] Kurtz, H. A.; Jordan, K. D. Ab Initio Study of the Positron Affinity of LiH. *J. Phys. B: Atom. Mol. Phys.* **1978**, *11*, L479, DOI: 10.1088/0022-3700/11/16/001 (pages 18, 64, 65).

[87] Kurtz, H. A.; Jordan, K. D. A Comparison of the Positron Energy Levels of (F^-; E+) and (CN^-; E+) with the Rydberg Levels of Na and NO. *J. Phys. B: Atom. Mol. Phys.* **1979**, *12*, L473, DOI: 10.1088/0022-3700/12/15/007 (pages 18, 64, 65).

[88] Kurtz, H. A.; Jordan, K. D. Theoretical Studies of Positron Complexes with Atomic Anions. *J. Chem. Phys.* **1980**, *72*, 493–503, DOI: 10.1063/1.438933 (pages 18, 64, 65).

[89] Kurtz, H. A.; Jordan, K. D. Theoretical Studies of Positron–Molecule Complexes. *J. Chem. Phys.* **1981**, *75*, 1876–1887, DOI: 10.1063/1.442211 (pages 18, 64, 65).

[90] Kurtz, H. A.; Jordan, K. D. Theoretical Study of Low-Energy Electron and Positron Scattering on Be, Mg and Ca. *J. Phys. B: Atom. Mol. Phys.* **1981**, *14*, 4361, DOI: 10.1088/0022-3700/14/22/019 (pages 18, 64, 65).

[91] Bishop, D. M. Non-Adiabatic Calculations for H_2^+, HD^+ and D_2^+. *Mol. Phys.* **1974**, *28*, 1397–1408, DOI: 10.1080/00268977400102681 (pages 18, 63).

[92] Bishop, D. M.; Cheung, L. M. Calculation of Transition Frequencies for H_2^+ and Its Isotopes to Spectroscopic Accuracy. *Phys. Rev. A* **1977**, *16*, 640–645, DOI: 10.1103/PhysRevA.16.640 (pages 18, 63).

[93] Pettitt, B. A. Hartree-Fock Theory of Proton States in Hydrides. *Chem. Phys. Lett.* **1986**, *130*, 399–402, DOI: 10.1016/0009-2614(86)80493-6 (pages 18, 63).

[94] Monkhorst, H. J. Chemical Physics without the Born-Oppenheimer Approximation: The Molecular Coupled-Cluster Method. *Phys. Rev. A* **1987**, *36*, 1544–1561, DOI: 10.1103/PhysRevA.36.1544 (pages 18, 63).

[95] Kozlowski, P. M.; Adamowicz, L. An Effective Method for Generating Nonadiabatic Many-body Wave Function Using Explicitly Correlated Gaussian-type Functions. *J. Chem. Phys.* **1991**, *95*, 6681–6698, DOI: 10.1063/1.461538 (pages 18, 63).

[96] Kozlowski, P. M.; Adamowicz, L. Implementation of Analytical First Derivatives for Evaluation of the Many-body Nonadiabatic Wave Function with Explicitly Correlated Gaussian Functions. *J. Chem. Phys.* **1992**, *96*, 9013–9024, DOI: 10.1063/1.462259 (pages 18, 63).

[97] Kozlowski, P. M.; Adamowicz, L. Newton–Raphson Optimization of the Many-body Nonadiabatic Wave Function Expressed in Terms of Explicitly Correlated Gaussian Functions. *J. Chem. Phys.* **1992**, *97*, 5063–5073, DOI: 10.1063/1.463827 (pages 18, 63).

[98] Kozlowski, P. M.; Adamowicz, Ludwik. Equivalent Quantum Approach to Nuclei and Electrons in Molecules. *Chem. Rev.* **1993**, *93*, 2007–2022, DOI: 10.1021/cr00022a003 (pages 18, 63).

[99] Kozlowski, P. M.; Adamowicz, L. Effctive Nonadiabatic Calculations on the Ground State of the HD^+ Molecule. *Int. J. Quantum Chem.* **1995**, *55*, 245–250, DOI: 10.1002/qua.560550305 (pages 18, 63).

[100] Kinghorn, D. B.; Adamowicz, L. Improved Nonadiabatic Ground-State Energy Upper Bound for Dihydrogen. *Phys. Rev. Lett.* **1999**, *83*, 2541–2543, DOI: 10.1103/PhysRevLett.83.2541 (pages 18, 63).

[101] Kinghorn, D. B.; Adamowicz, L. High Accuracy Non-Born–Oppenheimer Calculations for the Isotopomers of the Hydrogen Molecule with Explicitly Correlated Gaussian Functions. *J. Chem. Phys.* **2000**, *113*, 4203–4205, DOI: 10.1063/1.1288376 (pages 18, 63).

[102] Scheu, C. E.; Kinghorn, D. B.; Adamowicz, L. Non-Born–Oppenheimer Calculations on the LiH Molecule with Explicitly Correlated Gaussian Functions. *J. Chem. Phys.* **2001**, *114*, 3393–3397, DOI: 10.1063/1.1342757 (pages 18, 63).

[103] Cafiero, M.; Adamowicz, L. Nonadiabatic Calculations of the Dipole Moments of LiH and LiD. *Phys. Rev. Lett.* **2002**, *88*, 033002, DOI: 10.1103/PhysRevLett.88.033002 (pages 18, 63).

[104] Cafiero, M.; Adamowicz, L. Non-Born–Oppenheimer Calculations of the Polarizability of LiH in a Basis of Explicitly Correlated Gaussian Functions. *J. Chem. Phys.* **2002**, *116*, 5557–5564, DOI: 10.1063/1.1457435 (pages 18, 63).

[105] Cafiero, M.; Adamowicz, L. Non-Born-Oppenheimer Isotope Effects on the Polarizabilities of H_2. *Phys. Rev. Lett.* **2002**, *89*, 073001, DOI: 10.1103/PhysRevLett.89.073001 (pages 18, 63).

[106] Cafiero, M.; Bubin, S.; Adamowicz, L. Non-Born–Oppenheimer Calculations of Atoms and Molecules. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1491–1501, DOI: 10.1039/B211193D (pages 18, 63).

[107] Bubin, S.; Adamowicz, L. Variational Calculations of Excited States with Zero Total Angular Momentum (Vibrational Spectrum) of H2 without Use of the Born–Oppenheimer Approximation. *J. Chem. Phys.* **2003**, *118*, 3079–3082, DOI: 10.1063/1.1537719 (pages 18, 63).

[108] Bubin, S.; Adamowicz, L. Nonrelativistic Molecular Quantum Mechanics without Approximations: Electron Affinities of LiH and LiD. *J. Chem. Phys.* **2004**, *121*, 6249–6253, DOI: 10.1063/1.1786580 (pages 18, 63).

[109] Bednarz, E.; Bubin, S.; Adamowicz, L. Non-Born–Oppenheimer Variational Calculations of HT^+ Bound States with Zero Angular Momentum. *J. Chem. Phys.* **2005**, *122*, 164302, DOI: 10.1063/1.1884602 (pages 18, 63).

[110] Cafiero, M.; Adamowicz, L. Non-Born–Oppenheimer Molecular Structure and One-Particle Densities for H_2D^+. *J. Chem. Phys.* **2005**, *122*, 184305, DOI: 10.1063/1.1891707 (pages 18, 63).

[111] Pavanello, M.; Bubin, S.; Molski, M.; Adamowicz, L. Non-Born–Oppenheimer Calculations of the Pure Vibrational Spectrum of HeH^+. *J. Chem. Phys.* **2005**, *123*, 104306, DOI: 10.1063/1.2012332 (pages 18, 63).

[112] Bubin, S.; Adamowicz, L.; Molski, M. An Accurate Non-Born–Oppenheimer Calculation of the First Purely Vibrational Transition in LiH Molecule. *J. Chem. Phys.* **2005**, *123*, 134310, DOI: 10.1063/1.2047487 (pages 18, 63).

[113] Kędziera, D.; Stanke, M.; Bubin, S.; Barysz, M.; Adamowicz, L. Darwin and Mass-Velocity Relativistic Corrections in the Non-Born-Oppenheimer Calculations of Pure Vibrational States of H_2. *J. Chem. Phys.* **2006**, *125*, 014318, DOI: 10.1063/1.2209691 (pages 18, 63).

[114] Bubin, S.; Adamowicz, L. Non-Born-Oppenheimer Variational Calculation of the Ground-State Vibrational Spectrum of LiH^+. *J. Chem. Phys.* **2006**, *125*, 064309, DOI: 10.1063/1.2244563 (pages 18, 63).

[115] Kedziera, D.; Stanke, M.; Bubin, S.; Barysz, M.; Adamowicz, L. Darwin and Mass-Velocity Relativistic Corrections in Non-Born-Oppenheimer Variational Calculations. *J. Chem. Phys.* **2006**, *125*, 084303, DOI: 10.1063/1.2236113 (pages 18, 63).

[116] Stanke, M.; Kędziera, D.; Bubin, S.; Adamowicz, L. Relativistic Corrections to the Non-Born-Oppenheimer Energies of the Lowest Singlet Rydberg States of He_3 and He_4. *J. Chem. Phys.* **2007**, *126*, 194312, DOI: 10.1063/1.2735305 (pages 18, 63).

[117] Bubin, S.; Adamowicz, L. Calculations of the Ground States of BeH and BeH^+ without the Born-Oppenheimer Approximation. *J. Chem. Phys.* **2007**, *126*, 214305, DOI: 10.1063/1.2736699 (pages 18, 63).

[118] Stanke, M.; Kędziera, D.; Bubin, S.; Adamowicz, L. Electron Affinity of Li_7 Calculated with the Inclusion of Nuclear Motion and Relativistic Corrections. *J. Chem. Phys.* **2007**, *127*, 134107, DOI: 10.1063/1.2755767 (pages 18, 63).

[119] Stanke, M.; Kędziera, D.; Bubin, S.; Molski, M.; Adamowicz, L. Lowest Vibrational States of $^4He\ ^3He^+$: Non-Born-Oppenheimer Calculations. *Phys. Rev. A* **2007**, *76*, 052506, DOI: 10.1103/PhysRevA.76.052506 (pages 18, 63).

[120] Stanke, M.; Kędziera, D.; Bubin, S.; Adamowicz, L. Complete α^2 Relativistic Corrections to the Pure Vibrational Non-Born-Oppenheimer Energies of HeH^+. *Phys. Rev. A* **2008**, *77*, 022506, DOI: 10.1103/PhysRevA.77.022506 (pages 18, 63).

[121] Stanke, M.; Kędziera, D.; Bubin, S.; Molski, M.; Adamowicz, L. Orbit-Orbit Relativistic Corrections to the Pure Vibrational Non-Born-Oppenheimer Energies of H_2. *J. Chem. Phys.* **2008**, *128*, 114313, DOI: 10.1063/1.2834926 (pages 18, 63).

[122] Shigeta, Y.; Ozaki, Y.; Kodama, K.; Nagao, H.; Kawabe, H.; Nishikawa, K. Nonadiabatic Molecular Theory and Its Application. II. Water Molecule. *Int. J. Quantum Chem.* **1998**, *69*, 629–637, DOI: 10.1002/(SICI)1097-461X(1998)69:5<629::AID-QUA1>3.0.CO;2-X (pages 18, 63).

[123] Shigeta, Y.; Takahashi, H.; Yamanaka, S.; Mitani, M.; Nagao, H.; Yamaguchi, K. Density Functional Theory without the Born–Oppenheimer Approximation and Its Application. *Int. J. Quantum Chem.* **1998**, *70*, 659–669, DOI: 10.1002/(SICI)1097-461X(1998)70:4/5<659::AID-QUA12>3.0.CO;2-Y (pages 18, 63).

[124] Shigeta, Y.; Nagao, H.; Nishikawa, K.; Yamaguchi, K. A Formulation and Numerical Approach to Molecular Systems by the Green Function Method without the Born–Oppenheimer Approximation. *J. Chem. Phys.* **1999**, *111*, 6171–6179, DOI: 10.1063/1.479921 (pages 18, 63).

[125] Tachikawa, M.; Mori, K.; Nakai, H.; Iguchi, K. An Extension of Ab Initio Molecular Orbital Theory to Nuclear Motion. *Chem. Phys. Lett.* **1998**, *290*, 437–442, DOI: 10.1016/S0009-2614(98)00519-3 (pages 18, 63).

[126] Tachikawa, M.; Mori, K.; Osamura, Y. Isotope Effect of Hydrated Clusters of Hydrogen Chloride, HCl(H_2O)n and DCl(H_2O)n (n = 0–4): Application of Dynamic Extended Molecular Orbital Method. *Mol. Phys.* **1999**, *96*, 1207–1215, DOI: 10.1080/00268979909483065 (pages 18, 63).

[127] González, S. A.; Aguirre, N. F.; Reyes, A. Theoretical Investigation of Isotope Effects: The Any-Particle Molecular Orbital Code. *Int. J. Quantum Chem.* **2008**, *108*, 1742–1749, DOI: 10.1002/qua.21584 (pages 18, 63).

[128] Chakraborty, A.; Pak, M. V.; Hammes-Schiffer, S. Development of Electron-Proton Density Functionals for Multicomponent Density Functional Theory. *Phys. Rev. Lett.* **2008**, *101*, 153001, DOI: 10.1103/PhysRevLett.101.153001 (pages 18, 63).

[129] Chakraborty, A.; Pak, M. V.; Hammes-Schiffer, S. Inclusion of Explicit Electron-Proton Correlation in the Nuclear-Electronic Orbital Approach Using Gaussian-type Geminal Functions. *J. Chem. Phys.* **2008**, *129*, 014101, DOI: 10.1063/1.2943144 (pages 18, 63).

[130] Adamson, P. E.; Duan, X. F.; Burggraf, L. W.; Pak, M. V.; Swalina, C.; Hammes-Schiffer, S. Modeling Positrons in Molecular Electronic Structure Calculations with the Nuclear-Electronic Orbital Method. *J. Phys. Chem. A* **2008**, *112*, 1346–1351, DOI: 10.1021/jp7098015 (pages 18, 63, 64).

[131] Tachikawa, M. Simultaneous Optimization of Gaussian Type Function Exponents for Electron and Positron with Full-CI Wavefunction – Application to Ground and Excited States of Positronic Compounds with Multi-Component Molecular Orbital Approach. *Chem. Phys. Lett.* **2001**, *350*, 269–276, DOI: 10.1016/S0009-2614(01)01286-6 (pages 18, 63, 64).

[132] Tachikawa, M. Multi-Component Molecular Orbital Theory for Electrons and Nuclei Including Many-Body Effect with Full Configuration Interaction Treatment: Isotope Effects on Hydrogen Molecules. *Chem. Phys. Lett.* **2002**, *360*, 494–500, DOI: 10.1016/S0009-2614(02)00881-3 (pages 18, 63).

[133] Tachikawa, M.; Ishimoto, T.; Tokiwa, H.; Kasatani, H.; Deguchi, K. First-Principle Calculation on Isotope Effect in KH_2PO_4 and KD_2PO_4 of Hydrogen-Bonded Di-

electric Materials. Approach with Dynamic Extended Molecular Orbital Method. *Ferroelectrics* **2002**, *268*, 3–9, DOI: 10.1080/713715956 (pages 18, 63).

[134] Tachikawa, M. Multi-Component Molecular Orbital Study of Isotope Effects on Lithium Hydride Molecules with the Configuration Interaction Scheme. *Journal of Molecular Structure: THEOCHEM* **2003**, *630*, 75–79, DOI: 10.1016/S0166-1280(03)00147-7 (pages 18, 63).

[135] Tachikawa, M.; Buenker, R. J.; Kimura, M. Geometry Relaxation Effects for Molecules as a Result of Binding with a Positron. *J. Chem. Phys.* **2004**, *121*, 9191–9192, DOI: 10.1063/1.1805493 (pages 18, 63).

[136] Udagawa, T.; Ishimoto, T.; Tokiwa, H.; Tachikawa, M.; Nagashima, U. The Geometrical Isotope Effect of C–H$\cdots$O Type Hydrogen Bonds Revealed by Multi-Component Molecular Orbital Calculation. *Chem. Phys. Lett.* **2004**, *389*, 236–240, DOI: 10.1016/j.cplett.2004.03.091 (pages 18, 63).

[137] Ishimoto, T.; Tachikawa, M.; Tokiwa, H.; Nagashima, U. Isotope Effect on Hydrogen (Deuterium)-Absorbing Pt Clusters Calculated by the Multi-Component Molecular Orbital Method. *J. Phys. Soc. Jpn.* **2005**, *74*, 3112–3116, DOI: 10.1143/JPSJ.74.3112 (pages 18, 63).

[138] Ishimoto, T.; Tachikawa, M.; Tokiwa, H.; Nagashima, U. Kinetic and Geometrical Isotope Effects in Hydrogen-Atom Transfer Reaction, as Calculated by the Multi-Component Molecular Orbital Method. *Chemical Physics* **2005**, *314*, 231–237, DOI: 10.1016/j.chemphys.2005.03.007 (pages 18, 63).

[139] Shibl, M. F.; Tachikawa, M.; Kühn, O. The Geometric (H/D) Isotope Effect in Porphycene: Grid-Based Born–Oppenheimer Vibrational Wavefunctions vs. Multi-Component Molecular Orbital Theory. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1368–1373, DOI: 10.1039/B500620A (pages 18, 63).

[140] Ishimoto, T.; Tachikawa, M.; Nagashima, U. A Fragment Molecular-Orbital– Multicomponent Molecular-Orbital Method for Analyzing H/D Isotope Effects in Large Molecules. *J. Chem. Phys.* **2006**, *124*, 014112, DOI: 10.1063/1.2151897 (pages 18, 63).

[141] Udagawa, T.; Ishimoto, T.; Tokiwa, H.; Tachikawa, M.; Nagashima, U. Geometric Isotope Effect of Various Intermolecular and Intramolecular C-H· · ·O Hydrogen Bonds, Using the Multicomponent Molecular Orbital Method. *J. Phys. Chem. A* **2006**, *110*, 7279–7285, DOI: 10.1021/jp0615656 (pages 18, 63).

[142] Ishimoto, T.; Tachikawa, M.; Nagashima, U. Electron-Electron and Electron-Nucleus Correlation Effects on Exponent Values of Gaussian-type Functions for Quantum Protons and Deuterons. *J. Chem. Phys.* **2006**, *125*, 144103, DOI: 10.1063/1.2352753 (pages 18, 63).

[143] Udagawa, T.; Tachikawa, M. H/D Isotope Effect on Porphine and Porphycene Molecules with Multicomponent Hybrid Density Functional Theory. *J. Chem. Phys.* **2006**, *125*, 244105, DOI: 10.1063/1.2403857 (pages 18, 63).

[144] Tachikawa, M. The First-Principles Multi-Component Molecular Orbital Approach to Bound States of Positron with the 2-Deoxyglucose Molecule as a Reagent of Positron Emission Tomography. *J. Phys.: Condens. Matter* **2007**, *19*, 365235, DOI: 10.1088/0953-8984/19/36/365235 (pages 18, 63).

[145] Ishimoto, T.; Tachikawa, M.; Nagashima, U. Analytical Optimization of Exponent Values in Protonic and Deuteronic Gaussian-type Functions by Elimination of Translational and Rotational Motions from Multi-Component Molecular Orbital Scheme. *Int. J. Quantum Chem.* **2008**, *108*, 472–481, DOI: 10.1002/qua.21540 (pages 18, 63).

[146] Kikuta, Y.; Ishimoto, T.; Nagashima, U. Deuterium-Substituted Water–Ammonia Mixed Trimer Clusters, (H2O)N-3(NH3)n (N=0,1,2,3): Interaction Energy, Hydrogen Bond Structures, and Mulliken Population. *Chemical Physics* **2008**, *354*, 218–224, DOI: 10.1016/j.chemphys.2008.10.027 (pages 18, 63).

[147] Ishimoto, T.; Ishihara, Y.; Teramae, H.; Baba, M.; Nagashima, U. H/D Isotope Effect of Methyl Internal Rotation for Acetaldehyde in Ground State as Calculated from a Multicomponent Molecular Orbital Method. *J. Chem. Phys.* **2008**, *128*, 184309, DOI: 10.1063/1.2917149 (pages 18, 63).

[148] Ishimoto, T.; Ishihara, Y.; Teramae, H.; Baba, M.; Nagashima, U. H/D Isotope Effect in Methyl Torsional Interaction of Acetone as Calculated by a Multicomponent

Molecular Orbital Method. *J. Chem. Phys.* **2008**, *129*, 214116, DOI: 10.1063/1.3028540 (pages 18, 63).

[149] Nakai, H. Simultaneous Determination of Nuclear and Electronic Wave Functions without Born–Oppenheimer Approximation: Ab Initio NO+MO/HF Theory. *Int. J. Quantum Chem.* **2002**, *86*, 511–517, DOI: 10.1002/qua.1106 (pages 18, 63).

[150] Nakai, H.; Sodeyama, K. Many-Body Effects in Nonadiabatic Molecular Theory for Simultaneous Determination of Nuclear and Electronic Wave Functions: Ab Initio NOMO/MBPT and CC Methods. *J. Chem. Phys.* **2003**, *118*, 1119–1127, DOI: 10.1063/1.1528951 (pages 18, 63).

[151] Nakai, H.; Hoshino, M.; Miyamoto, K.; Hyodo, S. Elimination of Translational and Rotational Motions in Nuclear Orbital plus Molecular Orbital Theory. *J. Chem. Phys.* **2005**, *122*, 164101, DOI: 10.1063/1.1871914 (pages 18, 63).

[152] Sodeyama, K.; Miyamoto, K.; Nakai, H. Non-Born–Oppenheimer Effects Predicted by Translation-Free Nuclear Orbital plus Molecular Orbital Method. *Chem. Phys. Lett.* **2006**, *421*, 72–76, DOI: 10.1016/j.cplett.2006.01.064 (pages 18, 63).

[153] Hoshino, M.; Nakai, H. Elimination of Translational and Rotational Motions in Nuclear Orbital plus Molecular Orbital Theory: Application of Møller-Plesset Perturbation Theory. *J. Chem. Phys.* **2006**, *124*, 194110, DOI: 10.1063/1.2193513 (pages 18, 63).

[154] Miyamoto, K.; Hoshino, M.; Nakai, H. Elimination of Translational and Rotational Motions in Nuclear Orbital Plus Molecular Orbital Theory: Contribution of the First-Order Rovibration Coupling. *J. Chem. Theory Comput.* **2006**, *2*, 1544–1550, DOI: 10.1021/ct6002065 (pages 18, 63).

[155] Hoshino, M.; Tsukamoto, Y.; Nakai, H. Development of Analytic Energy Gradient Method in Nuclear Orbital plus Molecular Orbital Theory. *Int. J. Quantum Chem.* **2007**, *107*, 2575–2585, DOI: 10.1002/qua.21430 (pages 18, 63).

[156] Nakai, H.; Ikabata, Y.; Tsukamoto, Y.; Imamura, Y.; Miyamoto, K.; Hoshino, M. Isotope Effect in Dihydrogen-Bonded Systems: Application of the Analytical Energy Gradient Method in the Nuclear Orbital plus Molecular Orbital Theory. *Mol. Phys.* **2007**, *105*, 2649–2657, DOI: 10.1080/00268970701618416 (pages 18, 63).

[157] Imamura, Y.; Kiryu, H.; Nakai, H. Colle-Salvetti-type Correction for Electron–Nucleus Correlation in the Nuclear Orbital plus Molecular Orbital Theory. *J. Comput. Chem.* **2008**, *29*, 735–740, DOI: 10.1002/jcc.20840 (pages 18, 63).

[158] Webb, S. P.; Iordanov, T.; Hammes-Schiffer, S. Multiconfigurational Nuclear - Electronic Orbital Approach: Incorporation of Nuclear Quantum Effects in Electronic Structure Calculations. *J. Chem. Phys.* **2002**, *117*, 4106, DOI: 10.1063/1.1494980 (pages 18, 63).

[159] Iordanov, T.; Hammes-Schiffer, S. Vibrational Analysis for the Nuclear–Electronic Orbital Method. *J. Chem. Phys.* **2003**, *118*, 9489–9496, DOI: 10.1063/1.1569913 (pages 18, 63).

[160] Pak, M. V.; Hammes-Schiffer, S. Electron-Proton Correlation for Hydrogen Tunneling Systems. *Phys. Rev. Lett.* **2004**, *92*, 103002, DOI: 10.1103/PhysRevLett.92.103002 (pages 18, 63).

[161] Pak, M. V.; Swalina, C.; Webb, S. P.; Hammes-Schiffer, S. Application of the Nuclear–Electronic Orbital Method to Hydrogen Transfer Systems: Multiple Centers and Multiconfigurational Wavefunctions. *Chemical Physics* **2004**, *304*, 227–236, DOI: 10.1016/j.chemphys.2004.06.009 (pages 18, 63).

[162] Swalina, C.; Pak, M. V.; Hammes-Schiffer, S. Alternative Formulation of Many-Body Perturbation Theory for Electron–Proton Correlation. *Chem. Phys. Lett.* **2005**, *404*, 394–399, DOI: 10.1016/j.cplett.2005.01.115 (pages 18, 63).

[163] Swalina, C.; Pak, M. V.; Hammes-Schiffer, S. Analysis of the Nuclear-Electronic Orbital Method for Model Hydrogen Transfer Systems. *J. Chem. Phys.* **2005**, *123*, 014303, DOI: 10.1063/1.1940634 (pages 18, 63).

[164] Reyes, A.; Pak, M. V.; Hammes-Schiffer, S. Investigation of Isotope Effects with the Nuclear-Electronic Orbital Approach. *J. Chem. Phys.* **2005**, *123*, 064104, DOI: 10.1063/1.1990116 (pages 18, 63).

[165] Skone, J. H.; Pak, M. V.; Hammes-Schiffer, S. Nuclear-Electronic Orbital Nonorthogonal Configuration Interaction Approach. *J. Chem. Phys.* **2005**, *123*, 134108, DOI: 10.1063/1.2039727 (pages 18, 63).

[166] Swalina, C.; Hammes-Schiffer, S. Impact of Nuclear Quantum Effects on the Molecular Structure of Bihalides and the Hydrogen Fluoride Dimer. *J. Phys. Chem. A* **2005**, *109*, 10410–10417, DOI: 10.1021/jp053552i (pages 18, 63).

[167] Swalina, C.; Pak, M. V.; Chakraborty, A.; Hammes-Schiffer, S. Explicit Dynamical Electron-Proton Correlation in the Nuclear-Electronic Orbital Framework. *J. Phys. Chem. A* **2006**, *110*, 9983–9987, DOI: 10.1021/jp0634297 (pages 18, 63).

[168] Ohta, Y.; Demura, A.; Okamoto, T.; Hitomi, H.; Nagaoka, M. The Body-Centered Cubic Structure of Methyllithium Tetramer Crystal: Staggered Methyl Conformation by Electrostatic Stabilization via Intratetramer Multipolarization. *J. Phys. Chem. B* **2006**, *110*, 12640–12644, DOI: 10.1021/jp057014h (pages 18, 63).

[169] Kumarasiri, M.; Swalina, C.; Hammes-Schiffer, S. Anharmonic Effects in Ammonium Nitrate and Hydroxylammonium Nitrate Clusters. *J. Phys. Chem. B* **2007**, *111*, 4653–4658, DOI: 10.1021/jp065569m (pages 18, 63).

[170] Swalina, C.; Wang, Q.; Chakraborty, A.; Hammes-Schiffer, S. Analysis of Nuclear Quantum Effects on Hydrogen Bonding. *J. Phys. Chem. A* **2007**, *111*, 2206–2212, DOI: 10.1021/jp0682661 (pages 18, 63).

[171] Pak, M. V.; Chakraborty, A.; Hammes-Schiffer, S. Density Functional Theory Treatment of Electron Correlation in the Nuclear-Electronic Orbital Approach. *J. Phys. Chem. A* **2007**, *111*, 4522–4526, DOI: 10.1021/jp0704463 (pages 18, 63).

[172] Chakraborty, A.; Pak, M. V.; Hammes-Schiffer, S. Properties of the Exact Universal Functional in Multicomponent Density Functional Theory. *J. Chem. Phys.* **2009**, *131*, 124115, DOI: 10.1063/1.3236844 (pages 18, 63).

[173] Sirjoosingh, A.; Pak, M. V.; Hammes-Schiffer, S. Derivation of an Electron–Proton Correlation Functional for Multicomponent Density Functional Theory within the Nuclear–Electronic Orbital Approach. *J. Chem. Theory Comput.* **2011**, *7*, 2689–2693, DOI: 10.1021/ct200473r (pages 18, 63).

[174] Sirjoosingh, A.; Pak, M. V.; Hammes-Schiffer, S. Multicomponent Density Functional Theory Study of the Interplay between Electron-Electron and Electron-Proton Correlation. *J. Chem. Phys.* **2012**, *136*, 174114, DOI: 10.1063/1.4709609 (pages 18, 63).

[175] Yang, Y.; Brorsen, K. R.; Culpitt, T.; Pak, M. V.; Hammes-Schiffer, S. Development of a Practical Multicomponent Density Functional for Electron-Proton Correlation to Produce Accurate Proton Densities. *The Journal of Chemical Physics* **2017**, *147*, 114113, DOI: 10.1063/1.4996038 (pages 18, 63).

[176] Brorsen, K. R.; Yang, Y.; Hammes-Schiffer, S. Multicomponent Density Functional Theory: Impact of Nuclear Quantum Effects on Proton Affinities and Geometries. *J. Phys. Chem. Lett.* **2017**, *8*, 3488–3493, DOI: 10.1021/acs.jpclett.7b01442 (pages 18, 63).

[177] Brorsen, K. R.; Schneider, P. E.; Hammes-Schiffer, S. Alternative Forms and Transferability of Electron-Proton Correlation Functionals in Nuclear-Electronic Orbital Density Functional Theory. *J. Chem. Phys.* **2018**, *149*, 044110, DOI: 10.1063/1.5037945 (pages 18, 63).

[178] Yang, Y.; Culpitt, T.; Hammes-Schiffer, S. Multicomponent Time-Dependent Density Functional Theory: Proton and Electron Excitation Energies. *J. Phys. Chem. Lett.* **2018**, *9*, 1765–1770, DOI: 10.1021/acs.jpclett.8b00547 (pages 18, 63).

[179] Tao, Z.; Yang, Y.; Hammes-Schiffer, S. Multicomponent Density Functional Theory: Including the Density Gradient in the Electron-Proton Correlation Functional for Hydrogen and Deuterium. *J. Chem. Phys.* **2019**, *151*, 124102, DOI: 10.1063/1.5119124 (pages 18, 63).

[180] Culpitt, T.; Yang, Y.; Pavošević, F.; Tao, Z.; Hammes-Schiffer, S. Enhancing the Applicability of Multicomponent Time-Dependent Density Functional Theory. *J. Chem. Phys.* **2019**, *150*, 201101, DOI: 10.1063/1.5099093 (pages 18, 63).

[181] Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S. Multicomponent Coupled Cluster Singles and Doubles Theory within the Nuclear-Electronic Orbital Framework. *J. Chem. Theory Comput.* **2019**, *15*, 338–347, DOI: 10.1021/acs.jctc.8b01120 (pages 18, 63).

[182] Pavošević, F.; Hammes-Schiffer, S. Multicomponent Equation-of-Motion Coupled Cluster Singles and Doubles: Theory and Calculation of Excitation Energies for Positronium Hydride. *J. Chem. Phys.* **2019**, *150*, 161102, DOI: 10.1063/1.5094035 (pages 18, 63).

[183] Sirjoosingh, A.; Pak, M. V.; Brorsen, K. R.; Hammes-Schiffer, S. Quantum Treatment of Protons with the Reduced Explicitly Correlated Hartree-Fock Approach. *J. Chem. Phys.* **2015**, *142*, 214107, DOI: 10.1063/1.4921303 (pages 18, 63).

[184] Brorsen, K. R.; Sirjoosingh, A.; Pak, M. V.; Hammes-Schiffer, S. Nuclear-Electronic Orbital Reduced Explicitly Correlated Hartree-Fock Approach: Restricted Basis Sets and Open-Shell Systems. *J. Chem. Phys.* **2015**, *142*, 214108, DOI: 10.1063/1.4921304 (pages 18, 63).

[185] Sirjoosingh, A.; Pak, M. V.; Swalina, C.; Hammes-Schiffer, S. Reduced Explicitly Correlated Hartree-Fock Approach within the Nuclear-Electronic Orbital Framework: Theoretical Formulation. *J. Chem. Phys.* **2013**, *139*, 034102, DOI: 10.1063/1.4812257 (pages 18, 63).

[186] Fowler, D.; Brorsen, K. R. (T) Correction for Multicomponent Coupled-Cluster Theory for a Single Quantum Proton. *J. Chem. Theory Comput.* **2022**, *18*, 7298–7305, DOI: 10.1021/acs.jctc.2c00701 (pages 18, 63).

[187] Fajen, O. J.; Brorsen, K. R. Multicomponent MP4 and the Inclusion of Triple Excitations in Multicomponent Many-Body Methods. *J. Chem. Phys.* **2021**, *155*, 234108, DOI: 10.1063/5.0071423 (pages 18, 63).

[188] Alaal, N.; Brorsen, K. R. Multicomponent Heat-Bath Configuration Interaction with the Perturbative Correction for the Calculation of Protonic Excited States. *J. Chem. Phys.* **2021**, *155*, 234107, DOI: 10.1063/5.0076006 (pages 18, 63).

[189] Fajen, O. J.; Brorsen, K. R. Separation of Electron–Electron and Electron–Proton Correlation in Multicomponent Orbital-Optimized Perturbation Theory. *J. Chem. Phys.* **2020**, *152*, 194107, DOI: 10.1063/5.0006743 (pages 18, 63).

[190] Brorsen, K. R. Quantifying Multireference Character in Multicomponent Systems with Heat-Bath Configuration Interaction. *J. Chem. Theory Comput.* **2020**, *16*, 2379–2388, DOI: 10.1021/acs.jctc.9b01273 (pages 18, 63).

[191] Fajen, O. J.; Brorsen, K. R. Multicomponent CASSCF Revisited: Large Active Spaces Are Needed for Qualitatively Accurate Protonic Densities. *J. Chem. Theory Comput.* **2021**, *17*, 965–974, DOI: 10.1021/acs.jctc.0c01191 (pages 18, 63).

[192] Samsonova, I.; Tucker, G. B.; Alaal, N.; Brorsen, K. R. Hydrogen-Atom Electronic Basis Sets for Multicomponent Quantum Chemistry. *ACS Omega* **2023**, *8*, 5033–5041, DOI: 10.1021/acsomega.2c07782 (pages 18, 63).

[193] Pavošević, F.; Rousseau, B. J. G.; Hammes-Schiffer, S. Multicomponent Orbital-Optimized Perturbation Theory Methods: Approaching Coupled Cluster Accuracy at Lower Cost. *J. Phys. Chem. Lett.* **2020**, *11*, 1578–1583, DOI: 10.1021/acs.jpclett.0c00090 (pages 18, 63).

[194] Schneider, P. E.; Pavošević, F.; Hammes-Schiffer, S. Diagonal Born–Oppenheimer Corrections within the Nuclear–Electronic Orbital Framework. *J. Phys. Chem. Lett.* **2019**, *10*, 4639–4643, DOI: 10.1021/acs.jpclett.9b01803 (pages 18, 63).

[195] Pak, M. V.; Chakraborty, A.; Hammes-Schiffer, S. Calculation of the Positron Annihilation Rate in PsH with the Positronic Extension of the Explicitly Correlated Nuclear-Electronic Orbital Method. *J. Phys. Chem. A* **2009**, *113*, 4004–4008, DOI: 10.1021/jp810410y (pages 18, 63).

[196] Pedraza-González, L.; Charry, J.; Quintero, W.; Alí-Torres, J.; Reyes, A. Fast and Accurate Prediction of Proton Affinities: Revisiting the Extended Koopmans' Theorem for Protons. *Phys. Chem. Chem. Phys.* **2017**, *19*, 25324–25333, DOI: 10.1039/C7CP04936F (pages 18, 63).

[197] Charry, J.; Pedraza-González, L.; Reyes, A. On the Physical Interpretation of the Nuclear Molecular Orbital Energy. *J. Chem. Phys.* **2017**, *146*, 214103, DOI: 10.1063/1.4984098 (pages 18, 63).

[198] Kozlowski, P. M.; Adamowicz, L. Nonadiabatic Variational Calculations for the Ground State of the Positronium Molecule. *Phys. Rev. A* **1993**, *48*, 1903–1908, DOI: 10.1103/PhysRevA.48.1903 (pages 18, 64).

[199] Kozlowski, P. M.; Adamowicz, L. Lifetime of Positronium Molecule. Study with Boys' Explicitly Correlated Gaussians. *J. Phys. Chem.* **1996**, *100*, 6266–6271, DOI: 10.1021/jp9528166 (pages 18, 64).

[200] Tachikawa, M.; Mori, K.; Suzuki, K.; Iguchi, K. Full Variational Molecular Orbital Method: Application to the Positron-Molecule Complexes. *Int. J. Quantum Chem.*

1998, *70*, 491–501, DOI: 10.1002/(SICI)1097-461X(1998)70:3<491::AID-QUA5>3.0.CO;2-P (pages 18, 64).

[201] Ishida, M.; Tachikawa, M.; Tokiwa, H.; Mori, K.; Ishii, A. First Principles Calculation for Hydrogen/Positronium Adsorption on an Si(111) Surface Using the Dynamical Extended Molecular Orbital Method. *Surface Science* **1999**, *438*, 47–57, DOI: 10.1016/S0039-6028(99)00551-8 (pages 18, 64).

[202] Tachikawa, M. Isotope Effect and Cluster Size Dependence for Water and Hydrated Hydrogen Halide Clusters: Multi-Component Molecular Orbital Approach. *Mol. Phys.* **2002**, *100*, 881–901, DOI: 10.1080/00268970110099602 (pages 18, 64).

[203] Ishimoto, T.; Tachikawa, M.; Yamauchi, M.; Kitagawa, H.; Tokiwa, H.; Nagashima, U. Analysis of Isotope Effect of Hydrogen-Absorbing Pd Ultra-Fine Particle by X-ray Powder Diffraction and First Principle Multi-Component MO Calculation. *Chem. Phys. Lett.* **2003**, *372*, 503–507, DOI: 10.1016/S0009-2614(03)00414-7 (pages 18, 64).

[204] Itou, Y.; Mori, S.; Udagawa, T.; Tachikawa, M.; Ishimoto, T.; Nagashima, U. Quantum Treatment of Hydrogen Nuclei in Primary Kinetic Isotope Effects in a Thermal [1,5]-Sigmatropic Hydrogen (or Deuterium) Shift from (Z)-1,3-Pentadiene. *J. Phys. Chem. A* **2007**, *111*, 261–267, DOI: 10.1021/jp065759x (pages 18, 64).

[205] Pavošević, F.; Hammes-Schiffer, S. Multicomponent Coupled Cluster Singles and Doubles and Brueckner Doubles Methods: Proton Densities and Energies. *J. Chem. Phys.* **2019**, *151*, 074104, DOI: 10.1063/1.5116113 (pages 18, 64).

[206] Moncada, F.; Cruz, D.; Reyes, A. Muonic Alchemy: Transmuting Elements with the Inclusion of Negative Muons. *Chem. Phys. Lett.* **2012**, *539–540*, 209–213, DOI: 10.1016/j.cplett.2012.04.062 (pages 18, 64).

[207] Brorsen, K. R.; Pak, M. V.; Hammes-Schiffer, S. Calculation of Positron Binding Energies and Electron–Positron Annihilation Rates for Atomic Systems with the Reduced Explicitly Correlated Hartree–Fock Method in the Nuclear–Electronic Orbital Framework. *J. Phys. Chem. A* **2017**, *121*, 515–522, DOI: 10.1021/acs.jpca.6b10124 (pages 18, 64).

[208] Charry, J.; Varella, M. T. d. N.; Reyes, A. Binding Matter with Antimatter: The Covalent Positron Bond. *Angew. Chem. Int. Ed.* **2018**, *57*, 8859–8864, DOI: 10.1002/anie.201800914 (pages 18, 64).

[209] Romero, J.; Charry, J. A.; Flores-Moreno, R.; Varella, M. T. d. N.; Reyes, A. Calculation of Positron Binding Energies Using the Generalized Any Particle Propagator Theory. *J. Chem. Phys.* **2014**, *141*, 114103, DOI: 10.1063/1.4895043 (pages 18, 64).

[210] Charry, J.; Romero, J.; Varella, M. T. d. N.; Reyes, A. Calculation of Positron Binding Energies of Amino Acids with the Any-Particle Molecular-Orbital Approach. *Phys. Rev. A* **2014**, *89*, 052709, DOI: 10.1103/PhysRevA.89.052709 (pages 18, 64).

[211] Moncada, F.; Pedraza-González, L.; Charry, J.; Varella, M. T. d. N.; Reyes, A. Covalent Bonds in Positron Dihalides. *Chem. Sci.* **2019**, *11*, 44–52, DOI: 10.1039/C9SC04433G (pages 18, 64).

[212] Charry Martinez, J. A.; Barborini, M.; Tkatchenko, A. Correlated Wave Functions for Electron–Positron Interactions in Atoms and Molecules. *J. Chem. Theory Comput.* **2022**, *18*, 2267–2280, DOI: 10.1021/acs.jctc.1c01193 (pages 18, 64, 70, 71, 74).

[213] Charry, J.; Moncada, F.; Barborini, M.; Pedraza-González, L.; Varella, M. T. d. N.; Tkatchenko, A.; Reyes, A. The Three-Center Two-Positron Bond. *Chem. Sci.* **2022**, DOI: 10.1039/D2SC04630J (pages 18, 64).

[214] Bergami, M.; Santana, A. L. D.; Charry Martinez, J.; Reyes, A.; Coutinho, K.; Varella, M. T. d. N. Multicomponent Quantum Mechanics/Molecular Mechanics Study of Hydrated Positronium. *J. Phys. Chem. B* **2022**, *126*, 2699–2714, DOI: 10.1021/acs.jpcb.1c10124 (pages 18, 64).

[215] Varella, M. T. d. N.; Charry, J.; Romero, J.; Reyes, A. Positron-Molecule Interactions: Theory and Computation. *J. Phys.: Conf. Ser.* **2015**, *635*, 032119, DOI: 10.1088/1742-6596/635/3/032119 (pages 18, 64).

[216] Ishimoto, T.; Tachikawa, M.; Nagashima, U. Review of Multicomponent Molecular Orbital Method for Direct Treatment of Nuclear Quantum Effect. *Int. J. Quantum Chem.* **2009**, *109*, 2677–2694, DOI: 10.1002/qua.22069 (page 18).

[217] Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S. Multicomponent Quantum Chemistry: Integrating Electronic and Nuclear Quantum Effects via the Nuclear–Electronic Orbital Method. *Chem. Rev.* **2020**, *120*, 4222–4253, DOI: 10.1021/acs.chemrev.9b00798 (page 18).

[218] Ellis, B. H.; Aggarwal, S.; Chakraborty, A. Development of the Multicomponent Coupled-Cluster Theory for Investigation of Multiexcitonic Interactions. *J. Chem. Theory Comput.* **2016**, *12*, 188–200, DOI: 10.1021/acs.jctc.5b00879 (page 18).

[219] Flores-Moreno, R.; Posada, E.; Moncada, F.; Romero, J.; Charry, J.; Díaz-Tinoco, M.; González, S. A.; Aguirre, N. F.; Reyes, A. LOWDIN: The Any Particle Molecular Orbital Code. *Int. J. Quantum Chem.* **2014**, *114*, 50–56, DOI: 10.1002/qua.24500 (page 18).

[220] Sirjoosingh, A.; Pak, M. V.; Swalina, C.; Hammes-Schiffer, S. Reduced Explicitly Correlated Hartree-Fock Approach within the Nuclear-Electronic Orbital Framework: Applications to Positronic Molecular Systems. *J. Chem. Phys.* **2013**, *139*, 034103, DOI: 10.1063/1.4812259 (page 18).

[221] Moncada, F.; Cruz, D.; Reyes, A. Electronic Properties of Atoms and Molecules Containing One and Two Negative Muons. *Chem. Phys. Lett.* **2013**, *570*, 16–21, DOI: 10.1016/j.cplett.2013.03.004 (page 18).

[222] Posada, E.; Moncada, F.; Reyes, A. Negative Muon Chemistry: The Quantum Muon Effect and the Finite Nuclear Mass Effect. *J. Phys. Chem. A* **2014**, *118*, 9491–9499, DOI: 10.1021/jp501289s (page 18).

[223] Reyes, A.; Moncada, F.; Charry, J. The Any Particle Molecular Orbital Approach: A Short Review of the Theory and Applications. *Int. J. Quantum Chem.* **2019**, *119*, e25705, DOI: 10.1002/qua.25705 (page 18).

[224] Hao, H.; Shee, J.; Upadhyay, S.; Ataca, C.; Jordan, K. D.; Rubenstein, B. M. Accurate Predictions of Electron Binding Energies of Dipole-Bound Anions via Quantum

Monte Carlo Methods. *J. Phys. Chem. Lett.* **2018**, *9*, 6185–6190, DOI: 10.1021/acs.jpclett.8b02733 (pages 19, 38).

[225] Simons, J. Molecular Anions. *J. Phys. Chem. A* **2008**, *112*, 6401–6511, DOI: 10.1021/jp711490b (page 20).

[226] Lykke, K. R.; Mead, R. D.; Lineberger, W. C. Observation of Dipole-Bound States of Negative Ions. *Phys. Rev. Lett.* **1984**, *52*, 2221–2224, DOI: 10.1103/PhysRevLett.52.2221 (page 20).

[227] Sarre, P. J. The Diffuse Interstellar Bands: A Dipole-Bound State Hypothesis. *Monthly Notices of the Royal Astronomical Society* **2000**, *313*, L14–L16, DOI: 10.1046/j.1365-8711.2000.03388.x (page 20).

[228] Sarre, P. J. The Diffuse Interstellar Bands: A Major Problem in Astronomical Spectroscopy. *Journal of Molecular Spectroscopy* **2006**, *238*, 1–10, DOI: 10.1016/j.jms.2006.03.009 (page 20).

[229] Güthe, F.; Tulej, M.; Pachkov, M. V.; Maier, J. P. Photodetachment Spectrum of L-$C_3H_2^-$: The Role of Dipole Bound States for Electron Attachment in Interstellar Clouds. *ApJ* **2001**, *555*, 466, DOI: 10.1086/321465 (page 20).

[230] Maier, J. P.; Walker, G. A. H.; Bohlender, D. A.; Mazzotti, F. J.; Raghunandan, R.; Fulara, J.; Garkusha, I.; Nagy, A. Identification of H_2CCC as a Diffuse Interstellar Band Carrier. *ApJ* **2010**, *726*, 41, DOI: 10.1088/0004-637X/726/1/41 (page 20).

[231] McCall, B. J.; Oka, T.; Thorburn, J.; Hobbs, L. M.; York, D. G. A Critical Examination of the L-$C_3H_2^-$ Spectrum and the Diffuse Interstellar Bands. *ApJ* **2002**, *567*, L145, DOI: 10.1086/340111 (page 20).

[232] Fortenberry, R. C.; Crawford, T. D. Theoretical Prediction of New Dipole-Bound Singlet States for Anions of Interstellar Interest. *J. Chem. Phys.* **2011**, *134*, 154304, DOI: 10.1063/1.3576053 (page 20).

[233] Larsson, M.; Geppert, W. D.; Nyman, G. Ion Chemistry in Space. *Rep. Prog. Phys.* **2012**, *75*, 066901, DOI: 10.1088/0034-4885/75/6/066901 (page 20).

[234] Hendricks, J. H.; Lyapustina, S. A.; de Clercq, H. L.; Snodgrass, J. T.; Bowen, K. H. Dipole Bound, Nucleic Acid Base Anions Studied via Negative Ion Photoelectron

Spectroscopy. *J. Chem. Phys.* **1996**, *104*, 7788–7791, DOI: 10.1063/1.471482 (page 20).

[235] Buytendyk, A. M.; Buonaugurio, A. M.; Xu, S.-J.; Nilles, J. M.; Bowen, K. H.; Kirnosov, N.; Adamowicz, L. Computational and Photoelectron Spectroscopic Study of the Dipole-Bound Anions, Indole$(H_2O)_{1,2}^-$. *J. Chem. Phys.* **2016**, *145*, 024301, DOI: 10.1063/1.4954937 (page 20).

[236] Desfrançois, C.; Abdoul-Carime, H.; Schermann, J.-P. Ground-State Dipole-Bound Anions. *International Journal of Modern Physics B* **1996**, *10*, 1339–1395, DOI: 10.1142/S0217979296000520 (page 20).

[237] Hammer, N. I.; Diri, K.; Jordan, K. D.; Desfrançois, C.; Compton, R. N. Dipole-Bound Anions of Carbonyl, Nitrile, and Sulfoxide Containing Molecules. *J. Chem. Phys.* **2003**, *119*, 3650–3660, DOI: 10.1063/1.1590959 (page 20).

[238] Wang, L.-S.; Ding, C.-F.; Wang, X.-B.; Barlow, S. E. Photodetachment Photoelectron Spectroscopy of Multiply Charged Anions Using Electrospray Ionization. *Rev. Sci. Instrum.* **1999**, *70*, 1957–1966, DOI: 10.1063/1.1149694 (page 20).

[239] Gutowski, M.; Skurski, P.; Jordan, K. D.; Simons, J. Energies of Dipole-Bound Anionic States. *Int. J. Quantum Chem.* **1997**, *64*, 183–191, DOI: 10.1002/(SICI)1097-461X(1997)64:2<183::AID-QUA5>3.0.CO;2-S (page 20).

[240] Gutsev, G. L.; Nooijen, M.; Bartlett, R. J. Valence and Excited Dipole-Bound States of Polar Diatomic Anions: LiH^-, LiF^-, $LiCl^-$, NaH^-, NaF^-, $NaCl^-$, BeO^-, and MgO^-. *Chem. Phys. Lett.* **1997**, *276*, 13–19, DOI: 10.1016/S0009-2614(97)88028-1 (page 20).

[241] 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 20).

[242] Petruzielo, F. R.; Toulouse, J.; Umrigar, C. J. Approaching Chemical Accuracy with Quantum Monte Carlo. *J. Chem. Phys.* **2012**, *136*, 124116, DOI: 10.1063/1.3697846 (page 20).

[243] Motta, M.; Zhang, S. Ab Initio Computations of Molecular Systems by the Auxiliary-Field Quantum Monte Carlo Method. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1364, DOI: 10.1002/wcms.1364 (pages 20, 21, 31, 38).

[244] Suewattana, M.; Purwanto, W.; Zhang, S.; Krakauer, H.; Walter, E. J. Phaseless Auxiliary-Field Quantum Monte Carlo Calculations with Plane Waves and Pseudopotentials: Applications to Atoms and Molecules. *Phys. Rev. B* **2007**, *75*, 245123, DOI: 10.1103/PhysRevB.75.245123 (page 20).

[245] 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 20).

[246] Al-Saidi, W. A.; Krakauer, H.; Zhang, S. Auxiliary-Field Quantum Monte Carlo Study of TiO and MnO Molecules. *Phys. Rev. B* **2006**, *73*, 075103, DOI: 10.1103/PhysRevB.73.075103 (page 20).

[247] Al-Saidi, W. A.; Krakauer, H.; Zhang, S. Auxiliary-Field Quantum Monte Carlo Study of First- and Second-Row Post-d Elements. *J. Chem. Phys.* **2006**, *125*, 154110, DOI: 10.1063/1.2357917 (page 20).

[248] Al-Saidi, W. A.; Krakauer, H.; Zhang, S. A Study of H+H_2 and Several H-bonded Molecules by Phaseless Auxiliary-Field Quantum Monte Carlo with Plane Wave and Gaussian Basis Sets. *J. Chem. Phys.* **2007**, *126*, 194105, DOI: 10.1063/1.2735296 (page 20).

[249] Shee, J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. Chemical Transformations Approaching Chemical Accuracy via Correlated Sampling in Auxiliary-Field Quantum Monte Carlo. *J. Chem. Theory Comput.* **2017**, *13*, 2667–2680, DOI: 10.1021/acs.jctc.7b00224 (pages 20, 22).

[250] Møller, Chr.; 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 21, 30).

[251] 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 21, 32, 34).

[252] 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 (pages 21, 32, 34, 36).

[253] Frisch, M. J.; Head-Gordon, M.; Pople, J. A. A Direct MP2 Gradient Method. *Chem. Phys. Lett.* **1990**, *166*, 275–280, DOI: 10.1016/0009-2614(90)80029-D (page 21).

[254] Frisch, M. J.; Head-Gordon, M.; Pople, J. A. Semi-Direct Algorithms for the MP2 Energy and Gradient. *Chem. Phys. Lett.* **1990**, *166*, 281–289, DOI: 10.1016/0009-2614(90)80030-H (page 21).

[255] Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 Energy Evaluation by Direct Methods. *Chem. Phys. Lett.* **1988**, *153*, 503–506, DOI: 10.1016/0009-2614(88)85250-3 (page 21).

[256] Sæbø, S.; Almlöf, J. Avoiding the Integral Storage Bottleneck in LCAO Calculations of Electron Correlation. *Chem. Phys. Lett.* **1989**, *154*, 83–89, DOI: 10.1016/0009-2614(89)87442-1 (page 21).

[257] Head-Gordon, M.; Head-Gordon, T. Analytic MP2 Frequencies without Fifth-Order Storage. Theory and Application to Bifurcated Hydrogen Bonds in the Water Hexamer. *Chem. Phys. Lett.* **1994**, *220*, 122–128, DOI: 10.1016/0009-2614(94)00116-2 (page 21).

[258] 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 21, 36).

[259] Gordon, M. S.; Schmidt, M. W., Chapter 41 - Advances in Electronic Structure Theory: GAMESS a Decade Later In *Theory and Applications of Computational Chemistry*, Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005, pp 1167–1189, DOI: 10.1016/B978-044451719-7/50084-6 (pages 21, 36).

[260] Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. A. NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations. *Computer Physics Communications* **2010**, *181*, 1477–1489, DOI: 10.1016/j.cpc.2010.04.018 (page 21).

[261] 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 21, 31).

[262] Ceperley, D. M. Fermion Nodes. *J Stat Phys* **1991**, *63*, 1237–1267, DOI: 10.1007/BF01030009 (page 21).

[263] Needs, R. J.; Towler, M. D.; Drummond, N. D.; Ríos, P. L. Continuum Variational and Diffusion Quantum Monte Carlo Calculations. *J. Phys.: Condens. Matter* **2009**, *22*, 023201, DOI: 10.1088/0953-8984/22/2/023201 (page 21).

[264] Zhang, S., Auxiliary-Field Quantum Monte Carlo for Correlated Electron Systems In *Emergent Phenomena in Correlated Matter: Lecture Notes of the Autumn School Correlated Electrons 2013: At Forschungszentrum Jülich, 23-27 September 2013*, Pavarini, E., Koch, E., Schollwöck, U., Eds.; Schriften Des Forschungszentrums Jülich. Reihe Modeling and Simulation Band 3; Forschungszentrum, Zentralbibliothek: Jülich, 2013, pp 15.1–15.31 (page 21).

[265] Drummond, N. D.; Needs, R. J. Variance-Minimization Scheme for Optimizing Jastrow Factors. *Phys. Rev. B* **2005**, *72*, 085124, DOI: 10.1103/PhysRevB.72.085124 (page 21).

[266] Nelson, R. D.; Lide, D. R.; Maryott, A. A. In *Selected Values of Electric Dipole Moments for Molecules in the Gas Phase*, United States Department of Commerce, National Bureau of Standards: 1967 (page 22).

[267] Ard, S.; Garrett, W.; Compton, R.; Adamowicz, L.; Stepanian, S. Rotational States of Dipole-Bound Anions of Hydrogen Cyanide. *Chem. Phys. Lett.* **2009**, *473*, 223–226, DOI: 10.1016/j.cplett.2009.04.007 (page 22).

[268] Desfrançois, C. Determination of Electron Binding Energies of Ground-State Dipole-Bound Molecular Anions. *Phys. Rev. A* **1995**, *51*, 3667–3675, DOI: 10.1103/PhysRevA.51.3667 (page 22).

[269] Gottlieb, C. A.; Killian, T. C.; Thaddeus, P.; Botschwina, P.; Flügge, J.; Oswald, M. Structure of Propadienylidene, H_2CCC. *J. Chem. Phys.* **1993**, *98*, 4478–4485, DOI: 10.1063/1.465007 (page 22).

[270] Yokoyama, K.; Leach, G. W.; Kim, J. B.; Lineberger, W. C.; Boldyrev, A. I.; Gutowski, M. Autodetachment Spectroscopy and Dynamics of Vibrationally Excited Dipole-bound States of H_2CCC^-. *J. Chem. Phys.* **1996**, *105*, 10706–10718, DOI: 10.1063/1.472879 (page 22).

[271] Hammer, N. I.; Hinde, R. J.; Compton, R. N.; Diri, K.; Jordan, K. D.; Radisic, D.; Stokes, S. T.; Bowen, K. H. Dipole-Bound Anions of Highly Polar Molecules: Ethylene Carbonate and Vinylene Carbonate. *J. Chem. Phys.* **2004**, *120*, 685–690, DOI: 10.1063/1.1629669 (page 22).

[272] Xu, J.; Jordan, K. D. Application of the Diffusion Monte Carlo Method to the Binding of Excess Electrons to Water Clusters. *J. Phys. Chem. A* **2010**, *114*, 1364–1366, DOI: 10.1021/jp9066108 (page 23).

[273] Wang, F.-F.; Deible, M. J.; Jordan, K. D. Benchmark Study of the Interaction Energy for an $(H_2O)_{16}$ Cluster: Quantum Monte Carlo and Complete Basis Set Limit MP2 Results. *J. Phys. Chem. A* **2013**, *117*, 7606–7611, DOI: 10.1021/jp404541c (page 23).

[274] Voora, V. K.; Jordan, K. D. Nonvalence Correlation-Bound Anion States of Polycyclic Aromatic Hydrocarbons. *J. Phys. Chem. Lett.* **2015**, *6*, 3994–3997, DOI: 10.1021/acs.jpclett.5b01858 (page 28).

[275] Tulej, M.; Kirkwood, D. A.; Pachkov, M.; Maier, J. P. Gas-Phase Electronic Transitions of Carbon Chain Anions Coinciding with Diffuse Interstellar Bands. *ApJ* **1998**, *506*, L69, DOI: 10.1086/311637 (page 28).

[276] Voora, V. K.; Cederbaum, L. S.; Jordan, K. D. Existence of a Correlation Bound S-Type Anion State of C_{60}. *J. Phys. Chem. Lett.* **2013**, *4*, 849–853, DOI: 10.1021/jz400195s (pages 28, 30).

[277] Zen, A.; Brandenburg, J. G.; Klimeš, J.; Tkatchenko, A.; Alfè, D.; Michaelides, A. Fast and Accurate Quantum Monte Carlo for Molecular Crystals. *PNAS* **2018**, *115*, 1724–1729, DOI: 10.1073/pnas.1715434115 (page 28).

[278] Upadhyay, S.; Dumi, A.; Shee, J.; Jordan, K. D. The Role of High-Order Electron Correlation Effects in a Model System for Non-Valence Correlation-Bound Anions. *J. Chem. Phys.* **2020**, *153*, 224118, DOI: 10.1063/5.0030942 (pages 29, 64).

[279] Kairalapova, A.; Jordan, K. D.; Maienshein, D. N.; Fair, M. C.; Falcetta, M. F. Prediction of a Nonvalence Temporary Anion Shape Resonance for a Model $(H_2O)_4$ System. *J. Phys. Chem. A* **2019**, *123*, 2719–2726, DOI: 10.1021/acs.jpca.8b11881 (pages 30, 32).

[280] Kairalapova, A.; Jordan, K. D.; Falcetta, M. F.; Steiner, D. K.; Sutter, B. L.; Gowen, J. S. Prediction of a Non-Valence Temporary Anion State of $(NaCl)_2$. *J. Phys. Chem. B* **2019**, *123*, 9198–9205, DOI: 10.1021/acs.jpcb.9b07782 (page 30).

[281] Rogers, J. P.; Anstöter, C. S.; Verlet, J. R. R. Ultrafast Dynamics of Low-Energy Electron Attachment via a Non-Valence Correlation-Bound State. *Nat. Chem.* **2018**, *10*, 341–346, DOI: 10.1038/nchem.2912 (page 30).

[282] Rogers, J. P.; Anstöter, C. S.; Verlet, J. R. R. Evidence of Electron Capture of an Outgoing Photoelectron Wave by a Nonvalence State in $(C_6F_6)_n^-$. *J. Phys. Chem. Lett.* **2018**, *9*, 2504–2509, DOI: 10.1021/acs.jpclett.8b00739 (page 30).

[283] Bull, J. N.; Anstöter, C. S.; Verlet, J. R. R. Ultrafast Valence to Non-Valence Excited State Dynamics in a Common Anionic Chromophore. *Nat. Commun.* **2019**, *10*, 5820, DOI: 10.1038/s41467-019-13819-6 (page 30).

[284] Verlet, J. R. R.; Anstöter, C. S.; Bull, J. N.; Rogers, J. P. Role of Nonvalence States in the Ultrafast Dynamics of Isolated Anions. *J. Phys. Chem. A* **2020**, *124*, 3507–3519, DOI: 10.1021/acs.jpca.0c01260 (page 30).

[285] Voora, V. K.; Jordan, K. D. Nonvalence Correlation-Bound Anion States of Spherical Fullerenes. *Nano Lett.* **2014**, *14*, 4602–4606, DOI: 10.1021/nl5016574 (page 30).

[286] Choi, T. H.; Jordan, K. D. Model Potential Study of Non-Valence Correlation-Bound Anions of (C_{60})n Clusters: The Role of Electric Field-Induced Charge Transfer. *Faraday Discuss.* **2019**, *217*, 547–560, DOI: 10.1039/C8FD00199E (page 30).

[287] Sommerfeld, T.; Bhattarai, B.; Vysotskiy, V. P.; Cederbaum, L. S. Correlation-Bound Anions of NaCl Clusters. *J. Chem. Phys.* **2010**, *133*, 114301, DOI: 10.1063/1.3488228 (page 30).

[288] Bezchastnov, V. G.; Vysotskiy, V. P.; Cederbaum, L. S. Anions of Xenon Clusters Bound by Long-Range Electron Correlations. *Phys. Rev. Lett.* **2011**, *107*, 133401, DOI: 10.1103/PhysRevLett.107.133401 (page 30).

[289] 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 30, 45).

[290] Lochan, R. C.; Head-Gordon, M. Orbital-Optimized Opposite-Spin Scaled Second-Order Correlation: An Economical Method to Improve the Description of Open-Shell Molecules. *J. Chem. Phys.* **2007**, *126*, 164101, DOI: 10.1063/1.2718952 (page 30).

[291] Čížek, J.; Paldus, J.; Šroubková, L. Cluster expansion analysis for delocalized systems. *Int. J. Quantum Chem.* **1969**, *3*, 149–167, DOI: 10.1002/qua.560030202 (page 30).

[292] Dutta, A. K.; Saitow, M.; Demoulin, B.; Neese, F.; Izsák, R. A Domain-Based Local Pair Natural Orbital Implementation of the Equation of Motion Coupled Cluster Method for Electron Attached States. *J. Chem. Phys.* **2019**, *150*, 164123, DOI: 10.1063/1.5089637 (page 31).

[293] Grimm, R. C.; Storer, R. G. Monte-Carlo Solution of Schrödinger's Equation. *Journal of Computational Physics* **1971**, *7*, 134–156, DOI: 10.1016/0021-9991(71)90054-4 (page 31).

[294] Anderson, J. B. Quantum Chemistry by Random Walk. H ^{2}P, H_3^+ D_{3h} $^1A_1^{'}$, H_2 $^3\Sigma_u^+$, H_4 $^1\Sigma_g^+$, Be ^{1}S. *J. Chem. Phys.* **1976**, *65*, 4121–4127, DOI: 10.1063/1.432868 (page 31).

[295] 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 31).

[296] Blankenbecler, R.; Scalapino, D. J.; Sugar, R. L. Monte Carlo Calculations of Coupled Boson-Fermion Systems. I. *Phys. Rev. D* **1981**, *24*, 2278–2286, DOI: 10.1103/PhysRevD.24.2278 (pages 31, 38).

[297] Sugiyama, G.; Koonin, S. E. Auxiliary Field Monte-Carlo for Quantum Many-Body Ground States. *Annals of Physics* **1986**, *168*, 1–26, DOI: 10.1016/0003-4916(86)90107-7 (pages 31, 38).

[298] Scalapino, D. J.; Sugar, R. L. Monte Carlo Calculations of Coupled Boson-Fermion Systems. II. *Phys. Rev. B* **1981**, *24*, 4295–4308, DOI: 10.1103/PhysRevB.24.4295 (pages 31, 38).

[299] Zhang, S., Ab Initio Electronic Structure Calculations by Auxiliary-Field Quantum Monte Carlo In *Handbook of Materials Modeling : Methods: Theory and Modeling*, Andreoni, W., Yip, S., Eds.; Springer International Publishing: Cham, 2018, pp 1–27, DOI: 10.1007/978-3-319-42913-7_47-1 (pages 31, 38).

[300] Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38, DOI: 10.1016/0263-7855(96)00018-5 (page 33).

[301] Matthews, D. A.; Stanton, J. F. A New Approach to Approximate Equation-of-Motion Coupled Cluster with Triple Excitations. *J. Chem. Phys.* **2016**, *145*, 124102, DOI: 10.1063/1.4962910 (page 32).

[302] Kucharski, S. A.; Włoch, M.; Musiał, M.; Bartlett, R. J. Coupled-Cluster Theory for Excited Electronic States: The Full Equation-of-Motion Coupled-Cluster Single, Double, and Triple Excitation Method. *J. Chem. Phys.* **2001**, *115*, 8263–8266, DOI: 10.1063/1.1416173 (page 32).

[303] Kowalski, K.; Piecuch, P. The Active-Space Equation-of-Motion Coupled-Cluster Methods for Excited Electronic States: Full EOMCCSDt. *J. Chem. Phys.* **2001**, *115*, 643–651, DOI: 10.1063/1.1378323 (page 32).

[304] Stanton, J. F.; Gauss, J.; Cheng, L.; Harding, M. E.; Matthews, D. A.; Szalay, P. G. CFOUR, Coupled-Cluster Techniques for Computational Chemistry, a Quantum-Chemical Program Package (page 34).

[305] Matthews, D. A.; Cheng, L.; Harding, M. E.; Lipparini, F.; Stopkowicz, S.; Jagau, T.-C.; Szalay, P. G.; Gauss, J.; Stanton, J. F. Coupled-Cluster Techniques for Computational Chemistry: The CFOUR Program Package. *J. Chem. Phys.* **2020**, *152*, 214108, DOI: 10.1063/5.0004837 (page 34).

[306] Jastrow, R. Many-Body Problem with Strong Forces. *Phys. Rev.* **1955**, *98*, 1479–1484, DOI: 10.1103/PhysRev.98.1479 (page 34).

[307] 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 34, 54).

[308] 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.: Condens. Matter* **2018**, *30*, 195901, DOI: 10.1088/1361-648X/aab9c3 (pages 34, 36, 56).

[309] 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 (pages 34, 37).

[310] 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 34, 37).

[311] 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 34, 37).

[312] 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 (page 34).

[313] Bennett, M. C.; Melton, C. A.; Annaberdiyev, A.; Wang, G.; Shulenburger, L.; Mitas, L. A New Generation of Effective Core Potentials for Correlated Calculations. *J. Chem. Phys.* **2017**, *147*, 224106, DOI: 10.1063/1.4995643 (pages 36, 38, 52, 54).

[314] Annaberdiyev, A.; Wang, G.; Melton, C. A.; Bennett, M. C.; Shulenburger, L.; Mitas, L. A New Generation of Effective Core Potentials from Correlated Calculations: 3d Transition Metal Series. *J. Chem. Phys.* **2018**, *149*, 134108, DOI: 10.1063/1.5040472 (pages 36, 52, 54).

[315] Casula, M. Beyond the Locality Approximation in the Standard Diffusion Monte Carlo Method. *Phys. Rev. B* **2006**, *74*, 161102, DOI: 10.1103/PhysRevB.74.161102 (page 36).

[316] Kent, P. R. C.; Annaberdiyev, A.; Benali, A.; Bennett, M. C.; Landinez Borda, E. J.; Doak, P.; Hao, H.; Jordan, K. D.; Krogel, J. T.; Kylänpää, I.; Lee, J.; Luo, Y.; Malone, F. D.; Melton, C. A.; Mitas, L.; Morales, M. A.; Neuscamman, E.; Reboredo, F. A.; Rubenstein, B.; Saritas, K.; Upadhyay, S.; Wang, G.; Zhang, S.; Zhao, L. QMCPACK: Advances in the Development, Efficiency, and Application of Auxiliary Field and Real-Space Variational and Diffusion Quantum Monte Carlo. *J. Chem. Phys.* **2020**, *152*, 174105, DOI: 10.1063/5.0004860 (pages 36, 56).

[317] Barca, G. M. J.; Bertoni, C.; Carrington, L.; Datta, D.; De Silva, N.; Deustua, J. E.; Fedorov, D. G.; Gour, J. R.; Gunina, A. O.; Guidez, E.; Harville, T.; Irle, S.; Ivanic, J.; Kowalski, K.; Leang, S. S.; Li, H.; Li, W.; Lutz, J. J.; Magoulas, I.; Mato, J.; Mironov, V.; Nakata, H.; Pham, B. Q.; Piecuch, P.; Poole, D.; Pruitt, S. R.; Rendell, A. P.; Roskop, L. B.; Ruedenberg, K.; Sattasathuchana, T.; Schmidt, M. W.; Shen, J.; Slipchenko, L.; Sosonkina, M.; Sundriyal, V.; Tiwari, A.; Galvez Vallejo, J. L.; Westheimer, B.; Włoch, M.; Xu, P.; Zahariev, F.; Gordon, M. S. Recent Developments in the General Atomic and Molecular Electronic Structure System. *J. Chem. Phys.* **2020**, *152*, 154102, DOI: 10.1063/5.0005188 (page 36).

[318] 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 (page 36).

[319] Lee, J.; Malone, F. D.; Morales, M. A. Utilizing Essential Symmetry Breaking in Auxiliary-Field Quantum Monte Carlo: Application to the Spin Gaps of the C_{36} Fullerene and an Iron Porphyrin Model Complex. *J. Chem. Theory Comput.* **2020**, DOI: 10.1021/acs.jctc.0c00055 (page 38).

[320] Rudshteyn, B.; Coskun, D.; Weber, J. L.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A.; Shee, J. Predicting Ligand-Dissociation Energies of 3d Coordination Complexes with Auxiliary-Field Quantum Monte Carlo. *J. Chem. Theory Comput.* **2020**, *16*, 3041–3054, DOI: 10.1021/acs.jctc.0c00070 (page 38).

[321] Shee, J.; Rudshteyn, B.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. On Achieving High Accuracy in Quantum Chemical Calculations of 3d Transition Metal-Containing Systems: A Comparison of Auxiliary-Field Quantum Monte Carlo with Coupled Cluster, Density Functional Theory, and Experiment for Diatomic Molecules. *J. Chem. Theory Comput.* **2019**, *15*, 2346–2358, DOI: 10.1021/acs.jctc.9b00083 (page 38).

[322] Kumar, M.; Shee, J.; Rudshteyn, B.; Reichman, D. R.; Friesner, R. A.; Miller, C. E.; Francisco, J. S. Multiple Stable Isoprene–Ozone Complexes Reveal Complex Entrance Channel Dynamics in the Isoprene + Ozone Reaction. *J. Am. Chem. Soc.* **2020**, *142*, 10806–10813, DOI: 10.1021/jacs.0c02360 (page 38).

[323] Shee, J.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. Singlet–Triplet Energy Gaps of Organic Biradicals and Polyacenes with Auxiliary-Field Quantum Monte Carlo. *J. Chem. Theory Comput.* **2019**, *15*, 4924–4932, DOI: 10.1021/acs.jctc.9b00534 (page 38).

[324] Shee, J.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. Phaseless Auxiliary-Field Quantum Monte Carlo on Graphical Processing Units. *J. Chem. Theory Comput.* **2018**, *14*, 4109–4121, DOI: 10.1021/acs.jctc.8b00342 (page 38).

[325] Purwanto, W.; Krakauer, H.; Virgus, Y.; Zhang, S. Assessing Weak Hydrogen Binding on Ca+ Centers: An Accurate Many-Body Study with Large Basis Sets. *J. Chem. Phys.* **2011**, *135*, 164105, DOI: 10.1063/1.3654002 (page 38).

[326] O'boyle, N. M.; Tenderholt, A. L.; Langner, K. M. Cclib: A Library for Package-Independent Computational Chemistry Algorithms. *J. Comput. Chem.* **2008**, *29*, 839–845, DOI: 10.1002/jcc.20823 (page 39).

[327] Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K.-L. PySCF: The Python-based Simulations of Chemistry Framework. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1340, DOI: 10.1002/wcms.1340 (page 39).

[328] Sun, Q.; Zhang, X.; Banerjee, S.; Bao, P.; Barbry, M.; Blunt, N. S.; Bogdanov, N. A.; Booth, G. H.; Chen, J.; Cui, Z.-H.; Eriksen, J. J.; Gao, Y.; Guo, S.; Hermann, J.; Hermes, M. R.; Koh, K.; Koval, P.; Lehtola, S.; Li, Z.; Liu, J.; Mardirossian, N.; McClain, J. D.; Motta, M.; Mussard, B.; Pham, H. Q.; Pulkin, A.; Purwanto, W.; Robinson, P. J.; Ronca, E.; Sayfutyarova, E. R.; Scheurer, M.; Schurkus, H. F.; Smith, J. E. T.; Sun, C.; Sun, S.-N.; Upadhyay, S.; Wagner, L. K.; Wang, X.; White, A.; Whitfield, J. D.; Williamson, M. J.; Wouters, S.; Yang, J.; Yu, J. M.; Zhu, T.; Berkelbach, T. C.; Sharma, S.; Sokolov, A. Y.; Chan, G. K.-L. Recent Developments in the PySCF Program Package. *J. Chem. Phys.* **2020**, *153*, 024109, DOI: 10.1063/5.0006074 (page 39).

[329] Lebedev, V. I.; Laikov, D. N. A Quadrature Formula for the Sphere of the 131st Algebraic Order of Accuracy. *Dokl. Math.* **1999**, *59*, 477–481 (page 39).

[330] Schlömer, N.; Papior, N. R.; Arnold, D.; Ancellin, M.; Zetter, R. Nschloe/Quadpy v0.15.2, version v0.15.2, Zenodo, 2020, DOI: 10.5281/zenodo.3975551 (page 39).

[331] Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95, DOI: 10.1109/MCSE.2007.55 (pages 44, 56, 73).

[332] Lee, J.; Small, D. W.; Head-Gordon, M. Excited States via Coupled Cluster Theory without Equation-of-Motion Methods: Seeking Higher Roots with Application to Doubly Excited States and Double Core Hole States. *J. Chem. Phys.* **2019**, *151*, 214103, DOI: 10.1063/1.5128795 (page 45).

[333] Noga, J.; Bartlett, R. J. The Full CCSDT Model for Molecular Electronic Structure. *J. Chem. Phys.* **1987**, *86*, 7041–7050, DOI: 10.1063/1.452353 (page 45).

[334] Scuseria, G. E.; Schaefer, H. F. A New Implementation of the Full CCSDT Model for Molecular Electronic Structure. *Chem. Phys. Lett.* **1988**, *152*, 382–386, DOI: 10.1016/0009-2614(88)80110-6 (page 45).

[335] Watts, J. D.; Bartlett, R. J. The Coupled-cluster Single, Double, and Triple Excitation Model for Open-shell Single Reference Functions. *J. Chem. Phys.* **1990**, *93*, 6104–6105, DOI: 10.1063/1.459002 (page 45).

[336] Bomble, Y. J.; Stanton, J. F.; Kállay, M.; Gauss, J. Coupled-Cluster Methods Including Noniterative Corrections for Quadruple Excitations. *J. Chem. Phys.* **2005**, *123*, 054101, DOI: 10.1063/1.1950567 (page 45).

[337] Dumi, A.; Upadhyay, S.; Bernasconi, L.; Shin, H.; Benali, A.; Jordan, K. D. The Binding of Atomic Hydrogen on Graphene from Density Functional Theory and Diffusion Monte Carlo Calculations. *J. Chem. Phys.* **2022**, *156*, 144702, DOI: 10.1063/5.0085982 (pages 50, 51).

[338] Hughes, T. J.; Shaw, R. A.; Russo, S. P. Computational Investigations of Dispersion Interactions between Small Molecules and Graphene-like Flakes. *J. Phys. Chem. A* **2020**, *124*, 9552–9561, DOI: 10.1021/acs.jpca.0c06595 (page 51).

[339] Bartolomei, M.; Hernández, M. I.; Campos-Martínez, J.; Hernández-Lamoneda, R.; Giorgi, G. Permeation of Chemisorbed Hydrogen through Graphene: A Flipping Mechanism Elucidated. *Carbon* **2021**, *178*, 718–727, DOI: 10.1016/j.carbon.2021.02.056 (page 51).

[340] Ma, J.; Michaelides, A.; Alfè, D. Binding of Hydrogen on Benzene, Coronene, and Graphene from Quantum Monte Carlo Calculations. *J. Chem. Phys.* **2011**, *134*, 134701, DOI: 10.1063/1.3569134 (pages 51, 52).

[341] Shin, H.; Luo, Y.; Benali, A.; Kwon, Y. Diffusion Monte Carlo Study of O_2 Adsorption on Single Layer Graphene. *Phys. Rev. B* **2019**, *100*, 075430, DOI: 10.1103/PhysRevB.100.075430 (pages 51, 52, 54).

[342] Ma, J.; Michaelides, A.; Alfè, D.; Schimka, L.; Kresse, G.; Wang, E. Adsorption and Diffusion of Water on Graphene from First Principles. *Phys. Rev. B* **2011**, *84*, 033402, DOI: 10.1103/PhysRevB.84.033402 (pages 51, 52).

[343] Ahn, J.; Hong, I.; Lee, G.; Shin, H.; Benali, A.; Kwon, Y. Adsorption of a Single Pt Atom on Graphene: Spin Crossing between Physisorbed Triplet and Chemisorbed Singlet States. *Phys. Chem. Chem. Phys.* **2021**, *23*, 22147–22154, DOI: 10.1039/D1CP02473F (pages 51, 52).

[344] Geim, A. K. Graphene: Status and Prospects. *Science* **2009**, *324*, 1530–1534, DOI: 10.1126/science.1158877 (page 51).

[345] Novoselov, K. S.; Geim, A. K.; Morozov, S. V.; Jiang, D.; Katsnelson, M. I.; Grigorieva, I. V.; Dubonos, S. V.; Firsov, A. A. Two-Dimensional Gas of Massless Dirac Fermions in Graphene. *Nature* **2005**, *438*, 197–200, DOI: 10.1038/nature04233 (page 51).

[346] Alekseeva, O. K.; Pushkareva, I. V.; Pushkarev, A. S.; Fateev, V. N. Graphene and Graphene-Like Materials for Hydrogen Energy. *Nanotechnol Russia* **2020**, *15*, 273–300, DOI: 10.1134/S1995078020030027 (page 51).

[347] Niaz, S.; Manzoor, T.; Pandith, A. H. Hydrogen Storage: Materials, Methods and Perspectives. *Renewable and Sustainable Energy Reviews* **2015**, *50*, 457–469, DOI: 10.1016/j.rser.2015.05.011 (page 51).

[348] Fomkin, A.; Pribylov, A.; Men'shchikov, I.; Shkolin, A.; Aksyutin, O.; Ishkov, A.; Romanov, K.; Khozina, E. Adsorption-Based Hydrogen Storage in Activated Carbons and Model Carbon Structures. *Reactions* **2021**, *2*, 209–226, DOI: 10.3390/reactions2030014 (page 51).

[349] Ataca, C.; Aktürk, E.; Ciraci, S.; Ustunel, H. High-Capacity Hydrogen Storage by Metallized Graphene. *Appl. Phys. Lett.* **2008**, *93*, 043123, DOI: 10.1063/1.2963976 (page 51).

[350] Dimitrakakis, G. K.; Tylianakis, E.; Froudakis, G. E. Pillared Graphene: A New 3-D Network Nanostructure for Enhanced Hydrogen Storage. *Nano Lett.* **2008**, *8*, 3166–3170, DOI: 10.1021/nl801417w (page 51).

[351] Sha, X.; Jackson, B. First-Principles Study of the Structural and Energetic Properties of H Atoms on a Graphite (0001) Surface. *Surface Science* **2002**, *496*, 318–330, DOI: 10.1016/S0039-6028(01)01602-8 (page 51).

[352] Jeloaica, L.; Sidis, V. DFT Investigation of the Adsorption of Atomic Hydrogen on a Cluster-Model Graphite Surface. *Chem. Phys. Lett.* **1999**, *300*, 157–162, DOI: 10.1016/S0009-2614(98)01337-2 (page 51).

[353] Sahu, S.; Rout, G. C. Band Gap Opening in Graphene: A Short Theoretical Study. *Int Nano Lett* **2017**, *7*, 81–89, DOI: 10.1007/s40089-017-0203-5 (page 51).

[354] González-Herrero, H.; Gómez-Rodríguez, J. M.; Mallet, P.; Moaied, M.; Palacios, J. J.; Salgado, C.; Ugeda, M. M.; Veuillen, J.-Y.; Yndurain, F.; Brihuega, I. Atomic-Scale Control of Graphene Magnetism by Using Hydrogen Atoms. *Science* **2016**, *352*, 437–441, DOI: 10.1126/science.aad8038 (page 51).

[355] González-Herrero, H.; Río, E. C.-d.; Mallet, P.; Veuillen, J.-Y.; Palacios, J. J.; Gómez-Rodríguez, J. M.; Brihuega, I.; Ynduráin, F. Hydrogen Physisorption Channel on Graphene: A Highway for Atomic H Diffusion. *2D Mater.* **2019**, *6*, 021004, DOI: 10.1088/2053-1583/ab03a0 (page 51).

[356] Hummel, F.; Tsatsoulis, T.; Grüneis, A. Low Rank Factorization of the Coulomb Integrals for Periodic Coupled Cluster Theory. *J. Chem. Phys.* **2017**, *146*, 124105, DOI: 10.1063/1.4977994 (page 51).

[357] Mihm, T. N.; McIsaac, A. R.; Shepherd, J. J. An Optimized Twist Angle to Find the Twist-Averaged Correlation Energy Applied to the Uniform Electron Gas. *J. Chem. Phys.* **2019**, *150*, 191101, DOI: 10.1063/1.5091445 (page 51).

[358] Callahan, J. M.; Lange, M. F.; Berkelbach, T. C. Dynamical Correlation Energy of Metals in Large Basis Sets from Downfolding and Composite Approaches. *J. Chem. Phys.* **2021**, *154*, 211105, DOI: 10.1063/5.0049890 (page 51).

[359] Booth, G. H.; Grüneis, A.; Kresse, G.; Alavi, A. Towards an Exact Description of Electronic Wavefunctions in Real Solids. *Nature* **2013**, *493*, 365–370, DOI: 10.1038/nature11770 (page 51).

[360] Schäfer, T.; Libisch, F.; Kresse, G.; Grüneis, A. Local Embedding of Coupled Cluster Theory into the Random Phase Approximation Using Plane Waves. *J. Chem. Phys.* **2021**, *154*, 011101, DOI: 10.1063/5.0036363 (page 51).

[361] Benali, A.; Gasperich, K.; Jordan, K. D.; Applencourt, T.; Luo, Y.; Bennett, M. C.; Krogel, J. T.; Shulenburger, L.; Kent, P. R. C.; Loos, P.-F.; Scemama, A.; Caffarel, M. Toward a Systematic Improvement of the Fixed-Node Approximation in Diffusion Monte Carlo for Solids—A Case Study in Diamond. *J. Chem. Phys.* **2020**, *153*, 184111, DOI: 10.1063/5.0021036 (page 51).

[362] Schäfer, T.; Ramberger, B.; Kresse, G. Quartic Scaling MP2 for Solids: A Highly Parallelized Algorithm in the Plane Wave Basis. *J. Chem. Phys.* **2017**, *146*, 104101, DOI: 10.1063/1.4976937 (page 51).

[363] Brandenburg, J. G.; Zen, A.; Fitzner, M.; Ramberger, B.; Kresse, G.; Tsatsoulis, T.; Grüneis, A.; Michaelides, A.; Alfè, D. Physisorption of Water on Graphene: Subchemical Accuracy from Many-Body Electronic Structure Methods. *J. Phys. Chem. Lett.* **2019**, *10*, 358–368, DOI: 10.1021/acs.jpclett.8b03679 (page 52).

[364] 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 52).

[365] Perdew, J. P. In *Electron. Struct. Solids 91*, ed. by Ziesche, P.; Eschrig, H., Akademie Verlag: Berlin, 1991; Vol. 17, pp 11–20 (page 52).

[366] Roman, T.; Diño, W. A.; Nakanishi, H.; Kasai, H.; Sugimoto, T.; Tange, K. Hydrogen Pairing on Graphene. *Carbon* **2007**, *45*, 218–220, DOI: 10.1016/j.carbon.2006.09.027 (page 52).

[367] Šljivančanin, Ž.; Rauls, E.; Hornekær, L.; Xu, W.; Besenbacher, F.; Hammer, B. Extended Atomic Hydrogen Dimer Configurations on the Graphite(0001) Surface. *J. Chem. Phys.* **2009**, *131*, 084706, DOI: 10.1063/1.3187941 (pages 52, 56).

[368] Lehtinen, P. O.; Foster, A. S.; Ma, Y.; Krasheninnikov, A. V.; Nieminen, R. M. Irradiation-Induced Magnetism in Graphite: A Density Functional Study. *Phys. Rev. Lett.* **2004**, *93*, 187202, DOI: 10.1103/PhysRevLett.93.187202 (pages 52, 56).

[369] Lin, Y.; Ding, F.; Yakobson, B. I. Hydrogen Storage by Spillover on Graphene as a Phase Nucleation Process. *Phys. Rev. B* **2008**, *78*, 041402, DOI: 10.1103/PhysRevB.78.041402 (pages 52, 56).

[370] Casolo, S.; Løvvik, O. M.; Martinazzo, R.; Tantardini, G. F. Understanding Adsorption of Hydrogen Atoms on Graphene. *J. Chem. Phys.* **2009**, *130*, 054704, DOI: 10.1063/1.3072333 (pages 52, 56).

[371] Miwa, R. H.; Martins, T. B.; Fazzio, A. Hydrogen Adsorption on Boron Doped Graphene: An Ab Initio Study. *Nanotechnology* **2008**, *19*, 155708, DOI: 10.1088/0957-4484/19/15/155708 (pages 52, 56).

[372] Ishii, A.; Yamamoto, M.; Asano, H.; Fujiwara, K. DFT Calculation for Adatom Adsorption on Graphene Sheet as a Prototype of Carbon Nanotube Functionalization. *J. Phys.: Conf. Ser.* **2008**, *100*, 052087, DOI: 10.1088/1742-6596/100/5/052087 (page 52).

[373] Li, W.; Zhao, M.; He, T.; Song, C.; Lin, X.; Liu, X.; Xia, Y.; Mei, L. Concentration Dependent Magnetism Induced by Hydrogen Adsorption on Graphene and Single Walled Carbon Nanotubes. *Journal of Magnetism and Magnetic Materials* **2010**, *322*, 838–843, DOI: 10.1016/j.jmmm.2009.11.014 (page 52).

[374] Boukhvalov, D. W.; Katsnelson, M. I.; Lichtenstein, A. I. Hydrogen on Graphene: Electronic Structure, Total Energy, Structural Distortions and Magnetism from First-Principles Calculations. *Phys. Rev. B* **2008**, *77*, 035427, DOI: 10.1103/PhysRevB.77.035427 (page 52).

[375] Kim, M. A.; Sorescu, D. C.; Amemiya, S.; Jordan, K. D.; Liu, H. Real-Time Modulation of Hydrogen Evolution Activity of Graphene Electrodes Using Mechanical Strain. *ACS Appl. Mater. Interfaces* **2022**, *14*, 10691–10700, DOI: 10.1021/acsami.1c21821 (pages 52, 57).

[376] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104, DOI: 10.1063/1.3382344 (page 52).

[377] Monkhorst, H. J.; Pack, J. D. Special Points for Brillouin-zone Integrations. *Phys. Rev. B* **1976**, *13*, 5188–5192, DOI: 10.1103/PhysRevB.13.5188 (page 52).

[378] Marzari, N.; Vanderbilt, D.; De Vita, A.; Payne, M. C. Thermal Contraction and Disordering of the Al(110) Surface. *Phys. Rev. Lett.* **1999**, *82*, 3296–3299, DOI: 10.1103/PhysRevLett.82.3296 (page 52).

[379] 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).

[380] Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals. *J. Chem. Phys.* **2006**, *125*, 224106, DOI: 10.1063/1.2404663 (page 53).

[381] Peintinger, M. F.; Oliveira, D. V.; Bredow, T. Consistent Gaussian Basis Sets of Triple-Zeta Valence with Polarization Quality for Solid-State Calculations. *J. Comput. Chem.* **2013**, *34*, 451–459, DOI: 10.1002/jcc.23153 (page 53).

[382] Kruse, H.; Grimme, S. A Geometrical Correction for the Inter- and Intra-Molecular Basis Set Superposition Error in Hartree-Fock and Density Functional Theory Calculations for Large Systems. *J. Chem. Phys.* **2012**, *136*, 154101, DOI: 10.1063/1.3700154 (page 53).

[383] Brandenburg, J. G.; Alessio, M.; Civalleri, B.; Peintinger, M. F.; Bredow, T.; Grimme, S. Geometrical Correction for the Inter- and Intramolecular Basis Set Superposition Error in Periodic Density Functional Theory Calculations. *J. Phys. Chem. A* **2013**, *117*, 9282–9292, DOI: 10.1021/jp406658y (page 53).

[384] Giannozzi, P.; Baseggio, O.; Bonfà, P.; Brunato, D.; Car, R.; Carnimeo, I.; Cavazzoni, C.; de Gironcoli, S.; Delugas, P.; Ferrari Ruffino, F.; Ferretti, A.; Marzari, N.; Timrov, I.; Urru, A.; Baroni, S. Quantum ESPRESSO toward the Exascale. *J. Chem. Phys.* **2020**, *152*, 154105, DOI: 10.1063/5.0005082 (page 53).

[385] Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Corso, A. D.; Gironcoli, S. de; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello,

A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502, DOI: 10.1088/0953-8984/21/39/395502 (page 53).

[386] Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Nardelli, M. B.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M.; Colonna, N.; Carnimeo, I.; Corso, A. D.; Gironcoli, S. de; Delugas, P.; DiStasio, R. A.; Ferretti, A.; Floris, A.; Fratesi, G.; Fugallo, G.; Gebauer, R.; Gerstmann, U.; Giustino, F.; Gorni, T.; Jia, J.; Kawamura, M.; Ko, H.-Y.; Kokalj, A.; Küçükbenli, E.; Lazzeri, M.; Marsili, M.; Marzari, N.; Mauri, F.; Nguyen, N. L.; Nguyen, H.-V.; Otero-de-la-Roza, A.; Paulatto, L.; Poncé, S.; Rocca, D.; Sabatini, R.; Santra, B.; Schlipf, M.; Seitsonen, A. P.; Smogunov, A.; Timrov, I.; Thonhauser, T.; Umari, P.; Vast, N.; Wu, X.; Baroni, S. Advanced Capabilities for Materials Modelling with Quantum ESPRESSO. *J. Phys.: Condens. Matter* **2017**, *29*, 465901, DOI: 10.1088/1361-648X/aa8f79 (page 53).

[387] Dovesi, R.; Erba, A.; Orlando, R.; Zicovich-Wilson, C. M.; Civalleri, B.; Maschio, L.; Rérat, M.; Casassa, S.; Baima, J.; Salustro, S.; Kirtman, B. Quantum-Mechanical Condensed Matter Simulations with CRYSTAL. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1360, DOI: 10.1002/wcms.1360 (page 53).

[388] Dovesi, R.; Pascale, F.; Civalleri, B.; Doll, K.; Harrison, N. M.; Bush, I.; D'Arco, P.; Noël, Y.; Rérat, M.; Carbonnière, P.; Causà, M.; Salustro, S.; Lacivita, V.; Kirtman, B.; Ferrari, A. M.; Gentile, F. S.; Baima, J.; Ferrero, M.; Demichelis, R.; De La Pierre, M. The CRYSTAL Code, 1976–2020 and beyond, a Long Story. *J. Chem. Phys.* **2020**, *152*, 204111, DOI: 10.1063/5.0004892 (page 53).

[389] Aprà, E.; Bylaska, E. J.; de Jong, W. A.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Valiev, M.; van Dam, H. J. J.; Alexeev, Y.; Anchell, J.; Anisimov, V.; Aquino, F. W.; Atta-Fynn, R.; Autschbach, J.; Bauman, N. P.; Becca, J. C.; Bernholdt, D. E.; Bhaskaran-Nair, K.; Bogatko, S.; Borowski, P.; Boschen, J.; Brabec, J.; Bruner, A.; Cauët, E.; Chen, Y.; Chuev, G. N.; Cramer, C. J.; Daily, J.; Deegan, M. J. O.;

Dunning, T. H.; Dupuis, M.; Dyall, K. G.; Fann, G. I.; Fischer, S. A.; Fonari, A.; Früchtl, H.; Gagliardi, L.; Garza, J.; Gawande, N.; Ghosh, S.; Glaesemann, K.; Götz, A. W.; Hammond, J.; Helms, V.; Hermes, E. D.; Hirao, K.; Hirata, S.; Jacquelin, M.; Jensen, L.; Johnson, B. G.; Jónsson, H.; Kendall, R. A.; Klemm, M.; Kobayashi, R.; Konkov, V.; Krishnamoorthy, S.; Krishnan, M.; Lin, Z.; Lins, R. D.; Littlefield, R. J.; Logsdail, A. J.; Lopata, K.; Ma, W.; Marenich, A. V.; Martin del Campo, J.; Mejia-Rodriguez, D.; Moore, J. E.; Mullin, J. M.; Nakajima, T.; Nascimento, D. R.; Nichols, J. A.; Nichols, P. J.; Nieplocha, J.; Otero-de-la-Roza, A.; Palmer, B.; Panyala, A.; Pirojsirikul, T.; Peng, B.; Peverati, R.; Pittner, J.; Pollack, L.; Richard, R. M.; Sadayappan, P.; Schatz, G. C.; Shelton, W. A.; Silverstein, D. W.; Smith, D. M. A.; Soares, T. A.; Song, D.; Swart, M.; Taylor, H. L.; Thomas, G. S.; Tipparaju, V.; Truhlar, D. G.; Tsemekhman, K.; Van Voorhis, T.; Vázquez-Mayagoitia, Á.; Verma, P.; Villa, O.; Vishnu, A.; Vogiatzis, K. D.; Wang, D.; Weare, J. H.; Williamson, M. J.; Windus, T. L.; Woliński, K.; Wong, A. T.; Wu, Q.; Yang, C.; Yu, Q.; Zacharias, M.; Zhang, Z.; Zhao, Y.; Harrison, R. J. NWChem: Past, Present, and Future. *J. Chem. Phys.* **2020**, *152*, 184102, DOI: 10.1063/5.0004997 (page 53).

[390] Anderson, J. B. Quantum Chemistry by Random Walk: Higher Accuracy. *J. Chem. Phys.* **1980**, *73*, 3897–3899, DOI: 10.1063/1.440575 (page 53).

[391] 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 54).

[392] Umrigar, C. J.; Toulouse, J.; Filippi, C.; Sorella, S.; Hennig, R. G. Alleviation of the Fermion-Sign Problem by Optimization of Many-Body Wave Functions. *Phys. Rev. Lett.* **2007**, *98*, 110201, DOI: 10.1103/PhysRevLett.98.110201 (page 54).

[393] Zen, A.; Brandenburg, J. G.; Michaelides, A.; Alfè, D. A New Scheme for Fixed Node Diffusion Quantum Monte Carlo with Pseudopotentials: Improving Reproducibility and Reducing the Trial-Wave-Function Bias. *J. Chem. Phys.* **2019**, *151*, 134105, DOI: 10.1063/1.5119729 (page 54).

[394] 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 54).

[395] Lin, C.; Zong, F. H.; Ceperley, D. M. Twist-Averaged Boundary Conditions in Continuum Quantum Monte Carlo Algorithms. *Phys. Rev. E* **2001**, *64*, 016702, DOI: 10.1103/PhysRevE.64.016702 (page 54).

[396] Zen, A.; Sorella, S.; Gillan, M. J.; Michaelides, A.; Alfè, D. Boosting the Accuracy and Speed of Quantum Monte Carlo: Size Consistency and Time Step. *Phys. Rev. B* **2016**, *93*, 241118, DOI: 10.1103/PhysRevB.93.241118 (page 54).

[397] Cortés-Arriagada, D.; Gutiérrez-Oliva, S.; Herrera, B.; Soto, K.; Toro-Labbé, A. The Mechanism of Chemisorption of Hydrogen Atom on Graphene: Insights from the Reaction Force and Reaction Electronic Flux. *J. Chem. Phys.* **2014**, *141*, 134701, DOI: 10.1063/1.4896611 (pages 54, 58).

[398] Jiang, H.; Kammler, M.; Ding, F.; Dorenkamp, Y.; Manby, F. R.; Wodtke, Alec. M.; Miller, T. F.; Kandratsenka, A.; Bünermann, O. Imaging Covalent Bond Formation by H Atom Scattering from Graphene. *Science* **2019**, *364*, 379–382, DOI: 10.1126/science.aaw6378 (pages 54, 58).

[399] Krogel, J. T. Nexus: A Modular Workflow Management System for Quantum Simulation Codes. *Computer Physics Communications* **2016**, *198*, 154–168, DOI: 10.1016/j.cpc.2015.08.012 (page 56).

[400] Momma, K.; Izumi, F. VESTA 3 for Three-Dimensional Visualization of Crystal, Volumetric and Morphology Data. *J Appl Cryst* **2011**, *44*, 1272–1276, DOI: 10.1107/S0021889811038970 (page 56).

[401] Liu, A.; Chow, M.; Wildman, A.; Frisch, M. J.; Hammes-Schiffer, S.; Li, X. Simultaneous Optimization of Nuclear–Electronic Orbitals. *J. Phys. Chem. A* **2022**, *126*, 7033–7039, DOI: 10.1021/acs.jpca.2c05172 (page 66).

[402] Davidson, E. R. The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices. *Journal of Computational Physics* **1975**, *17*, 87–94, DOI: 10.1016/0021-9991(75)90065-0 (page 67).

[403] Mella, M.; Casalegno, M.; Morosi, G. Positron and Positronium Chemistry by Quantum Monte Carlo. VI. The Ground State of LiPs, NaPs, e^+Be, and e^+Mg. *J. Chem. Phys.* **2002**, *117*, 1450–1456, DOI: 10.1063/1.1486447 (pages 70, 71, 74).

[404] Mitroy, J. Structure of the LiPs and e^+ Be Systems. *JAMS* **2010**, *1*, 275–279, DOI: 10.4208/jams.071510.072110a (pages 70, 74).

[405] Hunt, R. J.; Szyniszewski, M.; Prayogo, G. I.; Maezono, R.; Drummond, N. D. Quantum Monte Carlo Calculations of Energy Gaps from First Principles. *Phys. Rev. B* **2018**, *98*, 075122, DOI: 10.1103/PhysRevB.98.075122 (page 72).

[406] Zhao, Y.; Guo, Y.; Yao, Y.; Dumi, A.; Mulvey, D. M.; Upadhyay, S.; Zhang, Y.; Jordan, K. D.; Yang, J.; Tang, X. In *2022 IEEE Int. Symp. High-Perform. Comput. Archit. HPCA*, 2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA), 2022, pp 726–740, DOI: 10.1109/HPCA53966.2022.00059 (page 77).

[407] Boixo, S.; Isakov, S. V.; Smelyanskiy, V. N.; Babbush, R.; Ding, N.; Jiang, Z.; Bremner, M. J.; Martinis, J. M.; Neven, H. Characterizing Quantum Supremacy in Near-Term Devices. *Nature Phys* **2018**, *14*, 595–600, DOI: 10.1038/s41567-018-0124-x (pages 78, 84).

[408] Monz, T.; Schindler, P.; Barreiro, J. T.; Chwalla, M.; Nigg, D.; Coish, W. A.; Harlander, M.; Hänsel, W.; Hennrich, M.; Blatt, R. 14-Qubit Entanglement: Creation and Coherence. *Phys. Rev. Lett.* **2011**, *106*, 130506, DOI: 10.1103/PhysRevLett.106.130506 (page 78).

[409] Zhong, H.-S.; Wang, H.; Deng, Y.-H.; Chen, M.-C.; Peng, L.-C.; Luo, Y.-H.; Qin, J.; Wu, D.; Ding, X.; Hu, Y.; Hu, P.; Yang, X.-Y.; Zhang, W.-J.; Li, H.; Li, Y.; Jiang, X.; Gan, L.; Yang, G.; You, L.; Wang, Z.; Li, L.; Liu, N.-L.; Lu, C.-Y.; Pan, J.-W. Quantum Computational Advantage Using Photons. *Science* **2020**, *370*, 1460–1463, DOI: 10.1126/science.abe8770 (page 78).

[410] IBM's Roadmap for Scaling Quantum Technology IBM Research Blog, https://research.ibm.com/blog/ibm-quantum-roadmap (page 78).

[411] Preskill, J. Quantum Computing in the NISQ Era and Beyond. *Quantum* **2018**, *2*, 79, DOI: 10.22331/q-2018-08-06-79 (page 78).

[412] Fatima, A.; Markov, I. L. In 2021 IEEE International Symposium on High-Performance Computer Architecture (HPCA), IEEE Computer Society: 2021, pp 194–207, DOI: 10.1109/HPCA51647.2021.00026 (pages 78, 83, 96, 106, 108, 109).

[413] Markov, I. L.; Fatima, A.; Isakov, S. V.; Boixo, S. In *2020 57th ACMIEEE Des. Autom. Conf. DAC*, 2020 57th ACM/IEEE Design Automation Conference (DAC), 2020, pp 1–6, DOI: 10.1109/DAC18072.2020.9218591 (page 78).

[414] Ding, Y.; Chong, F. T., *Quantum Computer Systems: Research for Noisy Intermediate-Scale Quantum Computers*; Synthesis Lectures on Computer Architecture; Springer International Publishing: Cham, 2020, DOI: 10.1007/978-3-031-01765-0 (pages 78, 82).

[415] Li, A.; Subasi, O.; Yang, X.; Krishnamoorthy, S. In *SC20 Int. Conf. High Perform. Comput. Netw. Storage Anal.* SC20: International Conference for High Performance Computing, Networking, Storage and Analysis, 2020, pp 1–15, DOI: 10.1109/SC41405.2020.00017 (pages 79, 82, 109).

[416] Li, Z.; Yuan, J. In *Cloud Comput. Secur.* Ed. by Sun, X.; Chao, H.-C.; You, X.; Bertino, E., Springer International Publishing: Cham, 2017, pp 85–97, DOI: 10.1007/978-3-319-68505-2_8 (pages 79, 109).

[417] Häner, T.; Steiger, D. S. In *Proc. Int. Conf. High Perform. Comput. Netw. Storage Anal.* Association for Computing Machinery: New York, NY, USA, 2017, pp 1–10, DOI: 10.1145/3126908.3126947 (pages 79, 83, 108).

[418] Wu, X.-C.; Di, S.; Dasgupta, E. M.; Cappello, F.; Finkel, H.; Alexeev, Y.; Chong, F. T. In *Proc. Int. Conf. High Perform. Comput. Netw. Storage Anal.* Association for Computing Machinery: New York, NY, USA, 2019, pp 1–24, DOI: 10.1145/3295500.3356155 (pages 79, 83, 108).

[419] Pednault, E.; Gunnels, J. A.; Nannicini, G.; Horesh, L.; Magerlein, T.; Solomonik, E.; Draeger, E. W.; Holland, E. T.; Wisnieff, R. Pareto-Efficient Quantum Circuit Simulation Using Tensor Contraction Deferral preprint (pages 79, 83, 106, 108).

[420] Doi, J.; Takahashi, H.; Raymond, R.; Imamichi, T.; Horii, H. In *Proc. 16th ACM Int. Conf. Comput. Front.* Association for Computing Machinery: New York, NY, USA, 2019, pp 85–93, DOI: 10.1145/3310273.3323053 (pages 79, 83, 85, 109).

[421] Aleksandrowicz, G.; Alexander, T.; Barkoutsos, P.; Bello, L.; Ben-Haim, Y.; Bucher, D.; Cabrera-Hernández, F. J.; Carballo-Franquis, J.; Chen, A.; Chen, C.-F.; Chow, J. M.; Córcoles-Gonzales, A. D.; Cross, A. J.; Cross, A.; Cruz-Benito, J.; Culver, C.; González, S. D. L. P.; Torre, E. D. L.; Ding, D.; Dumitrescu, E.; Duran, I.; Eendebak, P.; Everitt, M.; Sertage, I. F.; Frisch, A.; Fuhrer, A.; Gambetta, J.; Gago, B. G.; Gomez-Mosquera, J.; Greenberg, D.; Hamamura, I.; Havlicek, V.; Hellmers, J.; Herok, Ł.; Horii, H.; Hu, S.; Imamichi, T.; Itoko, T.; Javadi-Abhari, A.; Kanazawa, N.; Karazeev, A.; Krsulich, K.; Liu, P.; Luh, Y.; Maeng, Y.; Marques, M.; Martín-Fernández, F. J.; McClure, D. T.; McKay, D.; Meesala, S.; Mezzacapo, A.; Moll, N.; Rodríguez, D. M.; Nannicini, G.; Nation, P.; Ollitrault, P.; O'Riordan, L. J.; Paik, H.; Pérez, J.; Phan, A.; Pistoia, M.; Prutyanov, V.; Reuter, M.; Rice, J.; Davila, A. R.; Rudy, R. H. P.; Ryu, M.; Sathaye, N.; Schnabel, C.; Schoute, E.; Setia, K.; Shi, Y.; Silva, A.; Siraichi, Y.; Sivarajah, S.; Smolin, J. A.; Soeken, M.; Takahashi, H.; Tavernelli, I.; Taylor, C.; Taylour, P.; Trabing, K.; Treinish, M.; Turner, W.; Vogt-Lee, D.; Vuillot, C.; Wildstrom, J. A.; Wilson, J.; Winston, E.; Wood, C.; Wood, S.; Wörner, S.; Akhalwaya, I. Y.; Zoufal, C. Qiskit: An Open-Source Framework for Quantum Computing, version 0.7.2, Zenodo, 2019, DOI: 10.5281/zenodo.2562111 (pages 80, 83, 103).

[422] team, Q. A.; collaborators Qsim, Zenodo, 2020, DOI: 10.5281/zenodo.4023103 (pages 80, 107).

[423] QDK-Python, Microsoft, 2023 (pages 80, 83, 107).

[424] Nielsen, M. A.; Chuang, I. L., *Quantum Computation and Quantum Information: 10th Anniversary Edition*; Cambridge University Press: Cambridge, 2010, DOI: 10.1017/CBO9780511976667 (pages 80, 82).

[425] De Raedt, K.; Michielsen, K.; De Raedt, H.; Trieu, B.; Arnold, G.; Richter, M.; Lippert, Th.; Watanabe, H.; Ito, N. Massively Parallel Quantum Computer Simulator. *Computer Physics Communications* **2007**, *176*, 121–136, DOI: 10.1016/j.cpc.2006.08.007 (pages 82, 83).

[426] Aaronson, S.; Gottesman, D. Improved Simulation of Stabilizer Circuits. *Phys. Rev. A* **2004**, *70*, 052328, DOI: 10.1103/PhysRevA.70.052328 (pages 82, 83).

[427] McCaskey, A.; Dumitrescu, E.; Chen, M.; Lyakh, D.; Humble, T. Validating Quantum-Classical Programming Models with Tensor Network Simulations. *PLOS ONE* **2018**, *13*, e0206704, DOI: 10.1371/journal.pone.0206704 (page 83).

[428] Huang, C.; Zhang, F.; Newman, M.; Cai, J.; Gao, X.; Tian, Z.; Wu, J.; Xu, H.; Yu, H.; Yuan, B.; Szegedy, M.; Shi, Y.; Chen, J. Classical Simulation of Quantum Supremacy Circuits preprint (pages 83, 108).

[429] Lykov, D.; Schutski, R.; Galda, A.; Vinokur, V.; Alexeev, Y. In *2022 IEEE Int. Conf. Quantum Comput. Eng. QCE*, 2022 IEEE International Conference on Quantum Computing and Engineering (QCE), 2022, pp 582–593, DOI: 10.1109/QCE53715.2022.00081 (page 83).

[430] Markov, I. L.; Shi, Y. Simulating Quantum Computation by Contracting Tensor Networks. *SIAM J. Comput.* **2008**, *38*, 963–981, DOI: 10.1137/050644756 (page 83).

[431] Pednault, E.; Gunnels, J. A.; Nannicini, G.; Horesh, L.; Wisnieff, R. Leveraging Secondary Storage to Simulate Deep 54-Qubit Sycamore Circuits preprint (pages 83, 108).

[432] Smelyanskiy, M.; Sawaya, N. P. D.; Aspuru-Guzik, A. qHiPSTER: The Quantum High Performance Software Testing Environment preprint (pages 83, 106, 108).

[433] Li, R.; Wu, B.; Ying, M.; Sun, X.; Yang, G. Quantum Supremacy Circuit Simulation on Sunway TaihuLight preprint (page 83).

[434] Amariutei, A.; Caraiman, S. In *15th Int. Conf. Syst. Theory Control Comput.* 15th International Conference on System Theory, Control and Computing, 2011, pp 1–6 (pages 83, 109).

[435] Stair, N. H.; Huang, R.; Evangelista, F. A. A Multireference Quantum Krylov Algorithm for Strongly Correlated Electrons. *J. Chem. Theory Comput.* **2020**, *16*, 2236–2245, DOI: 10.1021/acs.jctc.9b01125 (page 84).

[436] Farhi, E.; Goldstone, J.; Gutmann, S. A Quantum Approximate Optimization Algorithm preprint (page 84).

[437] Koh, D. E. Further Extensions of Clifford Circuits and Their Classical Simulation Complexities. *QIC* **2017**, *17*, 262–282, DOI: 10.26421/QIC17.3-4-5 (page 84).

[438] Hein, M.; Dür, W.; Eisert, J.; Raussendorf, R.; Van den Nest, M.; Briegel, H.-J. Entanglement in Graph States and Its Applications. *Quantum Comput. Algorithms Chaos* **2006**, 115–218, DOI: 10.3254/978-1-61499-018-5-115 (page 84).

[439] Bravyi, S.; Gosset, D.; König, R. Quantum Advantage with Shallow Circuits. *Science* **2018**, *362*, 308–311, DOI: 10.1126/science.aar3106 (page 84).

[440] JavadiAbhari, A.; Patil, S.; Kudrow, D.; Heckey, J.; Lvov, A.; Chong, F. T.; Martonosi, M. ScaffCC: Scalable Compilation and Analysis of Quantum Programs. *Parallel Computing* **2015**, *45*, 2–17, DOI: 10.1016/j.parco.2014.12.001 (page 84).

[441] Bremner, M. J.; Jozsa, R.; Shepherd, D. J. Classical Simulation of Commuting Quantum Computations Implies Collapse of the Polynomial Hierarchy. *Proc. R. Soc. Math. Phys. Eng. Sci.* **2010**, *467*, 459–472, DOI: 10.1098/rspa.2010.0301 (pages 84, 85).

[442] Bremner, M. J.; Montanaro, A.; Shepherd, D. J. Average-Case Complexity Versus Approximate Simulation of Commuting Quantum Computations. *Phys. Rev. Lett.* **2016**, *117*, 080501, DOI: 10.1103/PhysRevLett.117.080501 (pages 84, 85).

[443] Gilliam, A.; Woerner, S.; Gonciulea, C. Grover Adaptive Search for Constrained Polynomial Binary Optimization. *Quantum* **2021**, *5*, 428, DOI: 10.22331/q-2021-04-08-428 (pages 84, 85).

[444] Simons Collaboration on the Many-Electron Problem; Motta, M.; Genovese, C.; Ma, F.; Cui, Z.-H.; Sawaya, R.; Chan, G. K.-L.; Chepiga, N.; Helms, P.; Jiménez-Hoyos, C.; Millis, A. J.; Ray, U.; Ronca, E.; Shi, H.; Sorella, S.; Stoudenmire, E. M.; White, S. R.; Zhang, S. Ground-State Properties of the Hydrogen Chain: Dimerization, Insulator-to-Metal Transition, and Magnetic Phases. *Phys. Rev. X* **2020**, *10*, 031058, DOI: 10.1103/PhysRevX.10.031058 (page 84).

[445] Al-Saidi, W. A.; Zhang, S.; Krakauer, H. Bond Breaking with Auxiliary-Field Quantum Monte Carlo. *J. Chem. Phys.* **2007**, *127*, 144101, DOI: 10.1063/1.2770707 (page 84).

[446] Hachmann, J.; Cardoen, W.; Chan, G. K.-L. Multireference Correlation in Long Molecules with the Quadratic Scaling Density Matrix Renormalization Group. *J. Chem. Phys.* **2006**, *125*, 144101, DOI: 10.1063/1.2345196 (page 84).

[447] Baiardi, A.; Reiher, M. The Density Matrix Renormalization Group in Chemistry and Molecular Physics: Recent Developments and New Challenges. *J. Chem. Phys.* **2020**, *152*, 040903, DOI: 10.1063/1.5129672 (page 84).

[448] Shen, Y.; Zhang, X.; Zhang, S.; Zhang, J.-N.; Yung, M.-H.; Kim, K. Quantum Implementation of the Unitary Coupled Cluster for Simulating Molecular Electronic Structure. *Phys. Rev. A* **2017**, *95*, 020501, DOI: 10.1103/PhysRevA.95.020501 (page 84).

[449] Bouland, A.; Fefferman, B.; Nirkhe, C.; Vazirani, U. On the Complexity and Verification of Quantum Random Circuit Sampling. *Nature Phys* **2019**, *15*, 159–163, DOI: 10.1038/s41567-018-0318-2 (page 84).

[450] Aschauer, H.; Dür, W.; Briegel, H.-J. Multiparticle Entanglement Purification for Two-Colorable Graph States. *Phys. Rev. A* **2005**, *71*, 012319, DOI: 10.1103/PhysRevA.71.012319 (page 84).

[451] *Quantum Error Correction*; Lidar, D. A., Brun, T. A., Eds.; Cambridge University Press: Cambridge, 2013, DOI: 10.1017/CBO9781139034807 (page 84).

[452] Shor, P. W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Rev.* **1999**, *41*, 303–332, DOI: 10.1137/S0036144598347011 (page 84).

[453] NVIDIA Profiler Users Guide https://docs.nvidia.com/cuda/profiler-users-guide/ (page 85).

[454] Li, G.; Ding, Y.; Xie, Y. In *Proc. Twenty-Fourth Int. Conf. Archit. Support Program. Lang. Oper. Syst.* Association for Computing Machinery: New York, NY, USA, 2019, pp 1001–1014, DOI: 10.1145/3297858.3304023 (page 96).

[455] Iten, R.; Moyard, R.; Metger, T.; Sutter, D.; Woerner, S. Exact and Practical Pattern Matching for Quantum Circuit Optimization. *ACM Transactions on Quantum Computing* **2022**, *3*, 4:1–4:41, DOI: 10.1145/3498325 (pages 96, 109).

[456] O'Neil, M. A.; Burtscher, M. In *Proc. Fourth Workshop Gen. Purp. Process. Graph. Process. Units*, Association for Computing Machinery: New York, NY, USA, 2011, pp 1–7, DOI: 10.1145/1964179.1964189 (page 101).

[457] Cross, A. W.; Bishop, L. S.; Smolin, J. A.; Gambetta, J. M. Open Quantum Assembly Language preprint (page 107).

[458] Zulehner, A.; Wille, R. Advanced Simulation of Quantum Computations. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2019**, *38*, 848–859, DOI: 10.1109/TCAD.2018.2834427 (page 108).

[459] Shi, Y.; Leung, N.; Gokhale, P.; Rossi, Z.; Schuster, D. I.; Hoffmann, H.; Chong, F. T. In *Proc. Twenty-Fourth Int. Conf. Archit. Support Program. Lang. Oper. Syst.* Association for Computing Machinery: New York, NY, USA, 2019, pp 1031–1044, DOI: 10.1145/3297858.3304018 (page 109).

[460] Gutierrez, E.; Romero, S.; Trenas, M. A.; Zapata, E. L. In *Proc. 7th Conf. 7th WSEAS Int. Conf. Syst. Theory Sci. Comput. - Vol. 7*, World Scientific and Engineering Academy and Society (WSEAS): Stevens Point, Wisconsin, USA, 2007, pp 121–126 (page 109).

[461] Zhang, P.; Yuan, J.; Lu, X. In *Algorithms Archit. Parallel Process.* Ed. by Wang, G.; Zomaya, A.; Martinez, G.; Li, K., Springer International Publishing: Cham, 2015, pp 241–256, DOI: 10.1007/978-3-319-27119-4_17 (page 109).

[462] Avila, A.; Maron, A.; Reiser, R.; Pilla, M.; Yamin, A. In *Proc. 29th Annu. ACM Symp. Appl. Comput.* Association for Computing Machinery: New York, NY, USA, 2014, pp 860–865, DOI: 10.1145/2554850.2554892 (page 109).