

UC Berkeley

UC Berkeley Previously Published Works

Title

Simulating the absorption spectra of helium clusters (N = 70, 150, 231, 300) using a charge transfer correction to superposition of fragment single excitations.

Permalink

<https://escholarship.org/uc/item/2p45b9gs>

Journal

The Journal of chemical physics, 146(4)

ISSN

0021-9606

Authors

Ge, Qinghui
Mao, Yuezhi
White, Alec F
[et al.](#)

Publication Date

2017

DOI

10.1063/1.4973611

Peer reviewed

Simulating the absorption spectra of helium clusters ($N=70, 150, 231, 300$) using a charge-transfer correction to superposition of fragment single excitations

Qinghui Ge,^{1,2} Yuezhi Mao,¹ Alec F. White,^{1,2} Evgeny Epifanovsky,³ Kristina D. Closser,⁴ and Martin Head-Gordon^{1,2}

¹*Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, CA 94720, USA*

²*Chemical Sciences Division, Lawrence Berkeley National Laboratory Berkeley, CA 94720, USA*

³*Q-Chem Inc., 6601 Owens Drive, Suite 105, Pleasanton, CA 94588, USA*

⁴*Molecular Foundry, Lawrence Berkeley National Laboratory Berkeley, CA 94720, USA*

(Dated: 19 December 2016)

Simulations of the $n = 2$ absorption spectra of He_N ($N = 70, 150, 231, 300$) clusters are reported, with nuclear configurations sampled by path integral molecular dynamics. The electronic structure is treated by a new approach, ALMO-CIS+CT, which is a formulation of configuration interaction singles (CIS) based on absolutely-localized molecular orbitals (ALMOs). The method generalizes the previously reported ALMO-CIS model (*J. Chem. Theory Comput.* **11**, 5791 (2015)) to include spatially localized charge transfer (CT) effects. It is designed to recover large numbers of excited states in atomic and molecular clusters, such as the entire $n = 2$ Rydberg band in helium clusters. ALMO-CIS+CT is shown to recover most of the error caused by neglecting charge transfer in ALMO-CIS, and has comparable accuracy to standard CIS for helium clusters. **For the $n = 2$ band, CT stabilizes states towards the blue edge by up to 0.5 eV.** ALMO-CIS+CT retains the formal cubic scaling of ALMO-CIS with respect to system size. With improvements to the implementation over that originally reported for ALMO-CIS, ALMO-CIS+CT is able to treat helium clusters with hundreds of atoms using modest computing resources. A detailed simulation of the absorption spectra associated with the $2s$ and $2p$ bands of helium clusters up to 300 atoms is reported, using path integral molecular dynamics with a spherical boundary condition to generate atomic configurations at 3 K. The main features of experimentally reported fluorescence excitation spectra for helium clusters are reproduced.

I. INTRODUCTION

Atomic and molecular clusters span the range between the gas phase and bulk limits, and they are useful for investigating fundamental differences between bulk and surface properties. Helium clusters are of particular interest: they are weakly interacting, superfluidic¹ and can be used as a spectroscopic medium^{2,3}. Experimentally, fluorescence spectra⁴⁻⁶ provide a way to understand how the excited electronic states are modified when a helium atom is adjacent to the others. Typical spectra of helium clusters consist of sharp atomic lines accompanied by blue-shifted wings, and their relative intensities depend on the cluster sizes.

Even for helium, the simplest many-electron atom, theoretical study of the excited electronic states remains a challenge for large clusters. A method must balance accuracy against efficiency to

be practically useful. Two of the simplest methods are time-dependent density functional theory (TDDFT)⁷⁻¹⁰ and configuration interaction singles (CIS)^{9,11}. TDDFT is more widely used, because in many cases where the dynamic electron correlation is important, it outperforms CIS in terms of accuracy. CIS often has quite large (~ 1 eV) errors in excitation energies due to the neglect of dynamic electron correlation.

Nonetheless, CIS does not suffer from the incorrect asymptotic potential and self-interaction error of TDDFT. For this reason, for relative energies of Rydberg states in helium clusters, CIS is the method of choice. Moreover, CIS can be improved systematically by adding higher substitutions to the CI wavefunction¹². Previous CIS calculations on small clusters involving up to 25 helium atoms shows that CIS is capable of elucidating the spectroscopy¹³ and the post-excitation dynamics of helium clusters¹⁴.

However, modern experiments mainly focus on helium clusters with hundreds to tens of thousands of atoms or even larger, which is beyond the normal capability of CIS due to its high order scaling with respect to system size: a conventional CIS calculation for all states scales as $O(M^6)$, with the number of (identical) atoms or molecules, M , in the cluster. When only a handful of the low-lying excited states are requested, the CIS eigen-equation can be solved iteratively using Davidson's method^{15,16}, which reduces the scaling to $O(M^4)$ in the molecular orbital (MO) basis per state. In the atomic orbital (AO) basis, matrix element sparsity reduces the cost to as low as $O(M^2)$ per state (with a large prefactor) for the rate-determining matrix-vector contraction.

Unfortunately, to directly compare with the experimental spectra of large homogeneous clusters, we need a full description of the energy bands of He clusters. The number of states required then grows at least linearly with the number of atoms (for instance, the $n = 2$ band of a 1000-atom helium cluster requires a minimum of 4000 states), suggesting that iterative methods are not preferred. On the other hand, direct solution for all states is not feasible, as already discussed.

Great effort has gone into reducing the scaling of CIS and TDDFT methods. By exploiting certain types of spatial locality or sparsity in matrix representations, some linear scaling methods have been developed¹⁷⁻¹⁹. However, rigorous linear scaling is only achieved in substantially large systems, especially since **the electronic density of excited states is often much more delocalized than that of the ground state (here we mainly focus on systems whose excitations cannot be localized in a certain region, such as homogeneous clusters)**. For this reason, these linear scaling methods are not fully applicable for systems consisting of hundreds of atoms, and different approximations seem to be desirable for these cases. A common feature of the above mentioned methods is that the localized molecular orbitals (LMOs) are obtained by localizing canonical molecular orbitals (CMOs) into confined physical regions. This "top-down" scheme is often found to be inefficient due to the difficulty of localizing virtual orbitals when system size increases^{20,21}.

Fragment-based methods, by contrast, follow a "bottom-up" scheme, where LMOs are obtained directly from subsystem calculations without computing the CMOs first. One example is the fragment molecular orbital (FMO) method developed by Kitaura and coworkers²². The idea is to divide the system into fragments and perform ab initio calculations of fragments and their dimers. The method

was extended to excited state calculations by employing the multilayer FMO method, in which the region of chemical interest is treated with CIS²³, CIS(D)²⁴ or TDDFT²⁵, while the environment is kept at the HF/DFT level. Wu *et al.* proposed a linear scaling TDDFT method through the use of fragment LMOs which are orthogonal but still well localized²⁶. More recently, Herbert and co-workers implemented the Frenkel-Davydov exciton model to study the excited states of aggregates, where an excitonic state is constructed from direct products of fragment configuration state function basis²⁷⁻²⁹.

Another bottom-up method, which is fully self-consistent, is the approach based on absolutely localized molecular orbitals (ALMOs), which were first introduced to speed up SCF calculations on weakly interacting systems³⁰⁻³⁴. ALMOs are defined by the constraint that the MO coefficient matrix is block diagonal between fragments. This type of constrained SCF procedure is commonly referred to as SCF for molecular interactions (SCF-MI). Subsequently ALMOs from SCF-MI have also proven useful in the development of energy decomposition analysis (EDA) methods³⁵⁻³⁹.

In a previous publication, we reported ALMO-CIS⁴⁰, an ALMO-based excited state method, and its application to helium clusters. The ALMO-CIS method scales as only $O(M^3)$ for the evaluation of $O(M)$ states. Two factors contribute to the reduction of the scaling: (a) The molecular orbitals (MOs) are linear combination of only AOs centered on a certain fragment, which greatly reduces the cost of transforming the electron-repulsion integrals (ERIs) from AOs to MOs; (b) the locality of MOs allows us to associate a single substitution to fragments, and the CIS equation can be truncated in a physically meaningful manner. In the ALMO-CIS model, only intrafragment single excitations are considered, and this reduces the matrix dimension from $O(M^2)$ to $O(M)$ in the CIS equation and thereby reduces the scaling of the method.

A TDDFT(MI) method has also been recently proposed by Liu and Herbert⁴¹, which shares the same spirit with ALMO-CIS. The major difference is that TDDFT(MI) computes several lowest excited states for the monomers first, then evaluates the supersystem excited states as the linear combination of these local states, using the Davidson algorithm. This current implementation is optimized for systems such as solvated chromophores, where a relatively small number of the excited states are of interest. With some modification of the algorithm, TDDFT(MI) could also be extended to obtain the full spectrum of molecular clusters.

Both ALMO-CIS and TDDFT(MI) restrict the excitations to be intrafragment, which means the charge transfer (CT) effect is neglected. For helium clusters, the ALMO-CIS method is found to have ~ 0.5 eV overestimation for excitation energies at the blue end of the $n = 2$ band, when measured against standard CIS. It also has systematic error for predicting the spectrum profiles⁴⁰. For many other systems with stronger interaction than helium, we expect the error will be more pronounced. All these considerations motivate us to seek a way to correct the ALMO-CIS model by at least partially reintroducing the neglected charge transfer class of excitations.

In this work, we add back the charge transfer effect by using a real-space distance cutoff. The model presented here is called ALMO-CIS+CT. The remainder of the paper is organized as follows: in the Theory section, we introduce the generalization of the ALMO-CIS model to ALMO-CIS+CT, and a Davidson-like variational method that is used to solve the eigenvalue problem. Next, we discuss some optimization we have done for the fast implementation of the ALMO-CIS+CT model. The accuracy and timing results are presented in the Result section. Finally, we apply ALMO-CIS+CT for the study of helium clusters, and the resulting spectra are compared with the experimental data. A brief description of a more rigorous simulation for the cluster geometries with path integral molecular dynamics can also be found in the Application section.

II. THEORY

A. Notation

The following notation is used throughout the paper. i, j, k, l : occupied MO indices; a, b : virtual MO indices; p, q, r : generic MO indices. χ_μ : atomic orbitals; χ_P, χ_Q : auxiliary basis functions. The ALMOs are denoted by ψ , and ϕ is used for projected virtual orbitals. Fragments are indicated by I, J or F_I, F_J . Unless otherwise specified, two orbital indices connected as \overline{pq} or $\overline{p\bar{q}}$ belong to two near-neighbor fragments (or the same fragment).

We use capital letters to indicate quantities that scale with system size, and lowercase letters for quantities that do not: O, V, N : total number of occupied/virtual/atomic orbitals in the system; o, v, n : the average numbers of occupied/virtual/atomic orbitals per fragment; $\bar{o}, \bar{v}, \bar{n}$: the average numbers of occupied/virtual/atomic orbitals within the near-

neighbor fragments defined by the distance criterion. For a system with identical fragments, the number of fragments M can be used to denote the size of the system.

We use standard tensor notation to work with nonorthogonal functions⁴². A covariant function is denoted by a subscript and a contravariant function is denoted by a superscript. The Einstein summation convention is also employed, where an index that occurs once covariant and once contravariant implies a sum.

B. ALMO-CIS and its generalization to include charge transfer

Let us begin by defining the ALMOs. The atoms in an atomic cluster (or molecules in a molecular cluster) are divided into **non-overlapping** subsets that are referred to as fragments. The atom centered AOs can thus be partitioned based on the fragments they belong to. In the ALMO formalism, each molecular orbital (MO) on a given fragment is a superposition of AOs centered on the same fragment exclusively, and this results in a block-diagonal MO coefficient matrix.

$$|\psi_p\rangle = \sum_{\mu \in F_I} |\chi_\mu\rangle C_p^\mu, \quad p \in F_I \quad (1)$$

The MO coefficients can be solved within the framework of self-consistent field theory with the constraint that the MO coefficient matrix should be fragment-blocked. The resulting ALMOs are orthogonal within a fragment but are nonorthogonal between fragments.

In the previous ALMO-CIS publication⁴⁰, we have derived the generalized CIS equations that apply to nonorthogonal molecular orbitals:

$$A_{ia,jb} t^{jb} = \omega_{CIS} S_{ab} S_{ij} t^{jb} \quad (2)$$

Here $\omega_{CIS} = E - E_{HF}$ is the excitation energy, and t^{jb} are the CIS amplitudes. The overlap metric S appears because of the interfragment nonorthogonality of ALMOs. We have chosen known **matrix elements** to be covariant, and the unknown amplitudes to be contravariant. **The (restricted) CIS Hamiltonian can be constructed from the Fock matrices f , overlap metric S and the two-electron integrals:**

$$A_{ia,jb} = f_{ab} S_{ij} - f_{ij} S_{ab} + 2(\psi_i \phi_a | \phi_b \psi_j) - (\psi_i \psi_j | \phi_a \phi_b) \quad (3)$$

All virtual orbitals above are technically “projected virtals”, which are defined by projecting out the

occupied space from the ALMO virtuals to ensure that occupieds and virtuals are orthogonal (the unprojected ALMO virtuals on a fragment are non-orthogonal to both occupied and virtual ALMOs on other fragments).

$$\begin{aligned} |\phi_a\rangle &= N_a(|\psi_a\rangle - \hat{P}_{occ}|\psi_a\rangle) \\ &= N_a(|\psi_a\rangle - |\psi_k\rangle(S^{-1})^{kl}\langle\psi_l|\psi_a\rangle) \\ &= N_a(|\psi_a\rangle - |\psi_k\rangle\langle\psi^k|\psi_a\rangle) \end{aligned} \quad (4)$$

where N_a is the normalization constant and $|\psi^k\rangle = |\psi_l\rangle(S^{-1})^{lk}$ are the contravariant occupied orbitals.

In the ALMO-CIS model, the CIS matrices are truncated by including only ‘‘intra-fragment’’ single substitutions (i.e. those that promote an electron from an occupied level to a virtual level assigned to same fragment), which can be denoted as follows:

$$A_{\underline{ia},\underline{jb}}t^{\underline{jb}} = \omega_{CIS}S_{\underline{ab}}S_{\underline{ij}}t^{\underline{jb}} = \omega_{CIS}G_{\underline{ia},\underline{jb}}t^{\underline{jb}} \quad (5)$$

G is the metric associated with the retained intra-fragment single substitutions. **It is intuitive that the ALMO-CIS model is free of charge-transfer contributions. As shown in the appendix, this can also be proven, in the sense that fragment Mulliken populations are unchanged from the ground state in the ALMO-CIS model.**

The ALMO-CIS equation, Eq. 5 can be generalized to include some charge transfer (CT) type single substitutions based on a selection of significant fragment pairs within which CT will be permitted. Specifically, we use a distance based cutoff (r_{cut}) to determine whether two fragments are to be considered as neighboring fragments. A significant fragment pair list can be created, comprising all pairs of fragments whose distances are smaller than r_{cut} . Pairs of a same fragment repeated twice are considered as zero distance pairs and thus are also included in the list.

As a result, when a single substitution has corresponding occupied and virtual orbitals that belong to a pair of fragments in the pair list, we will include it in the truncated matrices. In this way, the form of the working ALMO-CIS equation (eq. 5) remains the same, but the meaning of the contraction lines is generalized. In the ALMO-CIS model, a contraction line connects two indices that belong to a same fragment. Now, in the generalization that we will refer to as ALMO-CIS+CT, the two indices belong to a significant pair of fragments.

The cutoff distance is a user-defined parameter in the ALMO-CIS+CT model. At one extreme, in the

limit of $r_{cut} \rightarrow 0$, the fragment pair list will only include the pairs that contain the same fragment twice, and ALMO-CIS+CT reduces to the original ALMO-CIS model. At the other extreme, when a very large r_{cut} is chosen, all possible single substitutions are included, and the untruncated CIS equations (eq. 2) are recovered. For a weakly interacting system, such as a helium cluster, a cutoff that corresponds to including the first shell of neighboring atoms will be shown (in the Results section) to be sufficient to recover most of the error caused by neglecting CT in ALMO-CIS.

The oscillator strength of an excited state κ is defined as:

$$f_\kappa = \frac{2}{3}\omega_\kappa|\langle\psi_0|\hat{\mu}|\psi_\kappa\rangle|^2 \quad (6)$$

Within the ALMO-CIS+CT theory, the dipole matrix elements can be calculated as:

$$\begin{aligned} \langle\psi_0|\hat{\mu}|\psi_\kappa\rangle &= \sum_{\substack{F_I, F_A \in \\ \text{FrgPairList}}} \sum_{\substack{i \in F_I \\ a \in F_A}} t_\kappa^{ia} \langle\psi_i|\hat{\mu}|\phi_a\rangle \\ &= \sum_{\substack{F_I, F_A \in \\ \text{FrgPairList}}} \sum_{\substack{i \in F_I \\ a \in F_A}} t_\kappa^{ia} c_i^{\dagger\mu} \mu_{\mu\nu} c_{\nu a}^\nu \end{aligned} \quad (7)$$

These matrix elements will be used later in evaluating the absorption spectrum of helium clusters.

C. Davidson-like variational method

The eigenvalue problem of eq. 5 can be solved by a full diagonalization. In the ALMO-CIS model, this $O(M^3)$ scaling step was found to be a minor step because of its relatively small prefactor. However, when CT substitutions are included, the computational effort will increase by the cube of the factor by which the number of single substitutions has increased. For the nearest-neighbor cutoff, applied to a medium-sized helium cluster, the ALMO-CIS+CT matrix size is roughly five to six times larger than for ALMO-CIS, and as a result the diagonalization timing is about 200 times longer. Full diagonalization is still feasible, but becomes a dominant step. Thus, we propose a one-step Davidson-like variational method to alleviate this problem.

The eigenvalue problem of eq. 5 is now expressed in terms of the intra-fragment (local) subspace (denoted by l) and the charge transfer subspace (denoted by c). In the matrix elements below, each index l or c corresponds to an occupied-virtual pair ia .

The dimensions of the two subspaces are denoted by N_l and N_c , respectively. Likewise, the trailing state index in the amplitudes, \mathbf{t} , and eigenvalues, ω , can be associated with a state that is primarily local (l) or primarily CT (c).

$$\begin{bmatrix} \mathbf{A}_{ll} & \mathbf{A}_{lc} \\ \mathbf{A}_{cl} & \mathbf{A}_{cc} \end{bmatrix} \begin{bmatrix} \mathbf{t}_{ll} & \mathbf{t}_{lc} \\ \mathbf{t}_{cl} & \mathbf{t}_{cc} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{ll} & \mathbf{G}_{lc} \\ \mathbf{G}_{cl} & \mathbf{G}_{cc} \end{bmatrix} \begin{bmatrix} \mathbf{t}_{ll} & \mathbf{t}_{lc} \\ \mathbf{t}_{cl} & \mathbf{t}_{cc} \end{bmatrix} \begin{bmatrix} \omega_l & \mathbf{0} \\ \mathbf{0} & \omega_c \end{bmatrix} \quad (8)$$

Solving the above full-size eigenvalue problem can be avoided since we are wanting to obtain corrected roots only for the ALMO-CIS states (i.e. only for the “ l ” block). We will show that the dimension of the eigenvalue problem to be solved for those corrected roots can be reduced to just $2 \times N_l$ by using a one-step Davidson-like method.

We first solve an eigenvalue problem in the local subspace, i.e., solve $\mathbf{A}_{ll}\mathbf{t}_{ll}^{(0)} = \mathbf{G}_{ll}\mathbf{t}_{ll}^{(0)}\omega_l^{(0)}$ and obtain $\mathbf{t}_{ll}^{(0)}$ and $\omega_l^{(0)}$ as the uncorrected eigenvectors and eigenvalues. They are simply the ALMO-CIS eigenvectors and eigenvalues. Next, we form a transformation to a projected CT basis $\tilde{\mathbf{C}}$ that is constructed to be orthogonal to the local subspace.

$$\tilde{\mathbf{C}} = \begin{bmatrix} -\mathbf{t}_{ll}^{(0)\dagger}\mathbf{G}_{lc} \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{I} \end{bmatrix} \quad (9)$$

where \mathbf{I} is the $N_c \times N_c$ identity matrix. With the projected CT basis, the eigenvalue problem becomes:

$$\begin{bmatrix} \tilde{\mathbf{A}}_{ll} & \tilde{\mathbf{A}}_{lc} \\ \tilde{\mathbf{A}}_{cl} & \tilde{\mathbf{A}}_{cc} \end{bmatrix} \begin{bmatrix} \mathbf{t}_{ll} & \mathbf{t}_{lc} \\ \mathbf{t}_{cl} & \mathbf{t}_{cc} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{G}}_{ll} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{G}}_{cc} \end{bmatrix} \begin{bmatrix} \mathbf{t}_{ll} & \mathbf{t}_{lc} \\ \mathbf{t}_{cl} & \mathbf{t}_{cc} \end{bmatrix} \begin{bmatrix} \omega_l & \mathbf{0} \\ \mathbf{0} & \omega_c \end{bmatrix} \quad (10)$$

where the projected blocks are given by

$$\begin{aligned} \tilde{\mathbf{A}}_{ll} &= \mathbf{A}_{ll}; & \tilde{\mathbf{A}}_{lc} &= \mathbf{A}_{ll}\mathbf{C} + \mathbf{A}_{lc}; \\ \tilde{\mathbf{A}}_{cc} &= \tilde{\mathbf{C}}^\dagger\mathbf{A}\tilde{\mathbf{C}} = \mathbf{C}^\dagger\mathbf{A}_{ll}\mathbf{C} + \mathbf{A}_{cl}\mathbf{C} + \mathbf{C}^\dagger\mathbf{A}_{lc} + \mathbf{A}_{cc} \\ \tilde{\mathbf{G}}_{ll} &= \mathbf{G}_{ll}; & \tilde{\mathbf{G}}_{lc} &= \mathbf{0} \\ \tilde{\mathbf{G}}_{cc} &= \tilde{\mathbf{C}}^\dagger\mathbf{S}\tilde{\mathbf{C}} = \mathbf{C}^\dagger\mathbf{G}_{ll}\mathbf{C} + \mathbf{G}_{cl}\mathbf{C} + \mathbf{C}^\dagger\mathbf{G}_{lc} + \mathbf{G}_{cc} \end{aligned} \quad (11)$$

Similar to the Davidson method, we compute the correction vectors by

$$\boldsymbol{\delta} = (\tilde{\mathbf{A}}_{cc} - \bar{\omega}\tilde{\mathbf{G}}_{cc})^{-1}\mathbf{r} \quad (12)$$

where $\mathbf{r} = \tilde{\mathbf{A}}_{cl}\mathbf{t}_{ll}^{(0)}$ are the residue vectors of the uncorrected ALMO-CIS excited states. The scalar, $\bar{\omega}$, involved in the preconditioner is a parameter which should be chosen based on the states that one is interested in, and it should be close to the target eigenvalues. For example, in our study of helium clusters, we target the $n = 2$ manifold of states, and thus $\bar{\omega}$ is approximated by the average of the

uncorrected $n = 2$ eigenenergies $\omega_l^{(0)}$, so that separate preconditioning for each state can be avoided. It is thus required to solve an N_c -dimensional linear equation once to apply the preconditioner in Eq. 12. One can avoid this third-order scaling step by considering keeping only diagonal elements of the preconditioner. However, the cost of solving the linear equation has a quite small prefactor. In fact, for the helium cluster systems we have studied, it is actually not the dominant step in the whole Davidson-like procedure. Therefore, the reduction of computational cost from a simplified preconditioner does not seem to be worthwhile, considering the diminished accuracy that comes with it. Detailed timings for the Davidson-like method and tests showing the accuracy of different types of preconditioners are reported in the Results section.

We can now attach the correction vectors to the local subspace, and form \mathbf{A} and \mathbf{G} in a $2 \times N_l$ subspace spanned by $\mathbf{t}_{ll}^{(0)}$ and $\boldsymbol{\delta}$. The result is a generalized eigenvalue problem of dimension $2 \times N_l$, which is typically 3 times smaller than that of the full ALMO-CIS+CT model. The eigenstates and eigenvalues corresponding to intrafragment excited states (now corrected for CT) can be obtained by taking the lowest N_l eigen-solutions. This approach is like a one-step Davidson method, and the resulting excitation energies are variational upper bounds to the exact ALMO-CIS+CT eigenvalues that come from solving Eq. 5 exactly.

III. EFFICIENT IMPLEMENTATION AND SCALING

An efficient implementation of the ALMO-CIS+CT model has been completed within a development version of Q-Chem quantum chemistry program package^{43,44}. With CT states included, the size of the truncated Hamiltonian and overlap metric will increase by a factor that depends on the cut-off distance one chooses. However, the storage for these matrices still scales as $O(M^2)$. For the size of systems we have studied, this fact allows the matrices to be explicitly calculated and stored in memory, and Eq. 5 to be solved as a final step without memory issues. Thus, the algorithms for building the Hamiltonian reported in the implementation of the ALMO-CIS method⁴⁰ can be inherited without major modification. Nevertheless, some aspects of the two-electron integral evaluation have been reformulated to further increase computational efficiency. The following subsections will discuss these aspects

A. Strategy for Two-Electron Integral Evaluation

The two electron integrals appearing in Eq. 3, including a Coulomb-like term $(\psi_i \phi_a | \psi_j \phi_b)$ and an exchange-like term $(\psi_i \psi_j | \phi_a \phi_b)$, are expanded in terms of unprojected ALMO contributions and projection corrections:

$$\begin{aligned} (\psi_i \phi_a | \psi_j \phi_b) &= \mathcal{N}_a \mathcal{N}_b \{ (\psi_i \psi_a | \psi_j \psi_b) \\ &\quad - 2 (\psi_i \psi_a | \psi_j \psi_k) (\psi^k | \psi_b) \\ &\quad + (\psi_i \psi_k | \psi_j \psi_l) (\psi^k | \psi_a) (\psi^l | \psi_b) \} \\ J &= \mathcal{N}_a \mathcal{N}_b (J1 - 2 \cdot J23 + J4) \end{aligned} \quad (13)$$

$$\begin{aligned} J23 &= (\psi_i \psi_a | \psi_j \psi_k) (\psi^k | \psi_b) = (\psi_i \psi_a | \chi_P) (\chi^P | \chi^Q) (\chi_Q | \phi_j \psi_k) (\psi^k | \psi_b) \\ J4 &= (\psi_i \psi_k | \psi_j \psi_l) (\psi^k | \psi_a) (\psi^l | \psi_b) = (\psi^k | \psi_a) (\psi_i \psi_k | \chi_P) (\chi^P | \chi^Q) (\chi_Q | \phi_j \psi_l) (\psi^l | \psi_b) \\ K23 &= (\psi_i \psi_j | \psi_a \psi_k) (\psi^k | \psi_b) = (\psi_i \psi_j | \chi_P) (\chi^P | \chi^Q) (\chi_Q | \psi_a \psi_k) (\psi^k | \psi_b) \\ K4 &= (\psi_i \psi_j | \psi_k \psi_l) (\psi^k | \psi_a) (\psi^l | \psi_b) = (\psi_i \psi_j | \chi_P) (\chi^P | \chi^Q) (\chi_Q | \psi_k \psi_l) (\psi^k | \psi_a) (\psi^l | \psi_b) \end{aligned}$$

The number of three-center two-electron integrals needed does not depend on whether charge transfer states are included, since for both cases, the indices of three-center two-electron integrals should run over all the occupied and virtual orbitals, as well as all the auxiliary basis. Thus, the algorithm used in the ALMO-CIS model, namely, the “digestor” that transforms the three-center integrals from AO basis to MO basis can be adopted without any change. On the other hand, the following contraction steps need to be modified to account for the fact that more occupied-virtual pairs are now included.

The leading terms, $J1$ and $K1$, can be evaluated using the schemes that were presented previously⁴⁰. However, this may not be optimal and we have developed new algorithms that will improve the computational efficiency. In the next two subsections, we shall compare different possible schemes in build-

$$\begin{aligned} (\overline{\psi_i \psi_j | \phi_a \phi_b}) &= \mathcal{N}_a \mathcal{N}_b \{ (\overline{\psi_i \psi_j | \psi_a \psi_b}) \\ &\quad - 2 (\overline{\psi_i \psi_j | \psi_a \psi_k}) (\psi^k | \psi_b) \\ &\quad + (\overline{\psi_i \psi_j | \psi_k \psi_l}) (\psi^k | \psi_a) (\psi^l | \psi_b) \} \\ K &= \mathcal{N}_a \mathcal{N}_b (K1 - 2 \cdot K23 + K4) \end{aligned} \quad (14)$$

The expansion is exactly the same as in the ALMO-CIS model, except here two indices connected by a contraction line belong to two neighboring fragments, rather than being restricted to the same fragment. **The terms involving four-center two-electron integrals with three or four occupied indices ($J23$, $J4$, $K23$ and $K4$) are corrections to the four-center unprojected integrals and are essential for quantitative accuracy.**

We evaluate the correction terms using the resolution of the identity (RI) approximation^{45–47}, where the four-center two-electron integrals are decomposed into three-center integrals, $(\psi_i \psi_j | \chi_P)$ and $(\psi_i \psi_a | \chi_P)$, as well as the inverse of the Coulomb metric in the auxiliary basis $(\chi^P | \chi^Q)$:

ing the $J1$ and $K1$ terms, and discuss the necessity of adopting a better scheme when charge transfer states are present.

B. ALMO Coulomb Integral Evaluation

Let us begin by analyzing the formal scaling of the previous ALMO-CIS implementation. The most computationally significant step was to form the half-transformed integrals $J_{ia}^{\mu\nu} = (\chi_\mu \chi_\nu | \psi_i \psi_a)$, which were computed as a contraction of the AO integrals $(\mu\nu | \lambda\sigma)$ with the pseudo densities $P_{ia}^{\lambda\sigma} = c_i^\lambda c_a^{\dagger\sigma}$. When the excitations are restricted to be intrafragment, the only AO integrals we need to compute are those with μ and ν on the same fragment,

and λ and σ on a different fragment. Their number is only $NNnn$. Because of the block structure of the density, each integral will contract with ov densities, so the scaling for the contraction step is $NNnmov$.

With CT excitations between near-neighbors included, the AO integrals that are needed will grow to become $NN\tilde{n}\tilde{n}$, and the scaling will become $NN\tilde{n}\tilde{n}ov$. Therefore, although the scaling of the contraction remains quadratic with system size, the prefactor is $(\tilde{n}/n)^2$ times larger, which is roughly a factor of 25 for the case of helium clusters. Based on the ALMO-CIS timing results⁴⁰, this suggests that building the Coulomb integrals is likely to become a dominant step if no improvements are made. Thus, we have to look into other ways of doing the Coulomb-type contraction.

One possibility is to replace the contraction of AO integrals with densities by two quarter transforms to the ALMO representation. The AO integrals $(\mu\nu | \lambda\sigma)$ are first transformed with ALMO coefficients C_i^λ to form $(\mu\nu | i\sigma)$, and then the quarter transformed integrals are contracted with C_a^σ to form the half-transformed integrals $(\mu\nu | ia)$. In this scheme, with CT substitutions, the scaling for the first and second quarter transforms are given by $NN\tilde{n}\tilde{n}o$ and $NN\tilde{n}\tilde{n}ov$, respectively. Thus, by replacing the contraction with densities by two quarter transformations, the computational cost can be reduced by a factor of v , the average number of virtual orbitals on a single fragment. This is a significant improvement.

However, there is a price to be paid for the two quarter transformation approach. When contracting with the densities, an AO integral can be used multiple times because of the permutation symmetries. For example, $(\mu\nu | \lambda\sigma)$ is not only contracted with $P_{ia}^{\lambda\sigma}$, but also used as $(\lambda\sigma | \mu\nu)$ to contract with $P_{jb}^{\mu\nu}$. As a result, only about one-eighth of the integrals are computed since there are eight such permutations. On the other hand, if the AO integrals are quarter-transformed with MO coefficients, the permutation between the bra side and ket side cannot be readily used. To understand this, one can look at the batching scheme shown in Algorithm 1. The second quarter transform is performed after the first quarter transformation of a particular batch of bra values is done, and only a batch of quarter transformed integrals $(\mu\nu | i\sigma)$ is held in memory at one time. The first quarter transform must be completed by running over all pairs of ket indexes, so it is not possible to only loop over the ket batches with in-

Algorithm 1: (Half) integral transform for the Coulomb term, $J_{\mu\nu ia} = (\mu\nu | ia)$

```

for bra batches do
  for ket batches do
    for  $\mu\nu \in batch$  do
      for  $\lambda\sigma \in batch$  do
         $\lambda \rightarrow F_I, \sigma \rightarrow F_J$ 
        for  $i = 1, O_{F_I}$  do
           $(\mu\nu | i\sigma) += (\mu\nu | \lambda\sigma) C_i^{I\lambda}$ 
        for  $j = 1, O_{F_J}$  do
           $(\mu\nu | j\lambda) += (\mu\nu | \lambda\sigma) C_j^{J\sigma}$ 
    for  $\mu\nu \in batch$  do
      for  $ia \in \text{FrgPairList}$  do
         $\sigma \rightarrow F_A$ 
        for  $a = 1, V_{F_A}$  do
           $(\mu\nu | ia) += (\mu\nu | i\sigma) C_a^{A\sigma}$ 

```

Use the permutation between $\mu\nu$ to scatter $\mu\nu$ from shellpair form to matrix form.

indices smaller or larger than the bra batch index. If one want to use “upper triangle” or “lower triangle” loop structures for batches, the second half transform has to be moved outside the loop over bra batches. This, however, will demand all the quarter transformed integrals to be stored in memory, which is impractical for large systems, since the number of quarter transformed integrals is $NN\tilde{n}\tilde{n}o$. Therefore, in our current implementation, we employ the quarter transform at the price of computing twice as many AO integrals as we have before. The sacrifice is worthwhile if the cost of computing AO integrals is insignificant compared to the contraction steps, which is true for the Coulomb like integrals, as shown in the timing result in the Result section.

TABLE I: Comparison of the density contraction and the MO Coefficient contraction schemes for the half-transform to build J1

scheme	step	scaling	computing AO-ERI twice
a	$(\mu\nu \lambda\sigma) P_{ia}^{\lambda\sigma} \rightarrow (\mu\nu ia)$	$NN\tilde{n}\tilde{n}ov$	No
b	$(\mu\nu \lambda\sigma) C_i^\lambda \rightarrow (\mu\nu i\sigma)$ $(\mu\nu i\sigma) C_a^\sigma \rightarrow (\mu\nu ia)$	$NN\tilde{n}\tilde{n}o$ $NN\tilde{n}\tilde{n}ov$	Yes

The exchange term in ALMO-CIS was evaluated from half-transformed integrals $K_{\mu\lambda}^{ia} = (\chi_\mu\psi_i|\chi_\lambda\psi_a)$, which were formed by contracting the AO integrals with pseudo-density matrices $P_{ia}^{\nu\sigma} = c_i^\nu c_a^\sigma$. If this density contraction scheme is used for the case with charge transfer excitations, the scaling is predicted to be $(NN)\tilde{n}\tilde{n}o$. Here the parentheses on NN comes from the sparsity of $\mu\nu$ pairs, namely, the number of significant shell pairs grows linearly with system size, and thus we expect the overall scaling to form the half transformed integrals is $O(M)$ for large enough systems (a limit that is not easily reached with the very diffuse basis sets needed for excited states).

It is promising to consider using the two quarter transformations for the exchange term as well. There will be further improvement in both scaling and memory concerns if we transform the two occupied indexes first. The first step transforms $(\overline{\mu\nu}|\lambda\sigma)$ with C_i^μ to form $(i\nu|\lambda\sigma)$, with compute effort that scales as $(NN)\tilde{n}\tilde{n}o$. Next, the quarter transformed integrals are transformed with C_j^ν to form $(ij|\lambda\sigma)$, at a cost that scales as $(NN)\tilde{n}\tilde{d}o$. Therefore, the MO contraction scheme will reduce the cost of digesting the integrals by a factor of v as it did in the Coulomb case, and the memory storage for the half transformed integrals is also reduced by a factor of v/o .

However, we have not implemented that scheme yet. For the helium cluster systems we focus on, detailed timings show that, unlike the Coulomb-like terms, the cost of computing the AO integrals is comparable to the cost of the transformation steps. Thus the advantage in contraction is likely to be offset by computing twice as many integrals. For the special case of helium clusters, we implemented an alternative scheme for digestion, where the two occupied indices are transformed at the same time. The idea of this method is similar as using density matrices, one can view it as contracting with a density matrix built by occupied orbitals $P_{ij}^{\mu\nu} = c_i^\mu c_j^{\nu\dagger}$. Details of this scheme are shown in Algorithm 2.

Table II compares different aspects of the three schemes mentioned above. It is noteworthy that our current approach scales as $(NN)\tilde{n}\tilde{n}oo$. For the helium clusters, there is only one occupied orbital per fragment, i.e. $o = 1$. Therefore we actually achieve the same scaling as MO coefficient transformation

scheme(which scales as $(NN)\tilde{n}\tilde{n}o$), without the disadvantage of doubling the cost of computing AO integrals. When the cost for digestion step is dominant, which is likely to happen for systems with larger fragments, successive quarter transformation scheme will be preferable.

Algorithm 2: (Half) integral transform for the Exchange term, $K_{\mu\nu ij} = (\overline{\mu\nu}|ij)$

```

for bra batches do
  for ket batches do
    for  $\mu\nu \in \text{batch}$  do
      for  $\lambda\sigma \in \text{batch}$  do
         $\mu \rightarrow F_L, \nu \rightarrow F_J, \lambda \rightarrow F_K, \sigma \rightarrow F_L,$ 
        if  $F_I, F_K \in \text{FrgPairList}$  and
            $F_J, F_L \in \text{FrgPairList}$  then
          for  $k = 1, O_{F_K}$  do
            for  $l = 1, O_{F_L}$  do
               $(\overline{\mu\nu}|kl) += (\overline{\mu\nu}|\lambda\sigma) C_{kk}^{K\lambda} C_{ll}^{L\sigma}$ 
            for  $i = 1, O_{F_I}$  do
              for  $j = 1, O_{F_J}$  do
                 $(\lambda\sigma|ij) += (\overline{\mu\nu}|\lambda\sigma) C_{ii}^{I\mu} C_{jj}^{J\nu}$ 
          if  $F_I, F_L \in \text{FrgPairList}$  and
              $F_J, F_K \in \text{FrgPairList}$  then
            for  $k = 1, O_{F_K}$  do
              for  $l = 1, O_{F_L}$  do
                 $(\overline{\mu\nu}|lk) += (\overline{\mu\nu}|\lambda\sigma) C_{kk}^{K\lambda} C_{ll}^{L\sigma}$ 
              for  $i = 1, O_{F_I}$  do
                for  $j = 1, O_{F_J}$  do
                   $(\lambda\sigma|ji) += (\overline{\mu\nu}|\lambda\sigma) C_{ii}^{I\mu} C_{jj}^{J\nu}$ 
           $K_{\mu\nu ij} = K_{\mu\nu ij} + K_{\nu\mu ji}$ 

```

TABLE II: Comparison of different schemes for the half-transform to build K1

scheme	step	scaling	computing AO-ERI twice
a	$(\overline{\mu\nu} \lambda\sigma) P_{ia}^{\nu\sigma} \rightarrow (\overline{\mu i} \lambda a)$	$(NN)\tilde{n}\tilde{n}ov$	No
b	$(\overline{\mu\nu} \lambda\sigma) C_i^\mu \rightarrow (i\nu \lambda\sigma)$	$(NN)\tilde{n}\tilde{n}o$	Yes
	$(i\nu \lambda\sigma) C_j^\nu \rightarrow (ij \lambda\sigma)$	$(NN)\tilde{n}\tilde{d}o$	
c	$(\overline{\mu\nu} \lambda\sigma) C_i^\mu C_j^\nu \rightarrow (ij \lambda\sigma)$	$(NN)\tilde{n}\tilde{n}oo$	No

We have also implemented some optimizations with respect to the screening of integrals. The previous scheme for making the "mini-list" (significant quartet of AO basis shells) is summarized in Algorithm 3. The screening is based on the Schwarz inequality $|(\mu\nu|\lambda\sigma)| \leq (\mu\nu|\mu\nu)^{\frac{1}{2}} (\lambda\sigma|\lambda\sigma)^{\frac{1}{2}}$. This

selection step appears to be trivial, since all quantities needed in the algorithm can be pre-made, computations inside the loops are only a few multiplications and conditional evaluations. However, this is a quartic scaling step because of the loops over bra and ket shell pairs, and, for large systems, it actually dominates the evaluation of exchange integrals in the previous ALMO-CIS implementation. This is the reason that the overall scaling of computing exchange integrals was found to be 2.88, while it was supposed to be linear with system size.

Algorithm 3: previous scheme for mini-list selection in building K1

```

for bra shell pairs (IShl, JShl) do
  IShl → FI; JShl → FJ;
  for ket shell pairs (KShl, LShl) do
    KShl → FK; LShl → FL;
    if FI, FK ∈ FrgPairList and
      FJ, FL ∈ FrgPairList
    or FI, FL ∈ FrgPairList and
      FJ, FK ∈ FrgPairList then
      I1: max(μν | μν)1/2, μν ∈ bra shell pairs
      I2: max(λσ | λσ)1/2, λσ ∈ ket shell pairs
      if FI, FK ∈ FrgPairList and
        FJ, FL ∈ FrgPairList then
        Pmax: max(|Piaμλ|, |Piaνσ|)
      if FI, FL ∈ FrgPairList and
        FJ, FK ∈ FrgPairList then
        Pmax: max(|Piaμσ|, |Piaνλ|)
      if I1I2Pmax > thresh then
        add IShl, JShl, KShl, LShl to mini-list.

```

The current scheme for mini-list selection is described in Algorithm 4. We preselect fragment pairs based on the maximum value of occupied four center integrals ($ij | ij$) on a given pair. This allows us to have a selection before entering the loop of ket shell pairs. In this way, the scaling of the algorithm can be reduced to cubic, if the number of fragment pairs that has ($ij | ij$) bigger than thresh is linear with system size.

IV. ASSESSMENT OF THE ALMO-CIS+CT IMPLEMENTATION

A. Accuracy

We first study the accuracy of the ALMO-CIS+CT model with different choices of r_{cut} , in the hope that a relatively small r_{cut} can yield satisfactory accuracy, so that this model can be practically useful. Our test system is a small He₂₅ cluster

Algorithm 4: current scheme for mini-list selection in building K1

```

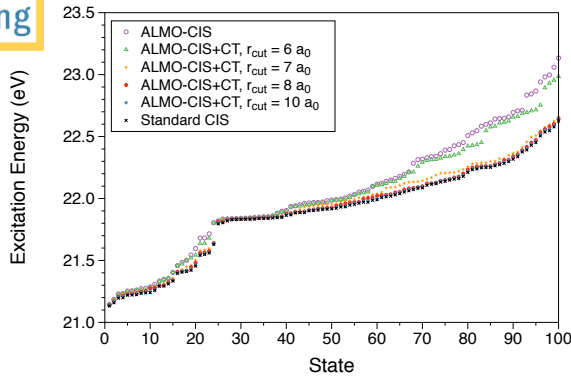
for bra shell pairs (IShl, JShl) do
  IShl → FI; JShl → FJ
  if max(ij | ij) < thresh(i ∈ FI, j ∈ FJ) then
    continue
  for ket shell pairs (KShl, LShl) do
    KShl → FK; LShl → FL;
    if FI, FK ∈ FrgPairList and FJ, FL ∈ FrgPairList
    or FI, FL ∈ FrgPairList and FJ, FK ∈ FrgPairList
    then
      add IShl, JShl, KShl, LShl to mini-list.

```

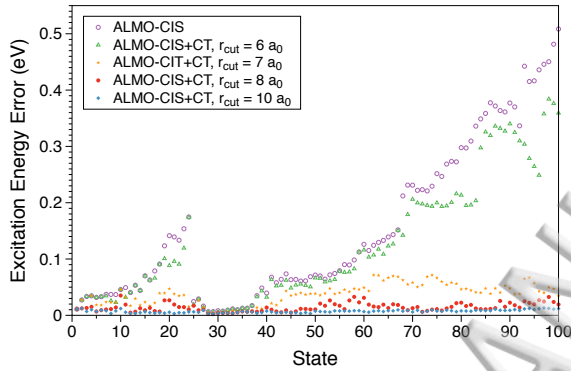
Note: The algorithm presented here is a simplified version of our actual implementation. In practice, considering the permutation between bra and ket side, we also need to select F_K, F_L based on $\max(kl | kl)$, and find F_I, F_J that connect with F_K, F_L using FrgPairList. This leads to a more complicated algorithm although it does not affect the scaling.

that has been studied with both standard CIS and ALMO-CIS⁴⁰. In this and all the following calculations for helium clusters, each helium atom will be treated as a fragment. Figure 1 shows the excitation energies of the first 100 states (which is the $n = 2$ manifold) when r_{cut} is chosen to be 6, 7, 8, 10 a_0 , along with the results obtained by standard CIS and ALMO-CIS. In this test, the RI approximation and the Davidson-like method are not applied, so that any error purely comes from the truncation of matrices in Eq. 5 (the error due to the use of ALMOs instead of CMOs is negligible in the case of helium clusters). The basis we use here is a modified 6-311(2+)G basis, which has 11 functions per helium atom. Thus in the absence of truncation, there will be 6250 single substitutions in total. By contrast, ALMO-CIS retains only the 250 intrafragment excitations. The choices of $r_{cut} = 6, 7, 8, 10 a_0$ lead to the number of retained excitations being 330, 730, 1030 and 1490, respectively. From Fig. 1, we find for $r_{cut} = 8, 10 a_0$, the excitation energies are almost identical to those obtained in standard CIS. We know that ALMO-CIS is least accurate at the high-energy edge of each band because these are the states with stronger CT character (some insights are provided in Sec. VD). For example, compared to standard CIS, ALMO-CIS exhibits errors of ~ 0.5 eV at the blue end of the $n = 2$ band. This error has been reduced to less than 0.02 eV in the ALMO-CIS+CT model, when $r_{cut} = 8 a_0$.

Figure 2 compares the spectrum of He₂₅ calculated by standard CIS, ALMO-CIS and ALMO-CIS+CT with $r_{cut} = 8 a_0$. We use the same He₂₅ geometries that were used for the spectrum computed by CIS and ALMO-CIS, which come from 100



(a)



(b)

FIG. 1: (a) $n = 2$ manifold excitation energies for He_{25} , evaluated by CIS, ALMO-CIS and ALMO-CIS+CT with $r_{\text{cut}} = 6, 7, 8, 10 a_0$. States 1-25 are of $2s$ character, and states 26-100 are of $2p$ character. (b) The absolute errors of excitation energies for ALMO-CIS and ALMO-CIS+CT model compared with standard CIS.

randomized clusters that are optimized subsequently at the MP2/6-311G level of theory¹³. It has been found that ALMO-CIS tends to give an overall more broadened spectrum, with the intensity at the low-energy edge (corresponding to surface states) being underestimated, and that at the high-energy edge (corresponding to bulk states) being overestimated. Encouragingly, when a cutoff distance of $8 a_0$ is employed, these errors are almost eliminated from the ALMO-CIS+CT spectrum.

Next we want to demonstrate the accuracy of the Davidson-like variational method for the excitation energies of the same He_{25} cluster. In Figure 3, we compare the ALMO-CIS and ALMO-CIS+CT re-

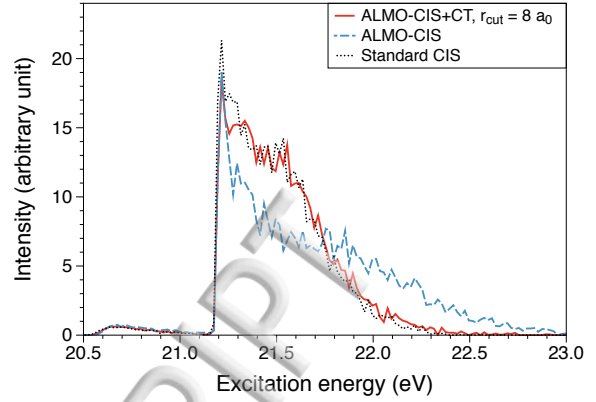


FIG. 2: Absorption spectrum of He_{25} cluster at the ALMO-CIS, ALMO-CIS+CT ($r_{\text{cut}} = 8 a_0$) and standard CIS level of theory. All spectrum profiles are shifted to the low-energy end by 0.625 eV to match the atomic peak of the experimental spectrum.

sults with and without the Davidson-like method ($r_{\text{cut}} = 8 a_0$). Two different preconditioners were examined, the full preconditioner \mathbf{P}_{full} and the diagonal preconditioner \mathbf{P}_{diag} .

$$\begin{aligned} \mathbf{P}_{\text{full}} &= (\tilde{\mathbf{A}}_{cc} - \bar{\omega} \tilde{\mathbf{G}}_{cc})^{-1} \\ \mathbf{P}_{\text{diag}} &= (\text{diag}(\tilde{\mathbf{A}}_{cc} - \bar{\omega} \tilde{\mathbf{G}}_{cc}))^{-1} \end{aligned} \quad (15)$$

We find with the full preconditioner, the Davidson-like variational method causes nearly no additional error, since the data points of \mathbf{P}_{full} almost overlap with those by solving Eq. 5 directly. When the diagonal preconditioner, \mathbf{P}_{diag} , is utilized, only about 75% of the ALMO-CIS model error can be eliminated. These facts suggest that the matrices $\tilde{\mathbf{A}}_{cc}$ and $\tilde{\mathbf{G}}_{cc}$ are dense, most likely because there is already a truncation in Eq. 5.

At this moment, we conclude that $r_{\text{cut}} = 8 a_0$ and a full preconditioner offer a good balance between accuracy and efficiency, and they will be used for all the calculations in the rest of the paper. Assuming that the structures of helium clusters more or less resemble that of the bulk system: an atom has 6 nearest neighbors⁴⁸ and the interatomic distance is 3.6 \AA ⁴⁹, the r_{cut} we have here will include the first shell of nearby atoms. This means that for the finite size clusters studied in this work, the matrix size of ALMO-CIS+CT will be no more than seven times (which is the bulk limit) of that in ALMO-CIS.

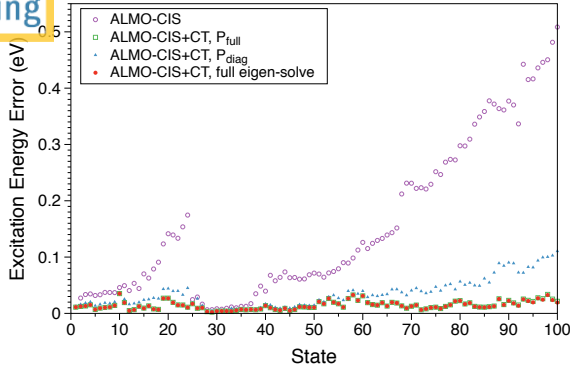


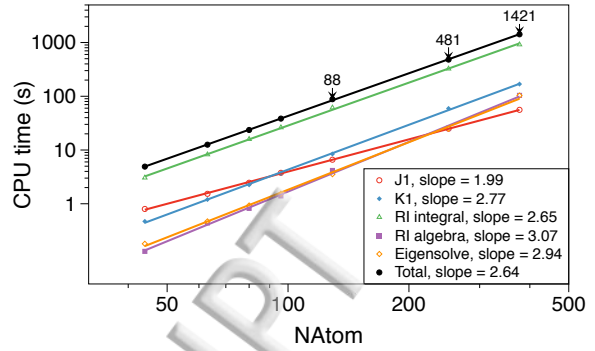
FIG. 3: Errors of the excitation energies for He_{25} using ALMO-CIS, ALMO-CIS+CT solved exactly, and ALMO-CIS+CT solved by the Davidson-like method, with P_{full} and P_{diag} as the preconditioners. $r_{cut} = 8 a_0$ for all ALMO-CIS+CT calculations

B. Timings

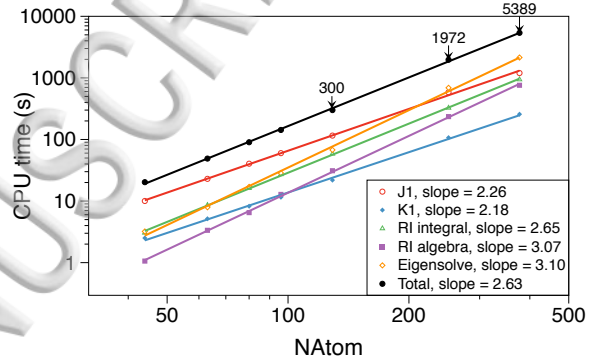
Two issues related to the efficiency of our models are of particular interest: the scaling versus system size, and the increase of computational effort when charge transfer is included. To address these two problems, we report timing results using helium clusters ranging from 44 to 377 regularly spaced atoms (4 \AA) with both ALMO-CIS and ALMO-CIS+CT ($r_{cut} = 8 a_0$) models. All calculations are performed with a single core on an AMD Opteron 6376 processor, using the same customized 6-311(2+)G basis as the previous subsection.

The relative costs of the major computational steps of ALMO-CIS and ALMO-CIS+CT are summarized in Figure 4. The slopes of the fitted lines show the scaling of each step. Overall, both methods have sub-cubic scaling, as we expected. It is noteworthy that forming the RI integrals, which is the most expensive step in ALMO-CIS, is exactly the same in ALMO-CIS+CT. The unchanged timing of this step is the main reason for the fact that the total cost of ALMO-CIS+CT is about only four times as large as that of ALMO-CIS, even though the matrix dimension is 5 – 6 times larger.

Detailed timing of forming the Coulomb-like and exchange-like integrals are tested with He_{129} and He_{251} clusters, and the results are presented in Table III and Table IV. For the Coulomb-like integrals, the relatively expensive cost of contraction steps (especially the first quarter transform) supports our ar-



(a) ALMO-CIS



(b) ALMO-CIS+CT

FIG. 4: CPU timing data for ALMO-CIS (a) and ALMO-CIS+CT with $r_{cut} = 8 a_0$ (b). The y axis is in logarithm scale of the CPU time.

gument that it is worthwhile to apply the successive MO transform scheme at the cost of computing AO integrals twice. We note that for building AO integrals in J1, an incompletely optimized new integral library is used. We expect roughly a four times speed-up of the AO integral computation in the future based on the performance of the old library. Also, we do not list the timing for mini-list selection as we do for the exchange-like integrals because the selection can be done at the shell-pair level, and it is a trivial step as it scales $O(M^2)$.

On the other hand, even with the optimizations described above, mini-list selection still takes about half of the time in forming the exchange integrals for ALMO-CIS+CT, and dominates in ALMO-CIS calculations where forming integrals and contractions are very cheap. The observed scaling of forming K1 is 2.77 for ALMO-CIS and 2.18 for ALMO-CIS+CT, instead of linear as is asymptotically possible. When

charge transfer is included, we expect a $(\tilde{n}/n)^2$ times growth of computational effort in contraction steps and forming AO integrals. Based on the ratio of number of states in ALMO-CIS+CT and ALMO-CIS, $(\tilde{n}/n)^2$ is estimated to be 25 and 35 for He₁₂₉ and He₂₅₁, respectively, and the observed growth in timing does not exceed these ratios.

To demonstrate the computational savings from the Davidson-like variational method, we also report the timing of the full eigen-solve for He₁₂₉ and He₂₅₁ in Table V. It can be seen that although the Davidson-like method does not change the cubic scaling, it is about 7 times faster than solving the full eigenvalue problem for these two systems. In addition, the preconditioning step (with full preconditioner) takes only about 1/6 of the total cost of the Davidson-like method. Thus, a simplified preconditioner such as \mathbf{P}_{diag} would not significantly reduce the computational cost. Considering the additional error it entails, the diagonal preconditioner is not recommended at this time.

V. CALCULATING THE ABSORPTION SPECTRA OF HELIUM CLUSTERS

A. PIMD simulation

In the previous ALMO-CIS work⁴⁰, we generated an ensemble of cluster geometries starting from randomized initial geometries which were then optimized using either MP2 theory, or a classical force field for bigger clusters where MP2 is inaccessible. The resulting geometries are a myriad of local minima because of the shallow potential wells and many degrees of freedom. However, with a standard Lennard-Jones potential ($\epsilon_0 = 10.7$ K, $r_0 = 2.9$ Å), we find the resulting clusters are too dense. To qualitatively match the known facts about helium clusters (for example, the average interatomic distance in bulk helium is 3.6 Å), we were forced to use a modified Lennard-Jones potential with a much shallower potential well ($\epsilon_0 = 0.05$ K) and a larger equilibrium distance ($r_0 = 3.6$ Å). Of course the origin of the failure of the standard classical force field is due to neglecting the nuclear zero point motion, which is critical for helium atoms.

Feynman's path integral theory⁵⁰ treats nuclei quantum mechanically by mapping each quantum nucleus onto a classical system comprising several fictitious particles connected by springs ("ring-polymers"). This provides an ideal computational technique for the simulation of helium clusters, and

there have been various successful applications of path integral Monte Carlo (PIMC)⁵¹ and path integral molecular dynamics (PIMD)^{52,53} to helium systems. In this work, we employ PIMD to generate configurations of helium clusters of different sizes, which are then used to calculate the absorption spectrum.

The temperature of helium clusters in spectral measurements may be as low as 0.4 K, where the Bose statistics of ⁴He atoms are significant. To properly account for the indistinguishability of identical particles, one has to sample the permutations as well. Ceperley *et al.* has employed a PIMC method that includes exchange to study helium clusters and found that superfluidity exists even in clusters as small as He₆₄⁵⁴. For simplicity of implementation, our current PIMD formalism neglects the exchange between particles. Therefore our simulations are performed at 3 K, which is above the superfluid transition temperature.

In a PIMD formulation, the partition function is given by:

$$Z = \lim_{P \rightarrow \infty} \left(\frac{1}{2\pi\hbar} \right)^{NP} \int d^{NP} \mathbf{q} \int d^{NP} \mathbf{p} e^{-\beta_P H_P(\mathbf{p}, \mathbf{q})} \quad (16)$$

where N is the number of distinguishable particles, P is the number of discretization points of quantum paths (or the number of chain particles in a ring polymer), $\beta_P = 1/Pk_B T$, and $H_P(\mathbf{p}, \mathbf{q})$ is the ring polymer Hamiltonian:

$$H_P(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^N \sum_{p=1}^P \left(\frac{\mathbf{p}_{i,p}^2}{2m_i} + \frac{1}{2} m_i \omega_P^2 [\mathbf{q}_{i,p} - \mathbf{q}_{i,p+1}]^2 \right) + \sum_{p=1}^P V(\mathbf{q}_{1,p}, \mathbf{q}_{2,p}, \dots, \mathbf{q}_{N,p}) \quad (17)$$

with $\omega_P = 1/\beta_P \hbar$. In our simulation, the interparticle potential V is described by the HFDHE2 potential⁵⁵. The time evolution of PIMD follows the normal mode algorithm, and a white-noise Langevin thermostat⁵⁶ is employed (see Ref. 57 for further details). We find $P = 64$ and a time step of 2 fs yield converged results.

At finite temperature and zero pressure, the helium cluster will always evaporate in the long time limit. To avoid this issue, we confine the system within a sphere of radius R_c , so that an equilibrium between liquid and vapor can be established. The parameter R_c will affect the density of the system, and the available zero temperature density profile of helium clusters provide guidance for choosing R_c . A reasonable value of R_c should not be too small, so

TABLE III: CPU time (s) for significant steps in the construction of J1 integrals.

Method	System	NState	AO integrals	$(\mu\nu \lambda\sigma) C_i^\lambda$ $\rightarrow (\mu\nu i\sigma)$	$(\mu\nu i\sigma) C_a^\sigma$ $\rightarrow (\mu\nu ia)$	Total
ALMO-CIS	He ₁₂₉	1290	3.75	2.31	0.17	6.60
	He ₂₅₁	2510	14.24	8.66	0.64	24.89
ALMO-CIS+CT	He ₁₂₉	6570	59.62	43.46	4.01	114.56
	He ₂₅₁	14870	298.20	219.72	20.99	577.48

TABLE IV: CPU time (s) for significant steps in the construction of K1 integrals.

Method	System	NState	Mini-list	AO integrals	$(\mu\nu \lambda\sigma) C_i^\mu C_j^\nu$ $\rightarrow (ij \lambda\sigma)$	Total
ALMO-CIS	He ₁₂₉	1290	7.81	0.26	0.34	9.48
	He ₂₅₁	2510	56.06	0.64	0.92	61.48
ALMO-CIS+CT	He ₁₂₉	6570	9.51	3.26	6.81	23.69
	He ₂₅₁	14870	63.9	9.90	17.93	110.79

TABLE V: CPU time (s) for the Davidson-like method and the full eigen-solve on He₁₂₉ and He₂₅₁ clusters.

System	Davidson-like method	Preconditioning	Full eigen-solve
He ₁₂₉	71.68	11.81	490.18
He ₂₅₁	696.87	131.48	6409.57

that the boundary will not have too much influence. It cannot be too big either, otherwise its role in preventing the system from evaporating is undermined. Unfortunately, we have no way to determine an optimal value of R_c . Therefore, for small clusters such as He₇₀ and He₁₅₀, we perform simulations using several different R_c , and examine how the resulting spectra change. Then, to study the size-dependence of the spectra, we pick a reasonable density, and run the simulation with R_c determined based on the fixed density for clusters of different sizes.

B. Dependence on R_c

Three different R_c are used for the simulation of He₇₀ and He₁₅₀, respectively ($R_c = 16, 18, 20 \text{ \AA}$ for He₇₀, $R_c = 20, 23, 26 \text{ \AA}$ for He₁₅₀). The simulations are run with 200,000 warm-up steps (400 ps) to ensure that equilibrium is reached, and properties of interest and geometries used for spectral calculations are extracted from the following 400,000 time steps (800 ps). The resulting geometries can be characterized by density profiles $\rho(r)$ and nearest neighbor distribution $nn(r)$ (as defined in Ref. 58), which are collected in Figure 5 along with the spectra. It is ev-

ident that the gas phase becomes increasingly dominant for larger R_c , which is reflected in a stronger atomic peak and smaller intensity at the high-energy edge of the spectra. This trend exists for both cluster sizes.

The kinetic energy (estimated by a virial estimator⁵⁹), potential energy and total energy per helium atom are listed in Table VI. The positive total energies may imply that the system will eventually dissociate into the gas phase. To address this issue, we perform a simulation of He₁₅₀ with $R_c = 24 \text{ \AA}$, using only the repulsive part of the HFDHE2 potential. The system quickly develops into an evenly distributed gas, and the corresponding spectrum is more like a single atomic peak. Snapshots of He₁₅₀ simulated with and without the attractive part of the potential (Figure 6) exhibit a clear difference. We conclude that liquid droplet structure exists at least for the time scale of our simulation. Meanwhile, Barnett *et al.* has performed similar PIMD simulation of He₇₀ with $R_c = 18.3 \text{ \AA}$ and He₁₅₀ with $R_c = 23.6 \text{ \AA}$ ⁵². The energies and density profiles we obtained qualitatively agree with those available in their publication.

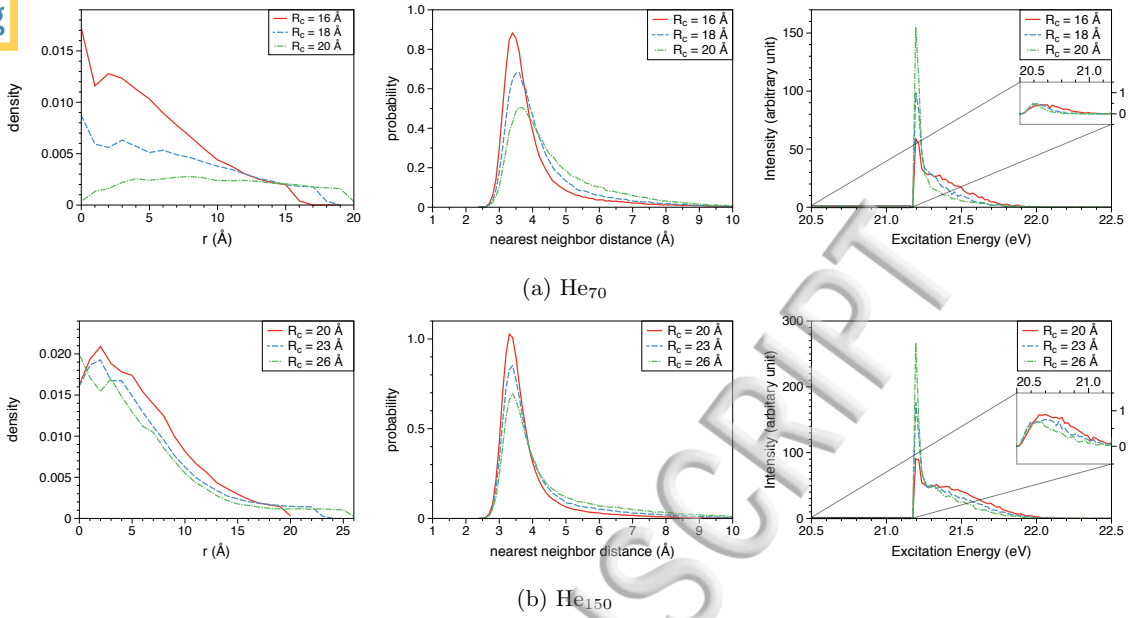


FIG. 5: Density profiles $\rho(r)$, nearest neighbor distributions $nn(r)$, and spectra of He_{70} (a) and He_{150} (b). The density profiles are plotted vs distance from the center of the confining sphere. Spectra are calculated by the ALMO-CIS+CT model, using 50 geometries from each PIMD simulation.

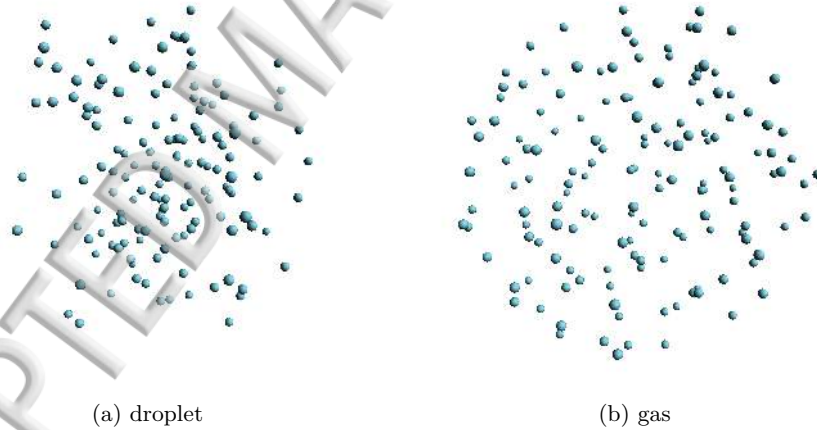


FIG. 6: Snapshots of He_{150} cluster at $t = 400$ ps when simulated with (a) and without (b) the attractive part of HFDHE2 potential. The resulting configurations are droplet and gas-like, respectively.

C. Size-dependence of the Spectra

To study the size-dependence of the spectra, we investigate four helium clusters of different sizes: He_{70} , He_{150} , He_{231} and He_{300} , and R_c is set to be 18, 23, 27, 29 Å, respectively, so that all four sys-

tems correspond to roughly the same density of $2.9 \times 10^{-3} \text{ \AA}^{-3}$. In Figure 7, we can see that the spectra of larger clusters have more intensity at the high-energy edge. This is reasonable since this “shoulder” next to the atomic peak results from the interaction between helium atoms, which should be stronger

TABLE VI: The density, per-particle kinetic, potential and total energy of He₇₀ and He₁₅₀ clusters simulated by PIMD at 3 K with different choices of R_c .

System	R_c (Å)	ρ (10^{-3} Å ⁻³)	T (K)	V (K)	E (K)
He ₇₀	16	4.08	7.44	-5.94	1.50
	18	2.87	6.26	-3.67	2.59
	20	1.57	5.57	-2.32	3.25
He ₁₅₀	20	4.48	8.50	-7.99	0.51
	23	2.94	7.61	-6.21	1.40
	26	2.04	6.85	-4.77	2.07

in larger clusters, as they have smaller interatomic spacing indicated by the density profile and nearest neighbor distribution. Moreover, the larger fraction of bulk atoms in the larger clusters also contributes to the broadening of the spectra.

D. Characterization of Excited States

A previous CIS study¹³ of small helium clusters (He₇ and He₂₅) concluded that the higher-energy excited states mainly come from bulk-type excitations, while the lower-energy states come from surface excitations. To further understand this, we compare the droplet (Figure 6(a)) and gas (Figure 6(b)) spectra of He₁₅₀ at $R_c = 16$ and 23 Å (Figure 8(a)). The droplet typically has a gradually decreasing density from the core region to the surface, while the gas structure has a roughly uniform density in the whole sphere. For both R_c 's, the droplet spectra are broader and more extended to the high-energy edge. This is because the droplet has a varying density, and its core density is higher than that of gas with the same R_c . It is also interesting to see, that when R_c decreases, the atomic peak diminishes in the droplet spectra, while for gas, we see a shift of the maximum peak instead.

The relation between the excitation energy and surface/bulk character of each state motivates us to find a mapping between these two. Previous state-by-state inspection¹³ of the attachment and detachment density is unfeasible for large clusters and multiple configurations. Alternatively, here we introduce a scalar quantity \bar{R} to represent the ‘‘average position’’ of the excitation relative to the center of

the cluster. For state κ , \bar{R}_κ is defined as:

$$\bar{R}_\kappa = \sum_{\substack{F_I, F_A \in \\ \text{FrgPairList}}} \sum_{\substack{i \in F_I \\ a \in F_A}} |\tilde{i}_\kappa^{ia}|^2 |(\vec{R}_I + \vec{R}_A)/2| \quad (18)$$

where \vec{R}_I (\vec{R}_A) represents the position of F_I (F_A) relative to the center of the cluster, and $\tilde{i} = G^{1/2}t$ so that \tilde{i} is orthonormal between states.

With the same spirit, we can define another quantity $\Delta\bar{R}$ as the weighted average of $|\vec{R}_I - \vec{R}_A|$, so that it serves as a metric of the charge transfer character of each state.

$$\Delta\bar{R}_\kappa = \sum_{\substack{F_I, F_A \in \\ \text{FrgPairList}}} \sum_{\substack{i \in F_I \\ a \in F_A}} |\tilde{i}_\kappa^{ia}|^2 |\vec{R}_I - \vec{R}_A| \quad (19)$$

For each system, \bar{R} and $\Delta\bar{R}$ can be plotted against the excitation energy. As shown in Figure 8(b) and (c), \bar{R} and $\Delta\bar{R}$ are strongly correlated with the excitation energy in droplet systems. For both the 2s and the 2p band, the low-energy excitations correspond to larger \bar{R} and smaller $\Delta\bar{R}$, which indicates that these excitations mainly come from the surface and have less CT character. As the excitation energy increases, bulk excitations with stronger CT character gain more importance, so we see the decrease of \bar{R} and the increase of $\Delta\bar{R}$. For gas systems, it is still true that CT raises the excitation energy, so we can still see the patterns of $\Delta\bar{R}$. However, because the density distribution of gas systems is more uniform, the correlation between excitation energies and the location of the excitation site is much weaker. Thus, the curves of \bar{R} are much flatter for the gas systems, except for the dips at the beginning of each band resulting from the boundary of the simulation sphere.

E. Comparison with Experiment

The size-dependence of helium droplet spectra have been studied by Möller and coworkers^{4,6}. In both Ref 4 and Ref 6, the authors observed that an increase in cluster size results in relative reduction of the atomic peak and relative enhancement of the hump at the high-energy edge. As was already demonstrated in Figure 7, this trend is reproduced by our computational simulation, and its origin is elucidated by the analysis of the surface/bulk character of each state introduced in Section V D. It is quite encouraging that the overall bandwidth seen

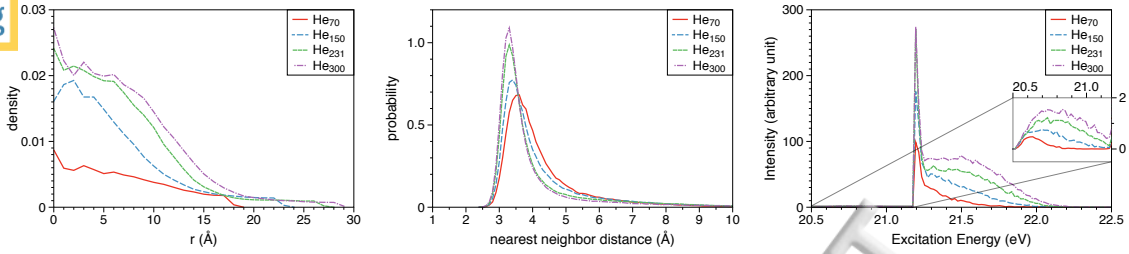


FIG. 7: Density profiles $\rho(r)$, nearest neighbor distributions $nn(r)$ and spectra calculated by ALMO-CIS+CT for He_{70} ($R_c = 18 \text{ \AA}$), He_{150} ($R_c = 23 \text{ \AA}$), He_{231} ($R_c = 27 \text{ \AA}$) and He_{300} ($R_c = 29 \text{ \AA}$) clusters

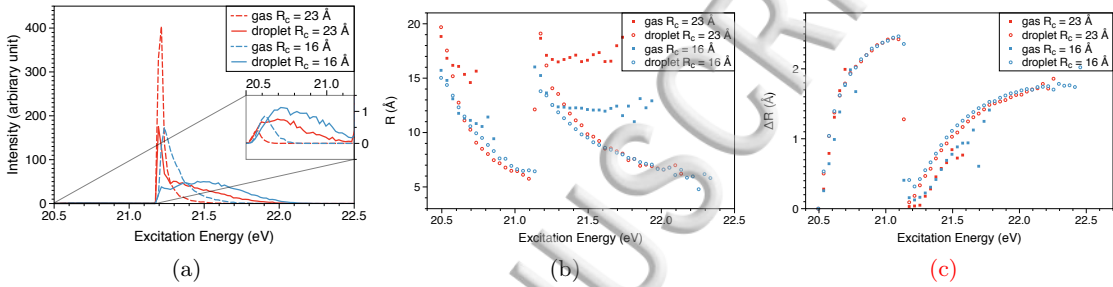


FIG. 8: (a) The spectra of droplet and gas-like He_{150} at $R_c = 16 \text{ \AA}$ and 23 \AA . (b) \bar{R} for the droplet and gas-like at $R_c = 16 \text{ \AA}$ and 23 \AA as a function of excitation energy in the $n = 2$ band. (c) $\Delta\bar{R}$ evaluated in the same way. (both quantities are averaged over 50 geometries for each choice of R_c as in the spectrum calculations).

experimentally is quite well reproduced by the simulation. This suggests that the destabilizing neighbor interactions responsible for the overall shift of the droplet excitations towards higher energy are quite well reproduced by the ALMO-CIS+CT treatment of the electronic structure.

However, it remains a challenge for the theoretically predicted spectra to quantitatively match the experimental results. There are several sources of difference between the simulations and the experiments. First is the fact that the experiments are technically fluorescence excitation spectra, while the calculations correspond to direct absorption. The fluorescence lifetime is on the order of 5-10 ns for “short-lived” excited states⁶, and there is an additional longer-time fluorescence. The nuclear wavefunction is not an eigenstate of the electronically excited state, and the non-stationary wavepacket will certainly evolve, which is likely to affect the spectrum. This is an interesting topic for future investigation. A second issue is that the current PIMD simulation does not treat the Bose statistics of helium nuclei, and thus cannot account for the superfluid-

ity that is present in the experiments. A third issue is that the current electronic structure model has limitations. Specifically, dynamic correlation is neglected in CIS, which results in an over-estimation of excitation energies. For this reason we have to shift all spectra to the left by 0.625 eV so that atomic peaks occur at the same position as in the experiments.

Uncertainty associated with experimental conditions adds further difficulty to a direct comparison of the calculated and observed spectra. The experiments usually prepare helium clusters by a nozzle expansion. The final cluster densities are sensitive to conditions such as nozzle temperature and pressure, and therefore the spectral profiles can vary in different experiments (for example when comparing cluster with similar sizes, spectra in Ref 4 have larger humps than those in Ref 6). As the actual density and other structure characters of helium clusters are hardly known, it is almost impossible to propose a simulation that can reproduce the condition at the experimental measurement. In addition, we believe that the systems are not at equilibrium in the experi-

ment, while in the simulations we impose a confining boundary to obtain reasonably converged results for a system in liquid-vapor equilibrium.

Additionally, the current simulation model appears to miss some detailed features of the experimental spectra. One example is that as stated in Ref 6, a small hump associated with the $2s$ band exists at around 20.95 eV. Experimentally, the position of this maximum is nearly invariant as the cluster size changes, while our calculations show a feature that depends on both cluster size and density (for instance, see the magnified region in Figure 7(c)). It is hard to say whether this is due to the limitation in the electronic structure model or the difference between the simulation and experimental conditions. Despite the above mentioned deficiencies, we think that the PIMD simulation is a clear physical improvement over our previous excited state studies of helium clusters^{13,40}. With the charge transfer effect included via the new ALMO-CIS+CT approach, our methodology also improves the accuracy rendering it comparable to standard CIS, while the former is dramatically more efficient.

VI. CONCLUSIONS

We have generalized the previously reported ALMO-CIS model to include charge transfer (CT) effects. Similar to ALMO-CIS, the ALMO-CIS+CT model presented in this work is a local variant of configuration interaction singles (CIS), and is formulated through the use of absolutely localized molecular orbitals (ALMOs). The difference is that unlike ALMO-CIS, where only intrafragment single excitations are involved, ALMO-CIS+CT retains interfragment excitations whose associated fragment pairs are within a cutoff distance. For helium clusters, the CT effects neglected by ALMO-CIS are mostly recovered when the first shell of neighboring atoms are included. When many excited states are requested (e.g. an entire band of states for a homogeneous cluster), ALMO-CIS+CT has a much lower computational cost (like ALMO-CIS) than standard CIS, while maintaining the same accuracy (in contrast to ALMO-CIS where systematic deviations are evident).

An efficient implementation has been described that minimizes the increase in computational cost versus ALMO-CIS. While the dimension of matrices in the eigenvalue problem is roughly six times larger than that in ALMO-CIS, the computational cost of building the Hamiltonian matrix does not grow enormously. This is partly because one of the most ex-

pensive steps (computing the three-center ERIs in the RI approximation) remains the same, and partly because the implementation of the Coulomb and Exchange integrals has been improved. Solving the eigenvalue problem can potentially be dominant in ALMO-CIS+CT calculations as its cost grows cubically with matrix size. This is remedied by employing a single step Davidson-like variational method without significant loss of accuracy. The overall scaling of ALMO-CIS+CT is third order with respect to system size, which is the same as ALMO-CIS. **For medium sized helium clusters, the computational cost of ALMO-CIS+CT is about four times as large compared to that of ALMO-CIS.** With our current implementation, systems with up to 377 helium atoms and 4147 atomic basis are reported in this work, using only standard workstation-level computer resources.

We apply ALMO-CIS+CT to study the $n = 2$ absorption spectra of helium clusters. To account for the quantum nature of helium nuclei, the geometries used for spectral calculations are generated from PIMD simulation (at 3 K to avoid the need to simulate the superfluid phase). This approach is not perfect due to the absence of indistinguishability between particles, but produces much more reasonable configurations than classical molecular dynamics. We report results on the size-dependence of the spectrum, as well as the effect of a confining boundary on the spectrum. We show that with reasonable choices of the confining radius, the system behaves as a liquid droplet in equilibrium with vapor on the timescale of the simulations. Broad features of the experimental spectrum, such as the bandwidth and size-dependence of the $2p$ band are qualitatively reproduced by the simulations. However, a number of finer details are not fully compatible between the simulations and reported experimental data.

As shown in the appendix, ALMO-CIS is CT-free because it preserves fragment populations in the Mulliken sense. Therefore differences between ALMO-CIS and ALMO-CIS+CT quantify the role of CT in cluster excited states. For helium cluster absorption spectra, states towards the blue edge of the $2p$ band are stabilized by up to 0.5 eV by CT contributions while the red edge is virtually unaffected. The $2s$ states also show CT contributions that increase with excitation energy while they are negligible at the red edge. Even in a system whose electron affinity is as unfavorable as helium, CT effects provide significant excited state stabilization beyond a superposition of atomic excitations, and should not be neglected.

Interesting topics for future work include correct-

ing the model for the neglected effect of dynamic correlations, either by incorporating higher substitutions into the wavefunction, or by extending the model to TDDFT. The applicability of this model to other molecular clusters with stronger interactions between monomers (for instance water clusters) is also potentially very interesting.

VII. ACKNOWLEDGEMENT

This work was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The authors thank Prof. Ryan Steele for the original implementation of PIMD in the Q-Chem package.

Appendix A: Proof that ALMO-CIS is free of charge transfer

Here we compute the Mulliken population on each fragment in the ALMO-CIS+CT model, and show that when ALMO-CIS+CT reduces to ALMO-CIS, the Mulliken population on each fragment is conserved during excitation. In this section, we follow the Einstein summation convention for indices that are summed over the whole space, but explicitly write out the summation if an index is only summed over a certain fragment. Also, a quantity that is built from projected virtuals is indicated by “ $\tilde{\sim}$ ”, for example, \tilde{C}_a^μ is the projected virtual coefficient while C_a^μ is the unprojected virtual coefficient.

The Mulliken population on fragment F_I is defined as

$$\rho_I = \sum_{\mu \in F_I} P^{\mu\nu} S_{\mu\nu} \quad (\text{A1})$$

where \mathbf{P} is the density matrix. It has been proven⁶⁰ that for ground state calculations with the ALMO constraint (SCF(MI)), each fragment’s Mulliken population remains the same as in isolation. So we just need to compute $\sum_{\mu \in F_I} \Delta P^{\mu\nu} S_{\mu\nu}$ (ΔP is the (unrelaxed) difference density matrix) to see if the excitation introduces any change in the fragment Mulliken population. In ALMO-CIS+CT, ΔP

is defined as follows:

$$\begin{aligned} \Delta P^{\mu\nu} &= \Delta P_{OO}^{\mu\nu} + \Delta P_{VV}^{\mu\nu} \\ &= C_i^\mu P_{OO}^{ij} C_j^\nu + \tilde{C}_a^\mu P_{VV}^{ab} \tilde{C}_b^\nu \\ P_{OO}^{ij} &= -t^{ia} \tilde{S}_{ab} t^{jb} \\ P_{VV}^{ab} &= t^{ia} S_{ij} t^{jb} \end{aligned} \quad (\text{A2})$$

The t^{ia} here is the ALMO-CIS+CT amplitude with i and a belonging to a pair of significant fragments.

The sum over the occupied part of Eq. A2 gives:

$$\begin{aligned} \sum_{\mu \in F_I} \Delta P_{OO}^{\mu\nu} S_{\mu\nu} &= - \sum_{\mu, i \in F_I} \sum_{a \in F_J(F_I)} C_i^\mu t^{ia} \tilde{S}_{ab} t^{jb} C_j^\nu S_{\mu\nu} \\ &= - \sum_{i \in F_I} \sum_{a \in F_J(F_I)} t^{ia} \tilde{S}_{ab} t^{jb} \sum_{\mu \in F_I} C_i^\mu S_{\mu\nu} C_j^\nu \\ &= - \sum_{i \in F_I} \sum_{a \in F_J(F_I)} t^{ia} \tilde{S}_{ab} t^{jb} S_{ij} \end{aligned} \quad (\text{A3})$$

where $F_J(F_I)$ means that F_J is a near-neighbor fragment of F_I . The last equality holds because the occupied MO coefficient matrix is block-diagonal, so the restriction on $\mu \in F_I$ can be dropped and $C_i^\mu S_{\mu\nu} C_j^\nu$ is just the occupied overlap metric, S_{ij} .

The virtual part of Eq. A2 is a little more complicated because projected virtuals are used, whose coefficients are not block-diagonal.

$$\begin{aligned} \sum_{\mu \in F_I} \Delta P_{VV}^{\mu\nu} S_{\mu\nu} &= \sum_{\mu \in F_I} \tilde{C}_a^\mu t^{ia} S_{ij} t^{jb} \tilde{C}_b^\nu S_{\mu\nu} \\ &= \sum_{\mu, a \in F_I} \sum_{i \in F_J(F_I)} \tilde{C}_a^\mu t^{ia} S_{ij} t^{jb} \tilde{C}_b^\nu S_{\mu\nu} \\ &+ \sum_{\mu \in F_I} \sum_{F_K \neq F_I} \sum_{a \in F_K} \sum_{i \in F_J(F_K)} \tilde{C}_a^\mu t^{ia} S_{ij} t^{jb} \tilde{C}_b^\nu S_{\mu\nu} \\ &= \sum_{a \in F_I} \sum_{i \in F_J(F_I)} t^{ia} S_{ij} t^{jb} \sum_{\mu \in F_I} \tilde{C}_a^\mu S_{\mu\nu} \tilde{C}_b^\nu \\ &+ \sum_{F_K \neq F_I} \sum_{a \in F_K} \sum_{i \in F_J(F_K)} t^{ia} S_{ij} t^{jb} \sum_{\mu \in F_I} \tilde{C}_a^\mu S_{\mu\nu} \tilde{C}_b^\nu \end{aligned} \quad (\text{A4})$$

For the first sum in Eq. A4, we notice $\tilde{C}_a^\mu = N_a C_a^\mu$, because the diagonal blocks of projected virtuals take contribution from only the unprojected virtuals. On the other hand, the overlap between projected virtual a and b is:

$$\begin{aligned} \tilde{S}_{ab} &= \tilde{C}_a^\mu S_{\mu\nu} \tilde{C}_b^\nu = N_a (C_a^\mu - C_l^\mu (\sigma^{-1})^{lk} S_{ka}) S_{\mu\nu} \tilde{C}_b^\nu \\ &= N_a C_a^\mu S_{\mu\nu} \tilde{C}_b^\nu = \sum_{\mu \in F_I} N_a C_a^\mu S_{\mu\nu} \tilde{C}_b^\nu \end{aligned} \quad (\text{A5})$$

In the second row of Eq. A5, the first equality holds because the projected virtual does not overlap with the occupied, and the second equality comes from the fact that the unprojected virtual coefficients are block-diagonal. In the end, the first term in Eq. A4 becomes $\sum_{a \in F_I} \sum_{i \in F_J(F_I)} t^{ia} S_{ij} t^{jb} \tilde{S}_{ab}$.

As for the second term in Eq. A4, since the off-diagonal projected virtual coefficient comes from the occupied, we have $\tilde{C}_a^\mu = -N_a C_l^\mu (\sigma^{-1})^{lk} S_{ka}$, $\mu, l \in F_I$, $a \in F_K \neq F_I$. The constraint $\mu \in F_I$ can then be dropped as C_l^μ is block diagonal. We then find an \tilde{S}_{lb} in the second term of Eq. A4, which is zero, so that the second term vanishes.

Combining these results, we see that for the ALMO-CIS model, where i and a belong to a same fragment, we have the change in Mulliken charge, $\sum_{\mu \in F_I} \Delta P^{\mu\nu} S_{\mu\nu} = 0$ because $-\sum_{\mu \in F_I} \Delta P_{OO}^{\mu\nu} S_{\mu\nu} = \sum_{\mu \in F_I} \Delta P_{VV}^{\mu\nu} S_{\mu\nu} = \sum_{i,a \in F_I} t^{ia} \tilde{S}_{ab} t^{jb} S_{ij}$.

For ALMO-CIS+CT, unless t^{ia} ($i \in F_I, a \in F_J$) = t^{ia} ($i \in F_J, a \in F_I$) (which could happen for a system consisting of two identical fragments), the occupied and virtual part cannot cancel each other, so the Mulliken population on each fragment is not conserved, unlike for ALMO-CIS.

- ¹S. Grebenev, J. P. Toennies, and A. F. Vilesov, *Science* **279**, 2083 (1998).
- ²C. C. Wang, O. Kornilov, O. Gessner, J. H. Kim, D. S. Peterka, and D. M. Neumark, *J. Phys. Chem. A* **112**, 9356 (2008).
- ³D. S. Peterka, J. H. Kim, C. C. Wang, and D. M. Neumark, *J. Phys. Chem. B* **110**, 19945 (2006).
- ⁴M. Joppien, R. Karnbach, and T. Möller, *Phys. Rev. Lett.* **71**, 2654 (1993).
- ⁵K. von Haeften, T. Laarmann, H. Wabnitz, and T. Möller, *Phys. Rev. Lett.* **87**, 153403 (2001).
- ⁶K. von Haeften, T. Laarmann, H. Wabnitz, T. Möller, and K. Fink, *J. Phys. Chem. A* **115**, 7316 (2011).
- ⁷M. A. Marques and E. K. U. Gross, *Annu. Rev. Phys. Chem.* **55**, 427 (2004).
- ⁸S. Van Gisbergen, J. Snijders, and E. Baerends, *Comput. Phys. Commun.* **118**, 119 (1999).
- ⁹A. Dreuw and M. Head-Gordon, *Chem. Rev.* **105**, 4009 (2005).
- ¹⁰M. E. Casida, *J. Mol. Struct-Theochem* **914**, 3 (2009).
- ¹¹J. B. Foresman, M. Head-Gordon, J. A. Pople, and M. J. Frisch, *J. Phys. Chem.* **96**, 135 (1992).
- ¹²M. Head-Gordon, R. J. Rico, M. Oumi, and T. J. Lee, *Chem. Phys. Lett.* **219**, 21 (1994).
- ¹³K. D. Closser and M. Head-Gordon, *J. Phys. Chem. A* **114**, 8023 (2010).
- ¹⁴K. D. Closser, O. Gessner, and M. Head-Gordon, *J. Chem. Phys.* **140**, 134306 (2014).
- ¹⁵E. R. Davidson, *J. Comput. Phys.* **17**, 87 (1975).
- ¹⁶R. E. Stratmann, G. E. Scuseria, and M. J. Frisch, *J. Chem. Phys.* **109**, 8218 (1998).
- ¹⁷Q. Li, Q. Li, and Z. Shuai, *Synth. Met.* **158**, 330 (2008).

- ¹⁸M. Miura and Y. Aoki, *J. Comput. Chem.* **30**, 2213 (2009).
- ¹⁹M. Miura and Y. Aoki, *Mol. Phys.* **108**, 205 (2010).
- ²⁰J. E. Subotnik, A. D. Dutoi, and M. Head-Gordon, *J. Chem. Phys.* **123**, 114108 (2005).
- ²¹Z. Li, H. Li, B. Suo, and W. Liu, *Acc. Chem. Res.* **47**, 2758 (2014).
- ²²K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, *Chem. Phys. Lett.* **313**, 701 (1999).
- ²³Y. Mochizuki, S. Koikegami, S. Amari, K. Segawa, K. Kitaura, and T. Nakano, *Chem. Phys. Lett.* **406**, 283 (2005).
- ²⁴Y. Mochizuki, K. Tanaka, K. Yamashita, T. Ishikawa, T. Nakano, S. Amari, K. Segawa, T. Murase, H. Tokiwa, and M. Sakurai, *Theor. Chem. Acc.* **117**, 541 (2007).
- ²⁵M. Chiba, D. G. Fedorov, and K. Kitaura, *Chem. Phys. Lett.* **444**, 346 (2007).
- ²⁶F. Wu, W. Liu, Y. Zhang, and Z. Li, *J. Chem. Theory Comput.* **7**, 3643 (2011).
- ²⁷A. F. Morrison, Z.-Q. You, and J. M. Herbert, *J. Chem. Theory Comput.* **10**, 5366 (2014).
- ²⁸A. F. Morrison and J. M. Herbert, *J. Phys. Chem. Lett.* **6**, 4390 (2015).
- ²⁹J. M. Herbert, X. Zhang, A. F. Morrison, and J. Liu, *Acc. Chem. Res.* **49**, 931 (2016).
- ³⁰H. Stoll, G. Wagenblast, and H. Preuß, *Theor. Chem. Acc.* **57**, 169 (1980).
- ³¹J. M. Cullen, *Int. J. Quantum Chem.* **40**, 193 (1991).
- ³²E. Gianinetti, A. Vandoni, A. Famulari, and M. Raimondi, *Adv. Quantum Chem.* **31**, 251 (1998).
- ³³T. Nagata, O. Takahashi, K. Saito, and S. Iwata, *J. Chem. Phys.* **115**, 3553 (2001).
- ³⁴R. Z. Khaliullin, M. Head-Gordon, and A. T. Bell, *J. Chem. Phys.* **124**, 204105 (2006).
- ³⁵Y. Mo, J. Gao, and S. D. Peyerimhoff, *J. Chem. Phys.* **112**, 5530 (2000).
- ³⁶R. Z. Khaliullin, E. A. Cobar, R. C. Lochan, A. T. Bell, and M. Head-Gordon, *J. Phys. Chem. A* **111**, 8753 (2007).
- ³⁷P. R. Horn, E. J. Sundstrom, T. A. Baker, and M. Head-Gordon, *J. Chem. Phys.* **138**, 134119 (2013).
- ³⁸J. Thirman and M. Head-Gordon, *J. Chem. Phys.* **143**, 084124 (2015).
- ³⁹P. R. Horn, Y. Mao, and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **18**, 23067 (2016).
- ⁴⁰K. D. Closser, Q. Ge, Y. Mao, Y. Shao, and M. Head-Gordon, *J. Chem. Theory Comput.* **11**, 5791 (2015).
- ⁴¹J. Liu and J. M. Herbert, *J. Chem. Phys.* **143**, 034106 (2015).
- ⁴²M. Head-Gordon, P. E. Maslen, and C. A. White, *J. Chem. Phys.* **108**, 616 (1998).
- ⁴³Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio, R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khaliullin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer, J. Kong, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **8**, 3172 (2006).

- Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormi, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuš, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyayev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. Gill, and M. Head-Gordon, *Mol. Phys.* **113**, 184 (2015).
- ⁴⁵X. Ren, P. Rinke, V. Blum, J. Wierfink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler, *New J. Phys.* **14**, 053020 (2012).
- ⁴⁶B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, *J. Chem. Phys.* **71**, 3396 (1979).
- ⁴⁷O. Vahtras, J. Almlöf, and M. W. Feyereisen, *Chem. Phys. Lett.* **213**, 514 (1993).
- ⁴⁸W. Keesom and K. Taconis, *Physica* **5**, 270 (1938).
- ⁴⁹B. E. Callicoatt, K. Forde, L. F. Jung, T. Ruchti, and K. C. Janda, *J. Chem. Phys.* **109**, 10195 (1998).
- ⁵⁰R. P. Feynman, A. R. Hibbs, and D. F. Styer, *Quantum mechanics and path integrals* (Courier Corporation, 2005).
- ⁵¹D. M. Ceperley, *Rev. Mod. Phys.* **67**, 279 (1995).
- ⁵²C. Cleveland, U. Landman, and R. Barnett, *Phys. Rev. B* **39**, 117 (1989).
- ⁵³F. Calvo, F. Naumkin, and D. Wales, *J. Chem. Phys.* **135**, 124308 (2011).
- ⁵⁴P. Sindzingre, M. L. Klein, and D. M. Ceperley, *Phys. Rev. Lett.* **63**, 1601 (1989).
- ⁵⁵R. Aziz, V. Nain, J. Carley, W. Taylor, and G. McConville, *J. Chem. Phys.* **70**, 4330 (1979).
- ⁵⁶G. Bussi and M. Parrinello, *Phys. Rev. E* **75**, 056707 (2007).
- ⁵⁷M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos, *J. Chem. Phys.* **133**, 124104 (2010).
- ⁵⁸D. S. Peterka, J. H. Kim, C. C. Wang, L. Poisson, and D. M. Neumark, *J. Phys. Chem. A* **111**, 7449 (2007).
- ⁵⁹M. Herman, E. Bruskin, and B. Berne, *J. Chem. Phys.* **76**, 5150 (1982).
- ⁶⁰R. Z. Khaliullin, A. T. Bell, and M. Head-Gordon, *J. Chem. Phys.* **128**, 184112 (2008).