

Approximate variational coupled cluster theory

James B. Robinson and Peter J. Knowles

Citation: *The Journal of Chemical Physics* **135**, 044113 (2011); doi: 10.1063/1.3615060

View online: <http://dx.doi.org/10.1063/1.3615060>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/135/4?ver=pdfcov>

Published by the [AIP Publishing](#)



Re-register for Table of Content Alerts

Create a profile.



Sign up today!



Approximate variational coupled cluster theory

James B. Robinson and Peter J. Knowles^{a)}

School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff CF10 3AT, United Kingdom

(Received 16 May 2011; accepted 1 July 2011; published online 27 July 2011)

We show that it is possible to construct an accurate approximation to the variational coupled cluster method, limited to double substitutions, from the minimization of a functional that is rigorously extensive, exact for isolated two-electron subsystems and invariant to transformations of the underlying orbital basis. This approximate variational coupled cluster theory is a modification and enhancement of our earlier linked pair functional theory. It is first motivated by the constraint that the inverse square root of the matrix that transforms the cluster amplitudes must exist. Low-order corrections are then included to enhance the accuracy of the approximation of variational coupled cluster, while ensuring that the computational complexity of the method never exceeds that of the standard traditional coupled cluster method. The effects of single excitations are included by energy minimization with respect to the orbitals defining the reference wavefunction. The resulting quantum chemical method is demonstrated to be a robust approach to the calculation of molecular electronic structure and performs well when static correlation effects are strong. © 2011 American Institute of Physics. [doi:10.1063/1.3615060]

I. INTRODUCTION

A central theoretical and computational problem in the fields of quantum chemistry and molecular physics is the purely *ab initio* prediction of molecular electronic structure. The standard and widely accepted approach to the treatment of this problem is to first make the Hartree-Fock^{1,2} (HF) approximation, solving the resulting equations by a self-consistent field procedure, and then to compute corrections to this approximation by some scheme. The broad term “electron correlation” is used to describe these corrections, which can be further conceptually divided into “dynamic” and “static” correlation. These terms refer, respectively, to the instantaneous correlated motion of the electrons that is not captured by the HF mean-field treatment and to the character of the system that cannot be captured easily from an underlying single-configuration reference wavefunction approximation.

Low-order corrections may be determined from some form of many-body perturbation theory, such as the Møller-Plesset method,³ with the HF Slater determinant playing the role of the zeroth-order wavefunction. However, the convergence of the perturbation series is often a troubling consideration⁴⁻⁶ and motivates the use of other theoretical methodologies, although these are often used in conjunction with perturbation theory for reasons of cost.

Of the alternative approaches, the conceptually (although not computationally) simplest is configuration interaction⁷ (CI), in which the exact wavefunction is expanded linearly in the basis of excited determinants. When all possible excitations up to the number of electrons are included it is referred to as full configuration interaction (FCI). Unfortunately, the determinantal series typically must be truncated for computational practicality, say, to include single and double excitations only (CISD). The method is then not extensive, meaning

that calculated energies do not scale linearly with the size of the system under consideration.

The widely used and highly successful coupled cluster⁸⁻¹⁰ (CC) method has the advantage over variational CI that its energy is rigorously extensive. However, when the CC equations are determined by projection of the time-independent Schrödinger equation against a basis (referred to as traditional coupled cluster (TCC) or simply CC), its energy is not a variational upper bound to the exact energy. Variational optimization of the energy expectation value using the CC wavefunction is referred to as variational coupled cluster¹¹ (VCC), but this is typically not computationally feasible.

Before the emergence of the CC theory, corrections of the CI method and other simple approximations, each of which attempts to eliminate the unphysical unlinked cluster contributions, were in use. These included the coupled pair functional,¹² the coupled electron pair approximation¹³⁻¹⁵ (CEPA), the linear coupled pair many-electron theory,^{16,17} also known as CEPA(0), and the very simple Davidson correction.¹⁸ CC has the advantage over the CEPA-type methods that it is both equivalent to FCI when the cluster operator is not truncated and is invariant to rotations of the underlying orbital spaces. However, more recent work¹⁹ has resulted in more sophisticated, orbitally invariant approaches derived from a VCC viewpoint.

When an underlying single-configuration reference wavefunction is a good approximation, coupled cluster with single and double excitations (CCSD) performs well. However, when static correlation effects become large and this approximation breaks down, such as in cases of the breaking of multiple bonds, CCSD can fail, predicting unphysical potential energy surfaces. On the other hand, a number of studies²⁰⁻²² have confirmed the superiority of VCCSD and related approaches over CCSD in such situations. This motivates a search for an approximation to VCC that remains

^{a)}Electronic mail: KnowlesPJ@Cardiff.ac.uk.

competitive with TCC both in terms of the computational cost and the methodological properties of the approach.

II. THEORY

A. Coupled cluster theory

The fundamental ansatz of coupled cluster theory is the exponential parameterization of the wavefunction

$$|\Phi\rangle = e^{\hat{T}}|\Phi_0\rangle, \quad (1)$$

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots, \quad (2)$$

where $|\Phi\rangle$ and $|\Phi_0\rangle$ are, respectively, the exact and (single configuration) reference wavefunctions. Each cluster operator, \hat{T}_n , generates a weighted sum of all possible n -fold excitations of the reference Slater determinant. An excitation is a replacement of an orbital occupied in the reference with one unoccupied in the reference. Typically, the cluster operator is truncated to include single and double excitations only, for reasons of computational cost

$$\hat{T}_1|\Phi_0\rangle = T_a^i|\Phi_i^a\rangle, \quad (3)$$

$$\hat{T}_2|\Phi_0\rangle = \frac{1}{4}T_{ab}^{ij}|\Phi_{ij}^{ab}\rangle. \quad (4)$$

The Einstein summation convention is used throughout this paper, so that a repeated index is understood to imply a summation over all possible values of that index. The indices i, j, k, \dots are used to label occupied orbitals and a, b, c, \dots to label virtual orbitals.

The traditional approach (TCC) to the approximate calculation of the ground-state Schrödinger energy eigenvalue is to insert the CC wavefunction ansatz into the Schrödinger equation and, in order to decouple the amplitude equations from the energy, premultiply by $e^{-\hat{T}}$,

$$e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = E_{\text{TCC}}|\Phi_0\rangle. \quad (5)$$

The energy, E_{TCC} , and equations to be solved for the cluster amplitudes can then be determined by projection. For example, in the CCSD method, for which $\hat{T} = \hat{T}_1 + \hat{T}_2$, there are three sets of equations that are acquired by projection with the reference determinant and each of the singly and doubly excited determinants

$$\begin{aligned} \langle\Phi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle &= E_{\text{TCC}}, \\ \langle\Phi_i^a|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle &= 0, \\ \langle\Phi_{ij}^{ab}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle &= 0. \end{aligned} \quad (6)$$

The TCC approach possesses many properties that make it a powerful method for the accurate calculation of ground-state molecular electronic structure, such as its rigorous extensivity and exactness (equivalence to FCI) when the cluster operator is not truncated. Unfortunately, it also has some troubling characteristics. Energies calculated by the TCC method are not variational upper bounds on the exact ground-state Schrödinger energy eigenvalue, and the symmetric generalized Hellmann-Feynman theorem²³ is not satisfied. These problems can be understood by noticing that Eq. (5) can be

thought of as a Schrödinger equation with a similarity transformed Hamiltonian, $\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$. The operator \bar{H} is not Hermitian, $\bar{H}^\dagger \neq \bar{H}$.

Despite the widespread acceptance of the TCC approach, a method that corrects or approximately corrects one or both of the above problematic properties would be highly desirable.

A more intuitive way to calculate an approximate ground-state energy within the coupled cluster ansatz would be to variationally minimize the quantum mechanical energy expectation value with respect to the cluster amplitudes. This is the VCC approach

$$E_{\text{VCC}} = \frac{\langle\Phi_0|e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}|\Phi_0\rangle}{\langle\Phi_0|e^{\hat{T}^\dagger}e^{\hat{T}}|\Phi_0\rangle} = \langle\Phi_0|e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}|\Phi_0\rangle_L. \quad (7)$$

In addition to extensivity and equivalence to FCI when the cluster operator is not truncated, this approach also has the additional advantages that, even with a truncated cluster operator, the VCC energy is an upper bound to the exact ground-state Schrödinger energy eigenvalue, and that it satisfies the symmetric generalized Hellmann-Feynman theorem. Indeed, it possesses *all* of the properties typically agreed to be important for a successful quantum chemical method, which makes VCC extremely desirable.

Unfortunately, VCC is typically far more computationally expensive than TCC. For example, consider the computational costs at the singles and doubles level. If N is the number of electrons, the computational complexity of CCSD is $\mathcal{O}(N^6)$ (more precisely $\mathcal{O}(o^2v^4)$, where o is the number of occupied orbitals and v is the number of virtual orbitals, where, typically $v \gg o$), whereas VCCSD scales as $\mathcal{O}(N!)$. The difference lies in the natural truncation of \bar{H} due to the Campbell-Baker-Hausdorff formula,²⁴ whereas the VCCSD functional in energy expectation value form does not truncate naturally except at the N -tuply excited limit. When the VCCSD denominator is cancelled with the unlinked parts of the numerator, uncanceled exclusion-principle-violating terms remain²⁵ and no natural truncation occurs at all; the expression is infinite.

The additional stability of VCCSD over CCSD can be attributed to the variational upper bound property, and the theoretical attractiveness of VCCSD motivates the search for viable approximations that may be computed within CCSD-like $\mathcal{O}(o^2v^4)$ complexity.

B. Approximating VCCD

Since the effect of single excitations is usually small, a first goal would be to approximate VCCD,

$$E_{\text{VCCD}} = \frac{\langle e^{\hat{T}_2^\dagger}\hat{H}e^{\hat{T}_2}\rangle}{\langle e^{\hat{T}_2^\dagger}e^{\hat{T}_2}\rangle} = \langle e^{\hat{T}_2^\dagger}\hat{H}e^{\hat{T}_2}\rangle_L. \quad (8)$$

In order to do this, the most obvious approach is to simply truncate the exponential operator, $e^{\hat{T}_2}$, at some level, such as to $1 + \hat{T}_2$. It is easily demonstrable that such a truncation yields a method that cannot be both extensive and exact (in the sense that it is equivalent to CID) in the limiting case of a two-electron system, simply by considering the VCCD

energy expression. Truncating the exponential to $1 + \hat{T}_2$ in the first form of the energy in Eq. (8) yields CID, which is known not to be extensive

$$E_{\text{CID}} = \frac{\langle \hat{H} \rangle + 2\langle \hat{H} \hat{T}_2 \rangle + \langle \hat{T}_2^\dagger \hat{H} \hat{T}_2 \rangle}{1 + \langle \hat{T}_2^\dagger \hat{T}_2 \rangle} \\ = \langle \hat{H} \rangle + \frac{2\langle \hat{H} \hat{T}_2 \rangle + \langle \hat{T}_2^\dagger (\hat{H} - \langle \hat{H} \rangle) \hat{T}_2 \rangle}{1 + \langle \hat{T}_2^\dagger \hat{T}_2 \rangle}. \quad (9)$$

On the other hand, the same truncation of the exponential in the second form of the energy yields CEPA(0), which is not exact even for two electrons,

$$E_{\text{CEPA(0)}} = \langle \hat{H} \rangle + 2\langle \hat{H} \hat{T}_2 \rangle + \langle \hat{T}_2^\dagger \hat{H} \hat{T}_2 \rangle_L \\ = \langle \hat{H} \rangle + 2\langle \hat{H} \hat{T}_2 \rangle + \langle \hat{T}_2^\dagger (\hat{H} - \langle \hat{H} \rangle) \hat{T}_2 \rangle. \quad (10)$$

One could conceive of truncating the exponential later in the infinite series, thus acquiring higher order VCCD terms. Disregarding the fact that this does not solve the above problem, even extending the truncation to $1 + \hat{T}_2 + (1/2!)\hat{T}_2^2$ would yield a method that scales as $\mathcal{O}(v^6)$ in computational complexity due to the presence of the term $(1/4)\langle \hat{T}_2^\dagger \hat{H} \hat{T}_2^2 \rangle$. This is much worse than CCSD.

Asymmetric expectation value expressions have also been proposed, such as improved coupled cluster²⁶ (ICC) and extended coupled cluster²⁷ (ECC). These approaches suggest hierarchies of methods that step systematically from TCC to VCC depending on the level of truncation employed. However, in addition to the increased computational complexity over TCC, these approaches have the disadvantage that one has no rigorous guarantee of extensivity at arbitrary truncation, except at the extremes of the hierarchies.

C. Linked pair functional theory

The recently proposed²⁸ doubles-only linked pair functional (LPDF) may be thought of as an approximation to VCCD that maintains the essential properties of exactness and extensivity. The LPDF ground-state energy is defined to be the minimum of the LPDF functional with respect to the cluster amplitudes T_{ab}^{ij} ,

$$E_{\text{LPDF}} = \langle \hat{H} \rangle + 2\langle \hat{H} \hat{T}_2 \rangle + \langle \hat{T}_2^\dagger (\hat{H} - \langle \hat{H} \rangle) \hat{T}_2 \rangle. \quad (11)$$

This functional is, of course, related to both CID and CEPA(0). The noteworthy difference is the presence of the ${}_q\hat{T}$ operators, which are modified double excitation operators. These are defined by their action on the reference wavefunction,

$${}_q\hat{T}|\Phi_0\rangle = \frac{1}{4}{}_qT_{ab}^{ij}|\Phi_{ij}^{ab}\rangle, \quad (12)$$

$${}_qT_{ab}^{ij} = \frac{1}{2}({}_q\mathbf{U}^{qP_u})_{kl}^{ij} T_{ab}^{kl}, \quad (13)$$

$${}_q\mathbf{U}^{ij} = \delta_{kl}^{ij} + {}_qS_u \Delta_{kl}^{ij}, \quad (14)$$

$$\Delta_{kl}^{ij} = \lambda \eta_{kl}^{ij} + \frac{1}{2}(1 - \lambda)(\delta_k^i \eta_l^j - \delta_l^j \eta_k^i - \delta_k^j \eta_l^i + \delta_l^i \eta_k^j) \\ = \lambda \eta_{kl}^{ij} + \frac{1}{2}(1 - \lambda)(1 - \tau_{ij})(1 - \tau_{kl})\delta_k^i \eta_l^j, \quad (15)$$

where τ_{ij} permutes the labels i, j in what follows, and

$$\eta_{kl}^{ij} = \langle \hat{T}^\dagger k l j^\dagger i^\dagger \hat{T} \rangle = \frac{1}{2} T_{ab}^{ij} T_{kl}^{ab}, \\ \eta_j^i = \langle \hat{T}^\dagger j i^\dagger \hat{T} \rangle = \frac{1}{2} T_{ab}^{ik} T_{jk}^{ab} \equiv \eta_{jk}^{ik}. \quad (16)$$

The quantities ${}_qP_u$ and ${}_qS_u$ are rational parameters chosen to satisfy certain conditions (discussed below), and the quantity λ is a continuously adjustable real scalar parameter.

One can define powers of ${}_q\mathbf{U}$, if one considers the indices labelling ${}_q\mathbf{U}$ to be the unique pair indices,

$$ij \in \{kl : k \in \{1 \dots N\}, l \in \{1 \dots (k-1)\}\},$$

where N is the number of electrons, in which case summations should be rewritten as summations over unique pairs. It is then clear that ${}_q\mathbf{U}$ is simply a $N_p \times N_p$ matrix, where $N_p = N(N-1)/2$, the number of electron pairs. Powers of ${}_q\mathbf{U}$ then are defined simply through the powers of this matrix.

The purpose of introducing the transformed cluster amplitudes ${}_qT_{ab}^{ij}$ is to ensure exactness in the limiting case of isolated two-electron subsystems, since then the matrix ${}_q\mathbf{U}$ contains the CID square norms of the isolated pairs. The powers, ${}_qP_u$, and coefficients, ${}_qS_u$, are fixed by the requirement that the square norms enter as local denominators for each term in the numerator containing the appropriate cluster amplitudes

$${}_qS_u = 1, \\ {}_qP_u = -\frac{q}{2}. \quad (17)$$

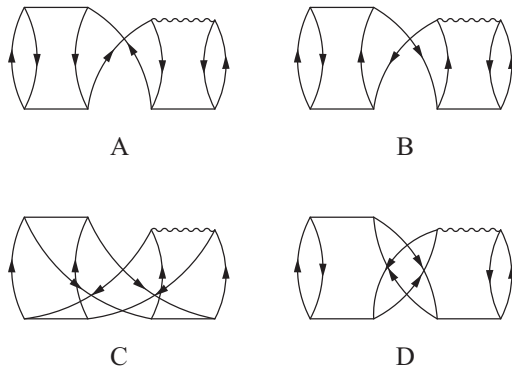
Independent of ${}_qP_u$ and ${}_qS_u$, this theory is extensive since no unlinked contributions are introduced. Since, in the case of a two-electron system, CID and VCCD are indistinguishable, LPDF is, in this sense, also an approximation to VCCD.

Of course, one may instead interpret the effect of the amplitude transformations as introducing an infinite series of terms via the binomial expansion of the powered ${}_q\mathbf{U}$,

$${}_q\mathbf{U}^{qP_u} = (\delta + {}_qS_u \delta)^{qP_u} \\ = \delta + {}_qS_u {}_qP_u \delta + \dots \quad (18)$$

It is in this conceptual form that the correspondences with VCCD are most clear. LPDF can be understood to capture a subset of the terms in the infinite series $\langle e^{\hat{T}_2^\dagger \hat{H} e^{\hat{T}_2}} \rangle_L$. In the two-electron limit, cancellations between VCCD terms occur within each order of T , and, independent of λ , LPDF always captures a subset of VCCD terms, appropriately weighted, that gives the correct two-electron behaviour. The value of λ controls the specific subset that is captured and the weightings of the terms within it.

This is most easily understood by considering the example of the $\mathcal{O}(T^3)$ VCCD terms arising from $\langle \hat{T}_2^\dagger \hat{H} \hat{T}_2^2 \rangle_L$. This is the first interesting term in VCCD, since the LPDF functional captures up to $\mathcal{O}(T^2)$ in VCCD exactly even without the amplitude transformations, since then it is just CEPA(0). The individual contributions to this term are given diagrammatically in Fig. 1. In the limit of two electrons, these terms

FIG. 1. $\mathcal{O}(T^3)$ terms contributing to VCCD.

obey the following relationships:

$$\begin{aligned} \mathbf{A} + \mathbf{D} &= 0, \\ \mathbf{B} + 2\mathbf{C} &= 0. \end{aligned} \quad (19)$$

The same two-electron contribution may then be captured in multiple ways

$$\mathbf{A} + \mathbf{B} + \mathbf{C} + \mathbf{D} = \mathbf{B} + \mathbf{C} = -\mathbf{C} = \frac{1}{2}\mathbf{B}. \quad (20)$$

Similar cancellations have been noted by Huntington and Nooijen²⁹ also to occur in the TCCSD residual.

By construction, all values of λ give rise to some combination of the diagrams **B** and **C**. For example, $\lambda = 0$ yields $(1/2)\mathbf{B}$ and $\lambda = +1$ yields $-\mathbf{C}$. If the goal is an approximate VCCD theory, $\lambda = -1$ is the best choice, since it yields both terms exactly as they appear in VCCD, $\mathbf{B} + \mathbf{C}$. This is the best that can be done from this ansatz since $\mathbf{A} + \mathbf{D}$ is not obtainable from ${}_q\mathbf{U}$.

Independent of λ , LPFD possesses most of the properties of electronic structure methods that are widely agreed to be either important or essential.

- It is exactly equivalent to CID in the case of two electrons, since in that case ${}_q\mathbf{U}$ is a scalar and equal to the norm of the wavefunction, $1 + \langle \hat{T}^\dagger \hat{T} \rangle$.
- The expression can be written in terms of linked diagrams only and is therefore rigorously extensive.
- The energy is a scalar that is invariant to rotations of the underlying orbital spaces $\{\psi_i\}$ and $\{\psi_a\}$.
- The energy is not an upper bound to the exact ground-state Schrödinger eigenvalue, but variational minimization of the functional gives a theory that satisfies the symmetric generalized Hellmann-Feynman theorem.

D. Positivity and the linked pair functional

Since much discussion will be given to different choices of λ in the LPFD theory, the notation LPFD(λ) is occasionally employed.

By inspection, it can be seen that the η matrices of Eq. (16) are Gram matrices, and so are necessarily positive-semidefinite. However, it does not immediately follow that $\mathbf{\Delta}$ is also positive-semidefinite. We are unable to construct

an argument for the positivity of ${}_q\mathbf{U}$ for arbitrary λ and, in fact, there exists a numerical counter-example for the case of $\lambda = -1$.

Since this is a compelling disproof of the positivity of ${}_q\mathbf{U}$ in the general case, one cannot be sure that its matrix inverse and inverse square root exist, implying a failure of the approach. This presents a strong constraint on the available choices of λ , one which it seems only $\lambda = +1$ satisfies within the current formulation of the theory, since then the positive-semidefiniteness of $\mathbf{\Delta}$ follows from its equality to η .

Rather than proceed with this choice, we consider a reformulation and simplification of the theory, which is suggested by considering the form of the first order contributions (of the binomial expansion of the powered ${}_q\mathbf{U}$) to the transformed amplitudes with the choice $\lambda = 0$,

$$\begin{aligned} \frac{1}{2}\Delta_{kl}^{ij} T_{ab}^{kl} &= \frac{1}{4}(1 - \tau_{ij})(1 - \tau_{kl})\delta_k^i \eta_l^j T_{ab}^{kl} \\ &= \frac{1}{2}(1 - \tau_{ij})\eta_k^i T_{ab}^{kj}. \end{aligned} \quad (21)$$

This suggests the following new definitions:

$${}_qU_j^i = \delta_j^i + {}_qS_u \Delta_j^i, \quad (22)$$

$$\Delta_j^i = \eta_j^i, \quad (23)$$

$${}_qT_{ab}^{ij} = \frac{1}{2}(1 - \tau_{ij}) ({}_q\mathbf{U}^{qP_u})_k^i T_{ab}^{kj}. \quad (24)$$

Since here $\Delta_j^i = \eta_j^i$, Δ_j^i is an element of a positive-semidefinite matrix. Obviously, δ_j^i is an element of a positive-definite matrix, the identity, for which all the eigenvalues are 1. Hence, ${}_q\mathbf{U}$ will be a positive-definite matrix and its inverse and inverse square root always exist.

This is very much like doing LPFD(0) in a positive-definite way. It has the aesthetic property that ${}_q\mathbf{U}$ is now indexed by electrons, rather than by electron pairs, and so the matrix is significantly smaller in general. Of course, this is *not* LPFD(0); the agreement between LPFD(0) and this new formulation of the theory is lost at high orders since the different definitions of matrix powers produce different terms. For example, in LPFD(0), $\eta_k^i \eta_l^j T_{ab}^{kl}$ is a contribution to $(1/2)(\mathbf{\Delta}^2)_{kl}^{ij} T_{ab}^{kl}$, but $(\eta^2)_j^i = \eta_k^i \eta_j^k$ and cannot produce this term. However, LPFD(0) and this new theory agree exactly to $\mathcal{O}(T^4)$ so, numerically, the differences are expected to be small.

All of the noteworthy properties of LPFD are preserved in this reformulation, so there is no obvious disadvantage to this approach. However, there is also no reason to believe it is, as an approximation to VCCD, a *better* theory than simply making use of LPFD(+1). For now, we adopt this new theory due to its theoretical and computational simplicity, but future work will address whether LPFD(+1) is a better underlying theory for approximate variational coupled cluster, to be presented in Sec. II F. In what follows, when we refer to LPFD(0) or simply LPFD, we mean this new approximate $\lambda = 0$ theory, unless otherwise stated.

E. The replacement of single excitations by orbital optimization

In order for LPF to be a competitive *ab initio* quantum chemical method, it must incorporate both single and double excitations. Extending the LPFD theory to explicitly include singles is not a trivial task, and we defer this problem to later work. In our earlier work,²⁸ the Brueckner orbitals from a preceding BCCD calculation³⁰ were instead used. These are a set of orbitals in which the optimum singles amplitudes in a CCSD calculation vanish. Thus, when these orbitals are used, a doubles-only functional that is equivalent to CID in the limiting case of two electrons (such as LPFD) is also equivalent to FCI.

While this approach suffices, it possesses two disadvantages over a more inclusive approach. First, the requirement of performing a calculation of the same computational complexity, $\mathcal{O}(N^6)$, as LPFD prior to the LPFD calculation itself undermines the simplicity of the approach. Second, the BCCD equations, Eq. (25), are derived from a projective approach analogous to CCSD, rather than the requirement that the energy is at a variational minimum,

$$\begin{aligned} \langle \Phi_0 | \hat{H}(1 + \hat{T}_2) | \Phi_0 \rangle &= E, \\ \langle \Phi_i^a | \hat{H}(1 + \hat{T}_2) | \Phi_0 \rangle &= 0, \\ \langle \Phi_{ij}^{ab} | \hat{H}(1 + \hat{T}_2 + \frac{1}{2} \hat{T}_2^2) | \Phi_0 \rangle &= T_{ij}^{ab} E. \end{aligned} \quad (25)$$

Both of these problems can be solved by asserting that the LPFD functional should be variationally minimized with respect to both the doubles amplitudes and the orbitals. The first of these conditions is satisfied by the requirement that the derivative of the functional with respect to the doubles vanishes, just as in the doubles-only case. The second condition can be derived and implemented straightforwardly as shown by Kollmar and Heßelmann³¹ in the case of CEPA and some other pair methods. The orbital gradient for LPFD is then obtained by setting, in the language of Ref. 31, $f_{ijkl} = 0$, to obtain the CEPA(0) orbital gradient, and then replacing each B in f_{ia} with ${}_2B$ or ${}_1B$ for terms linear and quadratic in B , respectively.

When the LPFD functional is optimized with respect to both the doubles amplitudes and the orbitals, we choose to denote it as OLPFD. We have also implemented optimization of the energy functional with respect to doubles amplitudes only, with the constraint that the Brueckner condition

$$\langle \Phi_i^a | \hat{H}(1 + {}_1\hat{T}_2) | \Phi_0 \rangle = 0 \quad (26)$$

is satisfied, and we denote this ansatz BLPFD. The OLPFD has the advantage that it is fully variational, making the calculation of properties simpler, since solution of the linear equations that determine the Lagrangian multipliers³² is unnecessary. However, BLPFD is computationally cheaper since the recalculation of integrals with three external indices in each iteration can be avoided.³³

F. Approximate variational coupled cluster theory

Regardless of the choice made between LPFD(+1) and LPFD(0) doubles-only theories, neither is a particularly good approximation to VCCD, each acquiring just one of the four unique $\mathcal{O}(T^3)$ VCCD terms. The LPFD(-1) would be superior since it acquires two of the four terms with the correct weightings. However, even if it had not been discarded already on the grounds of positivity, improvements could still be added.

In this section, low-order corrections are added to the doubles-only theory, to give the best possible correspondence with VCCD. These corrections do not increase the computational complexity of the theory above $\mathcal{O}(o^2v^4)$, so that it is still competitive with CCSD, and do not disturb any of the highly desirable properties of LPFD, particularly exactness for isolated two-electron subsystems and extensivity. This is ensured by formulating the corrections as additional transformation matrices. We denote this new method approximate variational coupled cluster (AVCC). In the present work, the functional is restricted to double excitations (AVCCD), and the effect of single excitations is taken into account by orbital optimization (OAVCCD), or through the Brueckner condition (BAVCCD).

The OAVCCD ground-state energy is the minimum of the AVCCD functional with respect to the cluster amplitudes and orbitals,

$$E_{AVCCD} = \langle \hat{H} \rangle + 2\langle \hat{H} {}_2\hat{T} \rangle + \langle {}_1\hat{T}^\dagger (\hat{H} - \langle \hat{H} \rangle) {}_1\hat{T} \rangle, \quad (27)$$

$${}_q\hat{T} | \Phi_0 \rangle = \frac{1}{4} {}_qT_{ab}^{ij} | \Phi_{ij}^{ab} \rangle. \quad (28)$$

In addition to the use of the ${}_q\mathbf{U}$ transformation, two new transformations are defined, the ${}_q\mathbf{W}$ and ${}_q\mathbf{V}$ transformations. Each transformation makes use of the $\boldsymbol{\eta}$ matrices, to which some new definitions are added,

$$\begin{aligned} \eta_{kl}^{ij} &= \langle \hat{T}^\dagger k l j^\dagger i^\dagger \hat{T} \rangle = \frac{1}{2} T_{ab}^{ij} T_{kl}^{ab}, \\ \eta_j^i &= \langle \hat{T}^\dagger j i^\dagger \hat{T} \rangle = \frac{1}{2} T_{ab}^{ik} T_{jk}^{ab} \equiv \eta_{jk}^{ik}, \\ \eta_{aj}^{ib} &= \langle \hat{T}^\dagger j b^\dagger a i^\dagger \hat{T} \rangle = T_{ac}^{ik} T_{jk}^{bc}, \\ \eta_a^b &= \langle \hat{T}^\dagger b^\dagger a \hat{T} \rangle = \frac{1}{2} T_{ac}^{ij} T_{ij}^{bc} \equiv \frac{1}{2} \eta_{ai}^{ib}. \end{aligned} \quad (29)$$

The ${}_q\mathbf{U}$ transformation remains the same as in the approximate LPFD(0) theory,

$${}_qU_j^i = \delta_j^i + {}_qS_u \Delta_j^i, \quad (30)$$

$$\Delta_j^i = \eta_j^i, \quad (31)$$

$$({}_q\mathbf{U} {}_qP_u \mathbf{T})_{ab}^{ij} = \frac{1}{2} (1 - \tau_{ij}) ({}_q\mathbf{U} {}_qP_u)_k^i T_{ab}^{kj}, \quad (32)$$

$${}_qS_u = 1, \quad {}_qP_u = -\frac{q}{2}. \quad (33)$$

The purpose of the ${}_q\mathbf{W}$ transformation is to make the method equivalent to LPFD(-1) to $\mathcal{O}(T^4)$. At $\mathcal{O}(T^3)$, for example, it introduces the additional terms $(1/2)\mathbf{B} + \mathbf{C}$, which adds with the $(1/2)\mathbf{B}$ from LPFD(0) to total $\mathbf{B} + \mathbf{C}$. This is the correct total contribution of the terms \mathbf{B} and \mathbf{C} , since this is how they

enter VCCD,

$${}_qW_{kl}^{ij} = \delta_{kl}^{ij} + {}_qS_w \Omega_{kl}^{ij}, \quad (34)$$

$$\Omega_{kl}^{ij} = \eta_{kl}^{ij} - \frac{1}{2}(1 - \tau_{ij})(1 - \tau_{kl}) \delta_k^i \eta_l^j, \quad (35)$$

$$({}_qW {}_qP_w \mathbf{T})_{ab}^{ij} = \frac{1}{2} ({}_qW {}_qP_w)_{kl}^{ij} T_{ab}^{kl}. \quad (36)$$

The purpose of the ${}_qV$ transformation is to go beyond LPPFD(-1) and add, at $\mathcal{O}(T^3)$ for example, the terms $\mathbf{A} + \mathbf{D}$,

$${}_qV_{aj}^{ib} = \delta_{aj}^{ib} + {}_qS_v \Gamma_{aj}^{ib}, \quad (37)$$

$$\Gamma_{aj}^{ib} = 2(\delta_j^i \eta_a^b - \eta_{aj}^{ib}), \quad (38)$$

$$({}_qV {}_qP_v \mathbf{T})_{ab}^{ij} = \frac{1}{4}(1 - \tau_{ij})(1 - \tau_{ab}) ({}_qV {}_qP_v)_{ak}^{ic} T_{cb}^{kj}. \quad (39)$$

We choose, for now, to apply these transformations sequentially; the ${}_qU$ -transformed amplitudes are used in the place of the untransformed amplitudes in the ${}_qW$ transformation, and the ${}_qW {}_qU$ -transformed amplitudes are used in place of the untransformed amplitudes in the ${}_qV$ transformation,

$${}_qT_{ab}^{ij} = ({}_qV {}_qP_v {}_qW {}_qP_w {}_qU {}_qP_u \mathbf{T})_{ab}^{ij}. \quad (40)$$

Note that, for two electrons, both matrices ${}_qW$ and ${}_qV$ reduce to the identity, or equivalently, Ω and Γ are zero matrices. Therefore, these transformations have no effect in the two-electron limit and the theory remains exact.

The choices of ${}_qS_w$, ${}_qP_w$, ${}_qS_v$, and ${}_qP_v$ are therefore somewhat arbitrary. However, they are constrained by the requirement of agreement with VCCD at $\mathcal{O}(T^3)$ to obey the relationships given in Eq. (41),

$${}_qS_w \cdot {}_qP_w = -{}_qS_v \cdot {}_qP_v = +\frac{q}{2}. \quad (41)$$

We therefore make the choices ${}_qP_w = +1$, ${}_qS_w = +\frac{q}{2}$ and ${}_qP_v = +1$, ${}_qS_v = -\frac{q}{2}$. This is the most obvious choice, based on two considerations. First, negative and non-integer powers of each of these matrices are ruled out on the grounds of positivity. Second, since neither matrix needs to be powered, the increase in computational difficulty from LPPFD to AVCCD is minimized. This is especially important for ${}_qV$, which can get quite large.

We now discuss the correspondences between VCCD and AVCCD. By construction, AVCCD is equivalent to VCCD at $\mathcal{O}(T^3)$. This is a significant improvement over all LPPFD theories, including LPPFD(-1).

At $\mathcal{O}(T^4)$ (where the relevant VCCD term is $(1/4)\langle(\hat{T}_2^\dagger)^2 \hat{H} \hat{T}_2^2\rangle_L$), AVCCD still omits some terms from VCCD. However, although AVCCD and VCCD do not agree to this order in general, the combination of terms in AVCCD is known to be reasonable since there is exact agreement to $\mathcal{O}(T^4)$ in the limiting case of two virtual spin-orbitals.

Although it is, in principle, also possible to add the terms omitted by AVCCD at $\mathcal{O}(T^4)$ without breaking any of the important properties of the method, doing so would violate the criterion that the method is computable in $\mathcal{O}(o^2v^4)$ time, since one of the omitted terms is precisely that term that makes $(1/4)\langle(\hat{T}_2^\dagger)^2 \hat{H} \hat{T}_2^2\rangle_L$ scale as $\mathcal{O}(v^6)$ in computational complexity.

III. NUMERICAL RESULTS

The BLPFD, OLPFD, BAVCCD, and OAVCCD have been implemented in the closed-shell CCSD program³³ of the MOLPRO (Ref. 34) quantum chemistry software package. We now present some benchmark results comparing these methods with CCSD. For reasonably small systems and basis sets, we are able to compare directly against VCCSD and FCI. When VCCSD and FCI calculations are prohibitively expensive, we instead compare against other methods, such as CCSD(T) (Ref. 35) or Davidson-corrected internally contracted multireference CI,³⁶ hereafter abbreviated to MRCI+Q.

We first examine some calculated atomic correlation energies, given in Table I. In the case of C, O, Si, and S, the reference wavefunction is one of the closed-shell single determinants contributing to the 1S and 1D states. The results for the OLPFD method are deceptively good. For example, it predicts a correlation energy for the carbon atom very close to the CCSD(T) value. However, it sometimes overshoots, as in the case of the oxygen atom. Where the single-determinant reference function is a good approximation (Ne and Ar), LPPFD is close to CCSD. The OAVCCD, however, gives correlation energies that are very close to those from CCSD in all cases, and this agreement persists on inclusion of the perturbative triples correction. We have also included the energies calculated by the BLPFD and BAVCCD methods, which differ by at most $0.1 mE_h$ from the results with variational orbital optimization.

We also examine the predicted equilibrium bond lengths and spectroscopic constants of some closed-shell diatomic molecules in Table II. We show results for HF and CO in a

TABLE I. Comparison of the valence-shell correlation energies (in millihartrees, mE_h) of some singlet-state atoms. All calculations were performed with the cc-pV5Z basis.

Atom	CCSD	BLPFD	OLPFD	BAVCCD	OAVCCD	CCSD(T)	OAVCCD(T)
C	-126.5	-132.3	-132.4	-125.2	-125.2	-132.2	-129.7
O	-217.7	-233.9	-234.0	-216.0	-216.0	-229.1	-225.0
Ne	-305.5	-305.3	-305.4	-305.2	-305.3	-311.6	-311.5
Si	-105.0	-111.0	-111.0	-103.4	-103.4	-110.0	-107.5
S	-184.4	-198.6	-198.7	-182.7	-182.8	-194.8	-191.6
Ar	-255.6	-258.0	-258.0	-255.5	-255.5	-264.9	-264.7

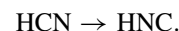
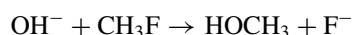
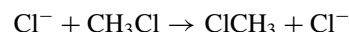
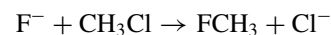
TABLE II. Comparison of equilibrium bond lengths (\AA) and spectroscopic constants (cm^{-1}) for some diatomic molecules.

Method	R_e	ω_e	$\omega_e x_e$
HF with the 6-31G basis			
CCSD	0.9462	3780.7	95.30
OLPFD	0.9458	3783.1	96.36
OAVCCD	0.9460	3785.0	95.11
VCCSD	0.9464	3778.3	95.45
FCI	0.9475	3757.3	96.45
HF with the cc-pV5Z basis			
CCSD	0.9154	4172.9	89.50
BLPFD	0.9151	4174.4	90.47
OLPFD	0.9152	4172.6	90.69
BAVCCD	0.9149	4181.0	89.26
OAVCCD	0.9151	4178.9	89.40
CCSD[T]	0.9185	4122.7	91.06
CCSD(T)	0.9182	4128.3	90.82
BLPFD(T)	0.9181	4125.4	92.07
OLPFD(T)	0.9182	4124.6	92.24
BAVCCD(T)	0.9179	4133.7	90.56
OAVCCD(T)	0.9180	4132.6	90.64
Empirical	0.9168	4138.3	89.88
CO with the STO-3G basis			
CCSD	1.1822	2179.4	9.64
OLPFD	1.1807	2191.8	9.80
OAVCCD	1.1816	2182.2	9.88
VCCSD	1.1820	2180.4	9.75
FCI	1.1925	2068.7	11.62
CO with the cc-pV5Z basis			
CCSD	1.1236	2235.5	12.08
BLPFD	1.1228	2246.7	11.83
OLPFD	1.1233	2240.8	11.98
BAVCCD	1.1223	2251.0	11.79
OAVCCD	1.1227	2244.9	11.96
CCSD[T]	1.1327	2136.7	14.19
CCSD(T)	1.1307	2164.5	13.11
BLPFD(T)	1.1309	2166.0	12.99
OLPFD(T)	1.1312	2161.2	13.15
BAVCCD(T)	1.1299	2174.4	12.89
OAVCCD(T)	1.1302	2169.5	13.07
Empirical	1.1283	2169.8	13.29
HCl with the cc-pV5Z basis			
CCSD	1.2735	3014.1	49.58
BLPFD	1.2744	3002.4	50.47
OLPFD	1.2745	3001.8	50.52
BAVCCD	1.2733	3016.6	49.44
OAVCCD	1.2734	3016.1	49.48
CCSD[T]	1.2759	2989.7	50.53
CCSD(T)	1.2758	2990.5	50.52
BLPFD(T)	1.2770	2975.4	51.65
OLPFD(T)	1.2770	2975.2	51.69
BAVCCD(T)	1.2757	2992.7	50.31
OAVCCD(T)	1.2757	2992.6	50.34
Empirical	1.2746	2990.9	52.82

minimal basis set, in order to allow comparison with FCI and VCCSD results. Results for HF, CO, and HCl in the cc-pV5Z basis³⁷ are also included. In this basis, we compare against both CCSD(T) and the empirically known values. We have additionally included results obtained by adding the standard (T) correction³⁵ evaluated with the converged doubles ampli-

tudes from the LPFD and AVCCD calculations using both variational and Brueckner orbitals. The conclusions are similar to those from the atoms; both OLPFD and OAVCCD produce similar results to CCSD, which is itself close to VCCSD. In the case of CO, the harmonic frequency from OLPFD is somewhat too high, but OAVCCD agrees very well with VCCSD.

Quantum chemical methods may also be benchmarked by the calculation and comparison of barrier heights, for which Truhlar *et al.*^{38,39} have given several representative test cases. Since OLPFD and OAVCCD have been implemented only in a closed-shell program, the number of test cases is limited, but we are able to calculate barrier heights for the following reactions:



In Table III, barrier heights calculated in the 3-21G basis are given. The CCSDT results obtained using the MRCC program⁴⁰ of Kállay demonstrate that the error due to the perturbative treatment of triple excitations in CCSD(T) is small relative to the error of CCSD. Barrier heights in the more representative cc-pVTZ basis³⁷ are given in Table IV, where the double-excitation methods can be compared with CCSD(T).

The conclusions from this test are no different; OAVCCD performs very much like CCSD. It is possibly the case that static correlation, for which OAVCCD may be a superior description, is weak for these barriers, which leads to the CCSD-like behaviour. Overall, we conclude that OAVCCD is very slightly inferior to CCSD for the treatment of dynamic correlation. However, we now turn to the examination of some systems where static correlation effects are strong.

We first examine the potential energy curve of BeO, given in Fig. 2, predicted by each of the methods with the STO-3G

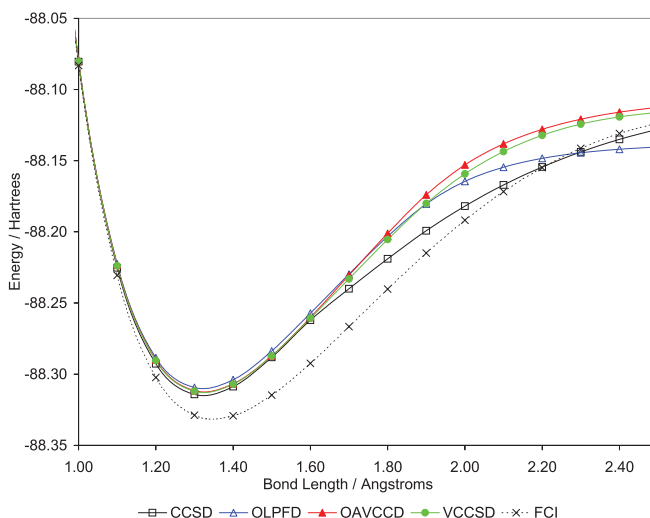


FIG. 2. Potential energy curve of BeO with the STO-3G basis set and all electrons correlated.

TABLE III. Calculated barrier heights (in millihartrees, mE_h) with the 3-21G basis.

Reaction	OLPFD	OAVCCD	CISD+Q	CEPA(0)	CCSD	CCSD(T)	CCSDT
$F^- + CH_3Cl \rightarrow FCH_3 + Cl^-$	-36.937	-36.541	-37.303	-39.008	-36.763	-38.454	-38.474
$Cl^- + CH_3Cl \rightarrow ClCH_3 + Cl^-$	2.045	2.278	1.615	0.533	1.978	0.747	0.599
$OH^- + CH_3F \rightarrow HOCH_3 + F^-$	-43.346	-42.802	-41.282	-46.328	-43.263	-46.137	-46.242
$HCN \rightarrow HNC$	110.415	111.179	109.096	105.556	110.703	109.182	108.888

basis. This system exemplifies the difference between CCSD and VCCSD. The shape of the CCSD curve is particularly poor at long bond lengths, crossing the FCI curve between 2.2 Å and 2.3 Å. The VCCSD performs significantly better, remaining fairly parallel to the FCI curve throughout the region investigated. Unlike CCSD, it does not cross the FCI curve, because the VCCSD energy functional is an upper bound on the ground-state energy. It is hard to say whether OLPFD performs better or worse than CCSD in this system; it remains closer to the VCCSD curve for a while but levels off quickly and ends up below the CCSD curve and significantly below the FCI. On the other hand, OAVCCD remains particularly close to the VCCSD curve even at long bond lengths and is thus qualitatively correct throughout. Although OAVCCD is not an upper bound on the exact ground-state energy, it still makes use of a variationally minimized energy functional and clearly inherits some stability from its parent theory, VCCSD.

The situation for BeO appears to remain much the same in the larger aug-cc-pVTZ basis,⁴¹ illustrated in Fig. 3, except now VCCSD is unavailable for comparison. Given the evidence from the smaller STO-3G basis and the lack of changes to the relative positioning of the OAVCCD and CCSD curves, OAVCCD may still mimic the VCCSD behaviour better than CCSD here, despite the improved CCSD performance.

The potential energy curve of C_2 with the cc-pVDZ basis³⁷ (d functions omitted), shown in Fig. 4, provides another interesting example. In this calculation, only the four electrons occupying the $1\pi_u$ orbitals in the reference are correlated, in order to explore the effect of the $1\pi_u - 3\sigma_g$ near degeneracy. None of the methods describes this system particularly well in comparison to FCI; the predicted equilibrium bond lengths are in each case roughly 0.1 Å too short. This can be attributed to strong dynamic correlation in the equilibrium region that is not captured by any of the methods due to the truncation of the cluster operator to the singles and doubles level. This is evidence that a variational approach (such as VCCSD) is no more useful than a non-variational approach (such as CCSD) in such situations. However, as the bond is stretched, static correlation effects also become strong and a clear difference between CCSD and VCCSD emerges. The

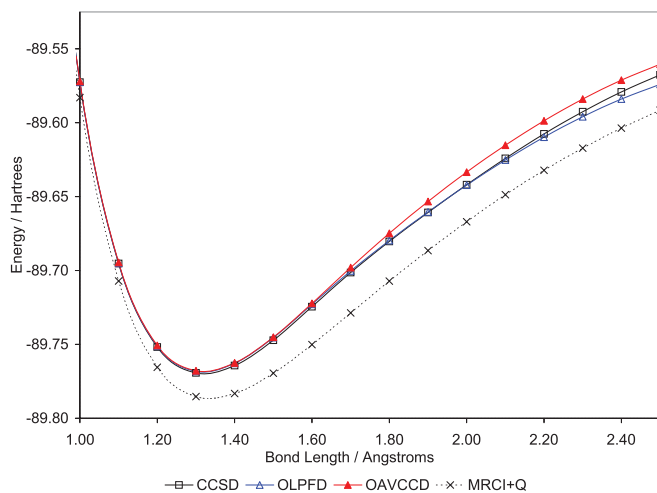
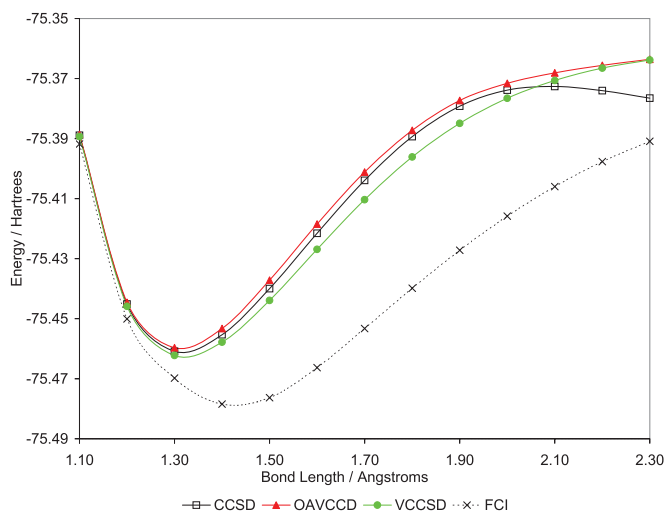


FIG. 3. Potential energy curve of BeO with the aug-cc-pVTZ basis set and all electrons correlated.

FIG. 4. Potential energy curve of C_2 with the cc-pVDZ basis set (d functions omitted). Only the four electrons occupying $1\pi_u$ in the reference wavefunction were correlated.TABLE IV. Calculated barrier heights (in millihartrees, mE_h) with the cc-pVTZ basis.

Reaction	OLPFD	OAVCCD	CISD+Q	CEPA(0)	CCSD	CCSD(T)
$F^- + CH_3Cl \rightarrow FCH_3 + Cl^-$	1.483	2.037	1.832	-2.180	1.682	-0.369
$Cl^- + CH_3Cl \rightarrow ClCH_3 + Cl^-$	22.079	22.898	23.054	18.193	22.517	19.807
$OH^- + CH_3F \rightarrow HOCH_3 + F^-$	-15.568	-14.683	-3.881	-21.508	-15.363	-20.276
$HCN \rightarrow HNC$	76.309	77.046	76.239	71.506	76.752	76.076

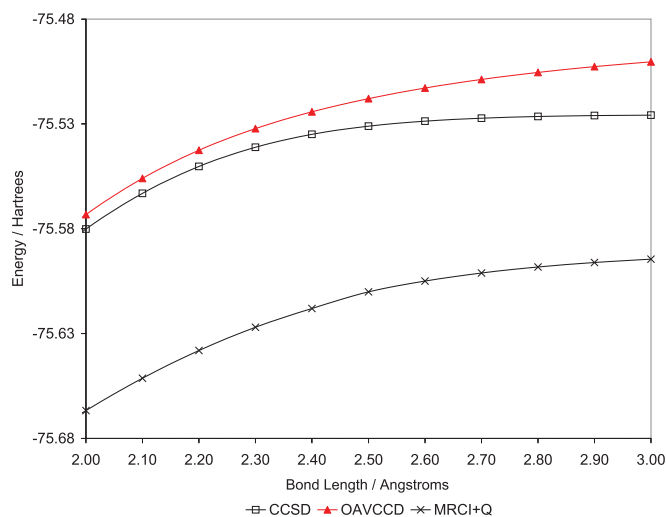


FIG. 5. Long bond length behaviour of the potential energy curve of C_2 with the aug-cc-pVTZ basis set and all electrons correlated.

CCSD predicts a maximum in the potential energy curve at ~ 2.0 Å. This behaviour is unphysical and qualitatively incorrect in comparison to the FCI. The VCCSD does not show this failure, and this can again be attributed to its variational upper bound property. The OAVCCD curve does not follow the VCCSD curve as precisely as in the case of BeO, but it still remains qualitatively correct and does not fail in the same catastrophic fashion as CCSD.

Finally, we examine C_2 at long bond lengths in the aug-cc-pVTZ basis, with all electrons correlated, illustrated in Fig. 5. The predicted unphysical maximum of CCSD has now vanished, but the methods predict the bonds to be fully broken (defined as the curve levelling off to close to zero gradient) at different bond lengths. The CCSD predicts a flat potential energy curve from around 2.8 Å onwards, whereas the OAVCCD curve remains highly parallel with the MRCI+Q curve, predicting a non-zero gradient over the 2.8 Å–3.0 Å region. We do not illustrate OLPFD on either of the C_2 plots since the method becomes qualitatively incorrect at bond lengths too short for any meaningful comparison to be made.

IV. CONCLUDING REMARKS

We have shown that it is possible to construct an approximation to VCCSD with computational demands that scale as $\mathcal{O}(\sigma^2 v^4)$, and which is exact for isolated two-electron subsystems, rigorously extensive, and possesses several other aesthetic qualities. The approximation agrees with VCCD to $\mathcal{O}(T^3)$ but replaces difficult diagrams at $\mathcal{O}(T^4)$ and beyond by an effective closed-form correction.

We have also demonstrated that the predictions of the OAVCCD method are extremely similar to CCSD when static correlation effects are weak. However, in cases of strong static correlation, OAVCCD is robust and close to VCCSD and does not suffer the same poor performance as CCSD. In cases dom-

inated by dynamic correlation, extending the cluster operator to include triple or quadruple excitations remains more important than the variational character of the method.

ACKNOWLEDGMENTS

We are indebted to C. Kollmar for discussions regarding his orbital optimization code. This work was performed using the computational facilities of the Advanced Research Computing @ Cardiff (ARCCA) Division, Cardiff University.

- ¹D. R. Hartree, *Proc. Cambridge Philos. Soc.* **24**, 89 (1928).
- ²V. Fock, *Z. Phys.* **61**, 126 (1930).
- ³C. Møller and M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- ⁴N. C. Handy, P. J. Knowles, and K. Somasundram, *Theor. Chem. Acc.* **68**, 87 (1985).
- ⁵O. Christiansen, J. Olsen, P. Jørgensen, H. Koch, and P.-Å. Malmqvist, *Chem. Phys. Lett.* **261**, 369 (1996).
- ⁶J. Olsen, P. Jørgensen, T. Helgaker, and O. Christiansen, *J. Chem. Phys.* **112**, 9736 (2000).
- ⁷P.-O. Löwdin, *Phys. Rev.* **97**, 1474 (1955).
- ⁸F. Coester, *Nucl. Phys.* **7**, 421 (1958).
- ⁹F. Coester and H. Kümmel, *Nucl. Phys.* **17**, 477 (1960).
- ¹⁰J. Čížek, *J. Chem. Phys.* **45**, 4256 (1966).
- ¹¹R. J. Bartlett and J. Noga, *Chem. Phys. Lett.* **150**, 29 (1988).
- ¹²R. Ahlrichs, P. Scharf, and C. Ehrhardt, *J. Chem. Phys.* **82**, 890 (1985).
- ¹³H. P. Kelly, *Phys. Rev.* **134**, A1450 (1964).
- ¹⁴W. Meyer, *J. Chem. Phys.* **58**, 1017 (1973).
- ¹⁵P. Pulay and S. Sæbø, *Chem. Phys. Lett.* **117**, 37 (1985).
- ¹⁶R. Ahlrichs, *Comput. Phys. Commun.* **17**, 31 (1979).
- ¹⁷S. Koch and W. Kutzelnigg, *Theor. Chim. Acta* **59**, 387 (1981).
- ¹⁸S. R. Langhoff and E. R. Davidson, *Int. J. Quantum Chem.* **8**, 61 (1974).
- ¹⁹M. Nooijen and R. J. LeRoy, *J. Mol. Struct.: THEOCHEM* **768**, 25 (2006).
- ²⁰T. Van Voorhis and M. Head-Gordon, *J. Chem. Phys.* **113**, 8873 (2000).
- ²¹P.-D. Fan, K. Kowalski, and P. Piecuch, *Mol. Phys.* **103**, 2191 (2005).
- ²²B. Cooper and P. J. Knowles, *J. Chem. Phys.* **133**, 234102 (2010).
- ²³P. G. Szalay, M. Nooijen, and R. J. Bartlett, *J. Chem. Phys.* **103**, 281 (1995).
- ²⁴E. Merzbacher, *Quantum Mechanics*, 2nd ed. (Wiley, New York, 1970).
- ²⁵I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, Cambridge, England, 2009).
- ²⁶W. Kutzelnigg, *Theo. Chim. Acta* **80**, 349 (1991).
- ²⁷J. Arponen, *Ann. Phys. (N.Y.)* **151**, 311 (1983).
- ²⁸P. J. Knowles and B. Cooper, *J. Chem. Phys.* **133**, 224106 (2010).
- ²⁹L. M. J. Huntington and M. Nooijen, *J. Chem. Phys.* **133**, 184109 (2010).
- ³⁰N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks, *Chem. Phys. Lett.* **164**, 185 (1989).
- ³¹C. Kollmar and A. Heßelmann, *Theor. Chem. Acc.* **127**, 311 (2009).
- ³²L. Adamowicz, W. D. Laidig, and R. J. Bartlett, *Int. J. Quantum Chem., Symp.* **18**, 245 (1984).
- ³³C. Hampel, K. Peterson, and H.-J. Werner, *Chem. Phys. Lett.* **190**, 1 (1992).
- ³⁴H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz *et al.*, “MOLPRO, version 2010.2, a package of *ab initio* programs” (2011), see <http://www.molpro.net>.
- ³⁵K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- ³⁶H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* **89**, 5803 (1988).
- ³⁷T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- ³⁸J. Zheng, Y. Zhao, and D. G. Truhlar, *J. Chem. Theory Comput.* **3**, 569 (2007).
- ³⁹J. Zheng, Y. Zhao, and D. G. Truhlar, *J. Chem. Theory Comput.* **5**, 808 (2009).
- ⁴⁰M. Kállay and P. R. Surján, *J. Chem. Phys.* **115**, 2945 (2001); see also <http://www.mrcc.hu>.
- ⁴¹R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).