Amplitude determinant coupled cluster with pairwise doubles

Luning Zhao1 and Eric Neuscamman1,2,*

1Department of Chemistry, University of California, Berkeley, California 94720, USA
2Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

(Dated: March 26, 2018)

Recently developed pair coupled cluster doubles (pCCD) theory successfully reproduces doubly occupied configuration interaction (DOCI) with mean field cost. However, the projective nature of pCCD makes the method non-variational and thus hard to improve systematically. As a variational alternative, we explore the idea of coupled-cluster-like expansions based on amplitude determinants and develop a specific theory similar to pCCD based on determinants of pairwise doubles. The new ansatz admits a variational treatment through Monte Carlo methods while remaining size-consistent and, crucially, polynomial cost. In the dissociations of LiH, HF, H2O and N2, the method performs very similarly to pCCD and DOCI, suggesting that coupled-cluster-like ansatzes and variational evaluation may not be mutually exclusive.

I. INTRODUCTION

The electronic correlation energy can be divided into two parts1. Dynamic correlation, small changes to mean field theory, can be described effectively through many-body perturbation theory2. Static or non-dynamic correlation becomes important when the ground state contains degenerate or near-degenerate configurations, i.e. where more than one determinant is needed for a qualitative description of the system. Such situations arise in bond breaking1, large π-conjugated molecules where the HOMO-LUMO gap is small3, transition metal complexes4 and superconductivity5. For such cases, a single Slater determinant gives qualitatively incorrect results.

The concept of seniority has a long history in nuclear physics6. Seniority number is the measurement of unpaired electrons in a determinant. It has been showed before that strong correlation is very often concentrated within the low seniority regions of Hilbert space7. Among them, the seniority zero wave function is the most studied due to the fact that it is both size consistent and capable of capturing a large amount of strong correlation.

The most elaborate seniority zero wave function is doubly occupied configuration interaction (DOCI)8, which is a full configuration interaction (FCI) wave function limited to seniority zero space. DOCI was thoroughly studied in the pioneering days of quantum chemistry8–10 and has many desirable features, being size-consistent, variational and effective for describing strong correlation. However, DOCI’s factorial cost scaling severely limits its applicability.

Coupled Cluster (CC) with single and double excitations (CCSD)11 and CCSD with perturbative triple excitations [CCSD(T)]12 provide a very powerful way to describe dynamic correlation for a polynomial scaling cost. Between its rigorous size consistency, its systematic improviability, and its exceptional accuracy even at the CCSD(T) level, CC has become one of the most trusted and even routine methods in quantum chemistry for weakly correlated systems that are not too large. Unfortunately, the main drawback of traditional CC is that the ansatz is optimized in a projective manner to achieve polynomial scaling, which makes the energy non-variational and leads to qualitative failure in strongly correlated regimes13–15. Examples include breaking of multiple bonds simultaneously, and the Hubbard model at large U, in which RHF-based CCSD or even CCSD(T) “overcorrelates” and the converged energy is well below the FCI energy16. While many approaches have been taken to resolve this issue, including multi-reference CC17, combinatorially scaling variational CC18, single-pair couple cluster19 and CC valence bond20, there remain no CC methods that can capture both static and dynamic correlation in large molecules at an affordable cost.

Recently, the antisymmetric-product of one-reference-orbital geminals (AP1roG)21 was introduced by Ayers and coworkers in an attempt to deal with static correlation at polynomial cost. Impressively, they showed that AP1roG gives results almost equivalent to those of DOCI. Later, Scuseria and coworkers recognized that AP1roG is a simplified version of coupled cluster doubles (CCD), in which electrons are paired and only pair excitations are allowed, which they termed pair-CCD (pCCD)13. It is quite remarkable that by eliminating the vast bulk of the cluster operator, the pCCD ansatz accurately reproduces the factorially complex DOCI. It is also important: since DOCI provides a powerful description of strong correlation in a wide variety of systems, so do AP1roG and pCCD. However, like most other CC methods, both pCCD and AP1roG are non-variational.

One may raise the question that since pCCD is already so close to DOCI, what is the point of achieving variationality? One should note that although pCCD does not have enough degrees of freedom to break variationality, once one breaks electron pairs and adds high seniority determinants to the ansatz, aiming to recover the dynamic correlations missing in pCCD, this theory

*Electronic mail: eneuscamman@berkeley.edu

arXiv:1608.05195v1 [physics.chem-ph] 18 Aug 2016
A. CC Incompatibility with VMC

In variational Monte Carlo, the energy expression is

\[ E = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \sum_{\vec{n}} |\langle \vec{i} | \Psi \rangle|^2 \frac{\langle \vec{i} | H | \Psi \rangle}{\langle \vec{i} | \Psi \rangle} \]

(1)

which may be approximated by sampling a set of occupations \( \xi \) from the wave function’s probability distribution \( |\langle \vec{i} | \Psi \rangle|^2 / \langle \Psi | \Psi \rangle \). The energy is then estimated as an average of local energies.

\[ E = \frac{1}{n_s} \sum_{\vec{n} \in \xi} \langle \vec{i} | H | \Psi \rangle = \frac{1}{n_s} \sum_{\vec{n} \in \xi} E_L (\vec{i}) \]

(2)

with \( n_s \) being the number of Monte Carlo samples in \( \xi \).

For our purposes, it is convenient to write the \textit{ab initio} Hamiltonian as

\[ H = \sum_{pq} t_{pq} \left( a^\dagger_p a_q + a^\dagger_q a_p \right) + \sum_{pqrs} (pq|rs) \left( \frac{1}{2} a^\dagger_p a^\dagger_q a_r a_s + \frac{1}{2} a^\dagger_p a_q a_r a_s + a^\dagger_p a_q a_r a^\dagger_s + a^\dagger_p a^\dagger_q a_s a_r \right) \]

(3)

where \((pq|rs)\) are the usual two-electron coulomb integrals in (11|22) order, \( a^\dagger_p \) (\( a_p^\dagger \)) and \( a_p \) (\( a^\dagger_p \)) are creation and destruction operators of an \( \alpha(\beta) \) electron in the \( p \)th spatial orbital, and \( t_{pq} \) are modified one-electron integrals,

\[ t_{pq} = h_{pq} - \frac{1}{2} \sum_r (pr|rq) \]

(4)

where \( h_{pq} \) are the standard one-electron integrals.

From the energy and Hamiltonian expressions, we see that if an ansatz can evaluate \( \langle \vec{i} | \Psi \rangle \) efficiently, this is sufficient for use with VMC in Fock space. For the pCCD ansatz with RHF as reference, its cluster amplitudes can be written in matrix form

\[ T = \begin{pmatrix}
\langle i | \pi \rangle & \langle i | \pi \rangle & \langle i | \pi \rangle \\
\langle j | \pi \rangle & \langle j | \pi \rangle & \langle j | \pi \rangle \\
\langle k | \pi \rangle & \langle k | \pi \rangle & \langle k | \pi \rangle \\
\vdots & \vdots & \vdots \\
\end{pmatrix} \]

(5)

in which index \( i, j, k \) represent occupied orbitals in the reference and \( a, b, c \) represent virtual orbitals in the reference.
The amplitude expression for pCCD ansatz can be written as

\[
\langle \vec{n} \rvert \Psi_{pCCD} \rangle = \langle \vec{n} \rvert \exp \left( \hat{T} \right) \rvert RHF \rangle = \langle \vec{n} \rvert RHF \rangle + \langle \vec{n} \rvert \hat{T} \rvert RHF \rangle + \frac{1}{2} \langle \vec{n} \rvert \hat{T}^2 \rvert RHF \rangle + \ldots \tag{6}
\]

in which \( \hat{T} \) is the cluster operator

\[
\hat{T} = \sum_{ia} t^{\sigma}_{ia} a_i^\dagger a_i a_i^\dagger \alpha_i \tag{7}
\]

Suppose the occupation number vector is a quadruply excited configuration, \( \vec{n}^{ab}_{ij} \), relative to the reference determinant. Then only the second order cluster operator term will not vanish. Therefore,

\[
\langle \vec{n}^{ab}_{ij} \rvert \Psi_{pCCD} \rangle = \frac{1}{2} \times 2 \times \left( t^{\sigma}_{ia} t^{\sigma}_{ja} + t^{\sigma}_{ia} t^{\sigma}_{jb} + t^{\sigma}_{ia} t^{\sigma}_{ja} + t^{\sigma}_{ja} t^{\sigma}_{jb} \right) \\
= \frac{1}{2} \times 2 \times \text{Perm} \left( t^{\sigma}_{ia} t^{\sigma}_{ja} t^{\sigma}_{ib} t^{\sigma}_{jc} \right) \\
= \text{Perm} \left[ T \left( \vec{n}^{ab}_{ij} \right) \right] \tag{8}
\]

in which “\text{Perm}” represents the permanent of a matrix and the factor 2 comes from the number of ways of permuting \( i, j, a, b \) indices, which cancels the 1/2 from the exponentials Taylor series. This result generalizes, with the amplitude of pCCD for a given occupation number vector given by a permanent of an \( n_{ex}/2 \) by \( n_{ex}/2 \) part of the cluster amplitude matrix, with \( n_{ex} \) the excitation level in the occupation number vector \[11\]. For \( N \)th level excitations, there are \( N! \) ways to permute the indices, therefore the exponential’s \( 1/N! \) is canceled and no constants appear before the permanent.

Thus, the difficulty in evaluating a pCCD occupation number coefficient is equivalent to the evaluation of a permanent. However, there are no known polynomial cost methods for permanent evaluation, and so pCCD, like CC in general, is incompatible with VMC.

B. Determinant Amplitude

We propose a different, but related ansatz, in which coefficients are defined to be determinants rather than permanents. Using the \( n_{ex}/2 \) by \( n_{ex}/2 \) matrix

\[
\hat{T} = \begin{pmatrix}
    t^{\sigma}_{ia} & t^{\sigma}_{ja} & t^{\sigma}_{ia} & \ldots \\
    -t^{\sigma}_{ia} & t^{\sigma}_{ja} & t^{\sigma}_{ia} & \ldots \\
    -t^{\sigma}_{ia} & -t^{\sigma}_{ia} & t^{\sigma}_{ja} & \ldots \\
    \ldots & \ldots & \ldots & \ldots
\end{pmatrix}
\]

we define the wave function coefficient to be

\[
\langle \vec{n} \rvert \Psi \rangle = \det \left[ \hat{T} \left( \vec{n} \right) \right] . \tag{10}
\]

These determinants can be evaluated through LU decomposition with \( \mathcal{O} \left( n_{ex}^3 \right) \) cost, making this pADCCD ansatz compatible with VMC.

We should note that our choice of determinant instead of permanent forms a different approximation than pCCD, and since pCCD is not exact, using a determinant is not necessarily better or worse. Moreover, as we will prove later, size consistency still holds, and so one of CC’s most important properties is maintained. And since we make the lower triangle negative, the coefficients obtained from pADCCD match those of pCCD through quadruple excitations. Unlike pCCD, pADCCD cannot be written in the compact exp(\( \hat{T} \)) form, but with size consistency nonetheless maintained, this seems a minor inconvenience.

C. Energy Expression

We will derive the energy expression for pADCCD in this subsection. Before considering our particular wave function, notice that the subset of Hamiltonian terms that contain only number and hole operators (i.e. \( a_i^\dagger a_i \) and \( a_i^\dagger a_i \)) will add a wave-function-independent contribution to the local energy,

\[
E_0 \left( \vec{n} \right) = \sum_i t_{ii} + \sum_i t_{ii} + \sum_{ij} \left( \langle ii \rangle \langle jj \rangle \right) \\
+ \frac{1}{2} \left[ \sum_{ij} \langle ii \rangle \langle jj \rangle + \sum_{ia} \langle ia \rangle \langle ai \rangle \right]
\]

The remaining one-electron terms in the Hamiltonian of the type \( a_i^\dagger a_i \) will not contribute to the local energy as it breaks seniority symmetry and pADCCD is strictly seniority zero. Likewise, the only two-electron terms in the Hamiltonian that will give a non-vanishing contribution to the local energy are the terms of the type \( a_i^\dagger a_i a_a^\dagger a_a^\dagger \). Therefore, the expression for the local energy is,

\[
E \left( \vec{n} \right) = E_0 \left( \vec{n} \right) + \sum_{ia} \langle ia \rvert \tilde{\vec{n}} \rvert \Psi \rangle \frac{\langle \tilde{\vec{n}}^\sigma_{ia} a_i a_a^\dagger a_a^\dagger \Psi \rangle}{\langle \tilde{\vec{n}} \rvert \Psi \rangle} \\
= E_0 \left( \vec{n} \right) + \sum_{ia} \langle ia \rvert \tilde{\vec{n}} \rvert \Psi \rangle \frac{\langle \tilde{\vec{n}}^\sigma_{ia} \Psi \rangle}{\langle \tilde{\vec{n}} \rvert \Psi \rangle} \\
= E_0 \left( \vec{n} \right) + \sum_{ia} \langle ia \rvert \tilde{\vec{n}} \rvert \Psi \rangle \frac{\det \left[ \hat{T} \left( \tilde{\vec{n}}^\sigma_{ia} \right) \right]}{\det \left[ \hat{T} \left( \vec{n} \right) \right]} \tag{12}
\]

in which \( \langle \tilde{\vec{n}}^\sigma_{ia} \rvert \Psi \rangle \) is obtained by exciting two electrons (\( \alpha \) and \( \beta \)) from the \( ith \) occupied orbital to the \( ath \) virtual orbital in \( \vec{n} \).
D. Proof of Size-Consistency

In this subsection, we will prove that our pADCCD ansatz is strictly size consistent. The definition of size-consistency states that the energy calculated with two non-interacting systems A and B together as a “super system” should be equal to the sum of the energies of systems A and B calculated separately. To prove this, consider two non-interacting systems A and B. Then the energy expression becomes

\[ E = \langle \vec{n}_{AB}|\Psi_{AB}\rangle \]

and it follows that,

\[ \langle \Psi_{AB}|\Psi_{AB}\rangle = \sum_{\vec{n}_{AB}} \langle \vec{n}_{AB}|\vec{n}_{AB}\rangle \langle \vec{n}_{AB}|\Psi_{AB}\rangle \]

\[ = \sum_{\vec{n}_{AB}} \langle \vec{n}_{AB}|\vec{n}_{AB}\rangle^2 \langle \vec{n}_{AB}|\Psi_{AB}\rangle^2 \]

Thus the energy expression becomes

\[ E_{AB} = \sum_{\vec{n}_{AB}} \frac{\langle \vec{n}_{AB}|\vec{n}_{AB}\rangle^2 \langle \vec{n}_{AB}|H_{A} + H_{B}|\Psi_{AB}\rangle}{\langle \vec{n}_{AB}|\Psi_{AB}\rangle} \]

\[ = \sum_{\vec{n}_{AB}} \frac{\langle \vec{n}_{AB}|H_{A}|\Psi_{A}\rangle \langle \vec{n}_{AB}|H_{B}|\Psi_{B}\rangle}{\langle \vec{n}_{AB}|\Psi_{AB}\rangle} \]

\[ + \sum_{\vec{n}_{AB}} \langle \vec{n}_{AB}|H_{A}|\Psi_{A}\rangle \langle \vec{n}_{AB}|H_{B}|\Psi_{B}\rangle \]

\[ = \sum_{\vec{n}_{A}} \langle \vec{n}_{A}|\vec{n}_{A}\rangle^2 \left( \frac{\langle \vec{n}_{A}|H_{A}|\Psi_{A}\rangle \langle \vec{n}_{A}|H_{B}|\Psi_{B}\rangle}{\langle \vec{n}_{A}|\Psi_{A}\rangle} \right) + \sum_{\vec{n}_{B}} \langle \vec{n}_{B}|H_{B}|\Psi_{B}\rangle \langle \vec{n}_{B}|\Psi_{B}\rangle \]

\[ = \sum_{\vec{n}_{A}} \langle \vec{n}_{A}|\vec{n}_{A}\rangle^2 \left( \frac{\langle \vec{n}_{A}|H_{A}|\Psi_{A}\rangle \langle \vec{n}_{A}|H_{B}|\Psi_{B}\rangle}{\langle \vec{n}_{A}|\Psi_{A}\rangle} \right) + \sum_{\vec{n}_{B}} \langle \vec{n}_{B}|H_{B}|\Psi_{B}\rangle \langle \vec{n}_{B}|\Psi_{B}\rangle \]

\[ = E_{A} + E_{B} \]

and so pADCCD is size consistent.

E. Derivative Ratios

Before discussing the variational minimization of the energy, we first lay the groundwork by developing efficient evaluations of terms that we call derivative ratios, which will make the optimization relatively simple to describe. Defining derivative notation \( \langle \Psi^x | H | \Psi^y \rangle \) with respect to the xth and yth wave function parameters \( \mu_x \) and \( \mu_y \), respectively, and \( |\Psi^0\rangle \equiv |\Psi\rangle \). After solving this eigenvalue problem for \( c \), one updates the parameters by

\[ \mu_x \leftarrow \mu_x + c_x / c_0 \]
The Hamiltonian and overlap matrices are built by Monte Carlo sampling

\[
\sum_{\vec{n} \in \xi} \sum_{y \in \{0,1,\ldots\}} \langle \Psi^y | \vec{n} \rangle \langle \vec{n} | H | \Psi^y \rangle \gamma_{y} \nonumber
\]

\[
= \lambda \sum_{\vec{n} \in \xi} \sum_{y \in \{0,1,\ldots\}} \langle \Psi^y | \vec{n} \rangle \langle \vec{n} | \Psi^y \rangle c_{y} \nonumber
\]

where we work at \( \vec{\kappa} = 0 \) by transforming the basis in \( \vec{\kappa} \) via a basis rotation. The one- and two-electron integrals.

Inspecting the above Monte Carlo approximations to the Hamiltonian and overlap matrices in the first derivative subspace makes clear that only derivative ratios \( D_n (\mu_x) \) and \( G_n (\mu_x) \) from Eqs. \ref{eq:17} and \ref{eq:19} are needed. Therefore, the LM can be applied efficiently to optimize the amplitudes by evaluating these ratios.

\[
\gamma_{pq} \equiv \langle \Psi | a^\dagger_p a_q + a^\dagger_q a_p | \Psi \rangle \nonumber
\]

\[
\Gamma^rs_{pq} \equiv \langle \Psi | \frac{1}{2} a^\dagger_p \sigma_r a^\dagger_q \sigma_s + a^\dagger_p \sigma_s a^\dagger_q \sigma_r + a^\dagger_p a_r a^\dagger_q a_s | \Psi \rangle \nonumber
\]

G. Orbital Optimization

Like DOCI and pCCD, pADCCD is not invariant to the choice of orbital basis, since all open-shell determinants are neglected. Thus optimal orbitals need to be found to fully minimize the energy. In this section we introduce an orbital optimization method for pADCCD. First consider the one-body anti-Hermitian operator

\[
\sum_{p > q} \sum_{\sigma} \kappa_{pq} \left( a^\dagger_{p\sigma} a_{q\sigma} - a^\dagger_{q\sigma} a_{p\sigma} \right)
\]

which, when exponentiated, creates unitary orbital rotations (here \( \sigma \) indexes spin).

Given this rotation operator, we can generalize the energy to be

\[
E (\kappa) = \langle \Psi \kappa \rangle \langle \kappa | H e^{\kappa} | \Psi \kappa \rangle
\]

We expand this energy to second order in \( \vec{\kappa} \),

\[
E (\vec{\kappa}) \simeq E (0) + \vec{\kappa}^T \left( \frac{\partial E (\vec{\kappa})}{\partial \vec{\kappa}} \right)_{\vec{\kappa}=0} + \frac{1}{2} \vec{\kappa}^T \left( \frac{\partial^2 E (\vec{\kappa})}{\partial \vec{\kappa}^2} \right)_{\vec{\kappa}=0} \vec{\kappa}
\]

where we work at \( \vec{\kappa} = 0 \) by transforming the basis in which we express the Hamiltonian (i.e. by transforming the one- and two-electron integrals).

We define the energy gradient and hessian,

\[
\omega_{pq} = \frac{\partial E (\kappa)}{\partial \kappa_{pq}}
\]

\[
A_{pq,rs} = \frac{\partial^2 E (\kappa)}{\partial \kappa_{pq} \partial \kappa_{rs}}
\]

both of which are functions of the one- and two-electron reduced density matrices

H. Scaling

Having presented our optimization method, we now analyze its cost. In each iteration, we need to loop over occupied and virtual orbitals to compute local energies, RDMs, and derivative ratios. As the evaluation of the relevant determinants scale as \( n_{ex}^3 \), in which \( n_{ex} \) is the pair excitation level, the overall cost scales as \( n_x n_o n_u n_{ex}^3 \), in which \( n_x, n_o \), and \( n_u \) are the number of samples, number of occupied and unoccupied orbitals in reference determinant, respectively. Although this \( \mathcal{O} \left( N^6 \right) \) scaling looks much steeper than the \( \mathcal{O} \left( N^3 \right) \) scaling of pCCD, one should note the highly excited configurations are rarely sampled, and so in many molecules the \( n_{ex}^3 \) term in the scaling may behave more like a constant. In such a regime, pADCCD’s scaling may appear closer to \( \mathcal{O} \left( N^3 \right) \).

At the end of each iteration, we need to reset \( \vec{\kappa} \) to 0 via a basis rotation. The one- and two-electron integrals needed to represent Hamiltonian in the new basis can be

| number of \( H_2 \) | pADCCD | CISD |
|-----------------|--------|--------|
| 1               | 0.000±0.001 | 0.00 |
| 2               | 0.000±0.0003 | 0.64 |
| 3               | 0.003±0.008 | 1.21 |
| 4               | 0.001±0.006 | 1.73 |
| 5               | 0.001±0.006 | 2.21 |

TABLE I: Size consistency errors (absolute values) for well separated \( H_2 \) in kcal/mol.
evaluated at an $\mathcal{O}(N^3)$ cost. However, as the basis rotation is required only once per LM iteration, rather than once per sample, its cost is typically negligible compared to that of the sampling effort involved in optimizing cluster amplitudes.

III. RESULTS

A. Computational Details

pADCCD results were obtained using our own software for VMC in Hilbert space, with one- and two-electron integrals for the Hamiltonian taken from PySCF\cite{28}. The full configuration interaction (FCI) results were obtained from Molpro\cite{29} and CISD results from Psi4\cite{30}. pCCD and DOCI results were kindly shared by Peter A. Limacher\cite{31}. We froze N and O 1s orbitals at the RHF level. Sample size is taken to be $3.6 \times 10^6$. All statistical uncertainties were converged to less than 0.01eV in all cases.

B. Size Consistency Check

Before showing our examples, we first check the size consistency of pADCCD by calculating the energy of up to 5 well separated H$_2$ molecules in a STO-3G basis. As shown in Table I the error per molecule does not grow with the increase of system size like in CISD. Instead, it remains zero within statistical uncertainty. This result demonstrates pADCCD’s size consistency, as we proved in section IID.

C. LiH

We begin our results with a simple example, the LiH molecule in the cc-pVDZ basis\cite{32}. The system has only two valence electrons, and both DOCI and pCCD deliver almost exact results compared to FCI. Due to our method’s similarity with pCCD, we also expect nearly exact results. Figure 1 shows our results at 17 bond lengths between 0.9 Å and 4.4 Å. The non-parallelity error (NPE) defined as the difference between the largest and smallest error with respect to FCI along the potential surface is about 2 mE$_h$, confirming our expectations.

FIG. 1: Dissociation of LiH in cc-pVDZ basis set

FIG. 2: Dissociation of HF in 6-31G basis set

FIG. 3: Symmetric Dissociation of H$_2$O in 6-31G basis set with a bond angle of 109.57°
D. HF

We next turn our attention to the dissociation of hydrogen fluoride in a 6-31G basis. Figure 2 shows the absolute energy of pADCCD, along with DOCI, pCCD and FCI results. The results show that pADCCD, DOCI and pCCD are energetically almost indistinguishable. A close analysis shows that the NPE of pADCCD and pCCD with respect to FCI is about 17 and 18 mE\(_h\), respectively. Unlike LiH, where seniority zero based wave functions are near exact, we see a larger energy gap between all seniority zero based ansatzes and FCI. This is a remainder that while seniority zero wave functions are often effective for strong correlations, they do not capture all the details of weak correlation.

In Figure 2, we also plot the results using only RHF canonical orbitals and no further orbital optimization. As one can see, the results are quite poor when using RHF orbitals, especially when one stretches the bonds and the optimal orbitals become more and more localized.

E. H\(_2\)O

Our next example is the symmetric double dissociation of H\(_2\)O, as shown in Figure 3. Again, pADCCD provides nearly identical energies compared to DOCI and pCCD, and the NPE with respect to FCI is about 19 mE\(_h\) for pADCCD and 16 mE\(_h\) for pCCD. The coincidence of the pADCCD with pCCD and DOCI is thus true not just for one pair of strongly correlated electrons, but for two pairs as well. However, we can still see from the plot that like pCCD and DOCI, a significant amount of dynamic correlation is clearly missing in all these methods.

In order to show the importance of orbital optimization, we plot optimized and RHF canonical orbitals for H\(_2\)O at bond length 2.2 Å in Figure 4. It is very clear that at this stretched geometry, the optimized orbitals are much more localized than canonical orbitals. This localization is also seen in pCCD. Indeed, the qualitative difference between optimized orbitals and canonical orbitals emphasizes the necessity of orbital optimization.

F. N\(_2\)

Our final example is the dissociation of N\(_2\). As Figure 5 reveals, the difference between pCCD and pADCCD is more noticeable in N\(_2\), which is to be expected now that hextuples (the first excitation level at which the
ansatz forms are different) are needed for a qualitatively correct description of the dissociation. Indeed, we find that disabling hexetuple and higher excitations raises the pADCCD energy by 0.5 eV. Although their NPEs now differ noticeably, 56 meV for pCCD versus 65 meV for pADCCD, they are of the same order of magnitude and neither are close to quantitative. Achieving a more quantitative accuracy clearly requires a more flexible cluster expansion, but, as is well known [1], this route leads to qualitatively incorrect variational violations when pursued in a traditional CC approach. As variational violations are not possible in a VMC-based approach, it will be interesting in future to investigate more flexible expansions within the amplitude determinant framework.

IV. CONCLUSIONS

We have presented amplitude determinant coupled cluster with pairwise doubles (pADCCD) as a variational cousin to pCCD. Unlike the permanent-based coefficients of pCCD, pADCCD defines its expansion coefficients as amplitude determinants. Combined with variational Monte Carlo methods, this choice produces a method that is exact for an electron pair, size-consistent, polynomial cost, and variational. Initial tests on the dissociations of LiH, HF, H₂O and N₂ reveal that pADCCD and pCCD produce similar results, suggesting that the leading approximation in both theories is their limitation to the seniority zero sector rather than the choice of permanent versus determinant for defining the cluster expansion.

Like pCCD and other seniority zero approaches, pADCCD proves effective for describing some strong electron correlations but is unable to deliver quantitative accuracy, a difficulty that may in future be addressed in two different ways. First, one may seek to increase the flexibility of the cluster expansion. We know that generalizing pCCD into CCSD greatly improves the recovery of weak correlation effects near equilibrium, but that it also leads to unacceptable variational violations as bonds are stretched. Analogous generalizations of pADCCD cannot suffer this problem, and so exploring more sophisticated amplitude-determinant-based cluster expansions is one attractive option. Another approach would be to use an amplitude-determinant-based expansion as the trial function in projector Monte Carlo, which the low per-sample cost of pADCCD for low-lying configurations suggests may be an especially effective pairing. Of course, these avenues are not mutually exclusive, and we look forward to investigating both in future research.

V. ACKNOWLEDGEMENT

The authors thank Peter A. Limacher for sharing with us his raw pCCD and DOCI data. We also acknowledge funding from the Office of Science, Office of Basic Energy Sciences, the US Department of Energy, Contract No. DE-AC02-05CH11231. Calculations were performed using the Berkeley Research Computing Savio cluster.

Appendix

For completeness, we include here expression for the pADCCD orbital rotation gradient and hessian. These should provide everything needed for the Newton-Raphson algorithm we use for orbital optimization. The expressions are similar to those in the pCCD paper [14].

The energy can be written as

$$E(\kappa) = \langle \Psi | e^{-\kappa H} e^{\kappa} | \Psi \rangle$$

where we evaluate the orbital rotation unitary transformation as $\exp(\kappa)$ and transform the integrals into this new basis.

The orbital gradient is

$$\left( \frac{\partial E(\kappa)}{\partial \kappa_{pq}} \right)_{\kappa=0} = \mathcal{P}_{pq} \sum_{\sigma} \langle [H, a_{p\sigma}^\dagger a_{q\sigma}] \rangle$$

$$= \mathcal{P}_{pq} \sum_{uvlt} [(uv|tp) \Gamma_{ut}^{uv} + (up|tv) \Gamma_{ut}^{vu} - (uv|qt) \Gamma_{up}^{vt} - (qv|tu) \Gamma_{pl}^{vu}]$$

with

$$\kappa = \sum_{p \neq q} \sum_{\sigma} \kappa_{pq} \left( a_{p\sigma}^\dagger a_{q\sigma} - a_{q\sigma}^\dagger a_{p\sigma} \right)$$

and the notation for the expectation value means

$$\langle O \rangle = \langle \Psi | \hat{O} | \Psi \rangle$$

Similarly, the Hessian is

$$A_{pq,rs} = \left( \frac{\partial^2 E(\kappa)}{\partial \kappa_{pq} \partial \kappa_{rs}} \right)_{\kappa=0}$$

$$= \frac{1}{2} \mathcal{P}_{pq} \mathcal{P}_{rs} \sum_{\sigma \tau} \langle [ [H, a_{p\sigma}^\dagger a_{q\tau}], a_{r\tau}^\dagger a_{s\sigma} ] \rangle$$

We obtain

$$A_{pq,rs} = \mathcal{P}_{pq} \mathcal{P}_{rs} \frac{1}{2} \sum_{uvlt} \delta_{qr} \left( (uv|tp) \Gamma_{ut}^{vu} + (up|tv) \Gamma_{ut}^{vu} \right.$$

$$+ (uv|st) \Gamma_{up}^{vt} + (sv|tu) \Gamma_{pl}^{vu} \right)$$

$$+ \delta_{ps} \left( (uv|qt) \Gamma_{ut}^{vt} + (qv|tu) \Gamma_{ru}^{vu} + (uv|tr) \Gamma_{ut}^{vu} + (qv|tu) \Gamma_{pl}^{vu} \right)$$

$$+ \sum_{uv} \left( (up|vr) \Gamma_{uv}^{qs} + (uv|tp) \Gamma_{uv}^{qs} \right)$$

$$+ (ut|qr) \Gamma_{up}^{ts} + (qv|tr) \Gamma_{up}^{ts} + (qr|tu) \Gamma_{pl}^{su}$$

$$+ (ut|sp) \Gamma_{ut}^{tp} + (up|st) \Gamma_{up}^{tp} + (su|tp) \Gamma_{ut}^{tp} + (sp|tu) \Gamma_{pl}^{tp}$$

$$- \sum_{tu} \left( (ut|qr) \Gamma_{up}^{ts} + (qv|tr) \Gamma_{up}^{ts} + (qr|tu) \Gamma_{pl}^{su} \right.$$
[1] T. Helgaker, P. Jøgensen, and J. Olsen, Molecular Electronic Structure Theory, John Wiley and Sons, Ltd, West Sussex, England, 2000.
[2] C. Møller and M. S. Plesset, Phys. Rev 46, 618 (1934).
[3] S. Grimme and M. Parac, ChemPhysChem 4, 292 (2003).
[4] E. Neuscamman, J. Chem. Theory Comput 12, 3149 (2016).
[5] A. Subedi, L. Zhang, D. J. Singh, and M. H. Du, Phys. Rev. B 78, 134514 (2008).
[6] P. Ring and P. Schuck, The Nuclear Many-Body Problem, Springer, Berlin, 2000.
[7] L. Bytautas, T. M. Henderson, C. A. Jiménez-Hoyos, J. K. Ellis, and G. E. Scuseria, J. Chem. Phys 135, 044119 (2011).
[8] F. Weinhold and E. B. W. Jr., J. Chem. Phys 46, 2752 (1967).
[9] T. L. Allen and H. Shull, J. Phys. Chem 66, 2281 (1962).
[10] D. W. Smith and S. J. Fogel, J. Phys. Chem 43, S91 (1965).
[11] G. D. P. III and R. J. Barlett, J. Chem. Phys. 76, 1910 (1982).
[12] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. 157, 479 (1989).
[13] T. Stein, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys 140, 214113 (2014).
[14] T. M. Henderson, I. W. Bulik, T. Stein, and G. E. Scuseria, J. Chem. Phys 141, 244104 (2014).
[15] I. W. Bulik, T. M. Henderson, and G. E. Scuseria, J. Chem. Theory Comput. 11, 3171 (2015).
[16] R. J. Bartlett and M. Musial, Rev. Mod. Phys 79, 291 (2007).
[17] D. I. Lyakh, M. Musial, V. F. Lotrlich, and R. J. Barlett, Chem. Rev 112, 182 (2012).
[18] B. Cooper and P. J. Knowles, J. Chem. Phys 133, 234102 (2010).
[19] J. A. Gomez, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys 144, 244117 (2016).
[20] D. W. Small, K. V. Lawler, and M. Head-Gordon, J. Chem. Theory Comput 10, 2027 (2014).
[21] P. A. Limacher et al., J. Chem. Theory Comput 9, 1394 (2013).
[22] J. B. Robinson and P. J. Knowles, J. Chem. Phys 136, 054114 (2012).
[23] C. J. Umrigar, J. Chem. Phys. 143, 164105 (2015).
[24] M. Casula and S. Sorella, J. Chem. Phys. 119, 6500 (2003).
[25] C. J. Umrigar and C. Filippi, Phys. Rev. Lett 94, 150201 (2005).
[26] M. P. Nightingale and V. Melik-Alaverdian, Phys. Rev. Lett. 87, 043401 (2001).
[27] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. 98, 110201 (2007).
[28] Psycf, a quantum chemistry package written in python, see http://chemists.princeton.edu/chan/software/pyscf.
[29] H.-J. Werner et al., MOLPRO, version 2012.1, a package of ab initio programs, see http://www.molpro.net.
[30] J. M. Turney et al., WIREs Comput. Mol. Sci. 2, 556 (2012).
[31] P. A. Limacher, personal communication, 2016.
[32] T. Dunning, J. Chem. Phys 90, 1007 (1989).
[33] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys 56, 2257 (1972).