

# Projected Coupled Cluster Amplitudes from a Different Basis Set As Initial Guess

Marco Caricato,\* Gary W. Trucks, and Michael J. Frisch

Gaussian, Inc., 340 Quinipiac St. Bldg. 40, Wallingford, Connecticut 06492, United States

**S** Supporting Information

**ABSTRACT:** Many model chemistry schemes require a sequence of high level calculations, often at the coupled cluster (CC) level, with increasingly larger basis sets. The CC equations are solved iteratively, and the rate of convergence strongly depends on the quality of the initial guess. Here, we propose a strategy to define a better guess from a previous CC calculation with a different basis set by employing the concept of corresponding orbitals.<sup>1,2</sup> Only the part of the converged amplitudes from the previous calculation that has a large correspondence to the space spanned by the new basis set is projected and used as a new starting point. The computational time for the projection is negligible and significantly reduces the number of cycles necessary for convergence in comparison to the standard guess based on the first order wave function. Numerical results are presented for ground and excited state calculations with the CC singles and doubles (CCSD) and the equation of motion CCSD (EOM-CCSD) methods with the restricted and unrestricted Hartree–Fock (HF) reference functions. However, this approach is more general and can be applied to any truncation of the cluster expansion and reference function.

## 1. INTRODUCTION

Single reference coupled cluster (CC) methods currently represent the most computationally balanced level of theory in quantum chemistry to obtain accurate and consistent descriptions of ground and excited state wave functions for small- and medium-sized molecules.<sup>3</sup> Various levels of truncation of the wave operator can be defined and used, often depending on what is computationally affordable for a particular system. Among these, the most widely used approximation of the complete excitation manifold includes single and double excitation operators (CCSD).<sup>4</sup> Another factor that plays an important role in the accuracy of the calculation is the finite one-electron expansion basis set. Indeed, CC methods show a significant sensitivity to the quality of the basis set, and often large bases are necessary to reach convergence in the description of properties.<sup>5</sup> This is particularly true for excited states where diffuse functions are frequently unavoidable even for a qualitatively sound description of electronic transitions. Also in this respect, performing production calculations in a reasonable time poses a limit to the size of the basis sets that can be effectively used.

In recent years, many efforts have been carried out to extend the range of applications of CC theory to larger systems and/or larger basis expansions.<sup>3</sup> Examples of these include introducing further approximations in the wave operator,<sup>6–8</sup> combining various levels of theory in hybrid methods,<sup>9–11</sup> considering explicit electron correlation to improve basis set convergence,<sup>12–14</sup> and defining protocols for thermochemistry and basis set extrapolations.<sup>15–20</sup> In most cases, CC equations are solved iteratively, stopping when certain convergence criteria are met. The rate of convergence is fundamentally influenced by two factors: the quality of the initial guess and the extrapolation technique used to define the amplitudes in the next iterative cycle. The latter is normally dealt with using standard techniques

like the direct inversion of the iterative subspace (DIIS)<sup>21–23</sup> or the reduced linear equation (RLE)<sup>24,25</sup> methods for ground state calculations and extended Davidson algorithms for excited states.<sup>26–28</sup> In this article, we concentrate on the former issue, i.e., how to define a better starting point for a CC calculation assuming that a former calculation with another, usually smaller, basis set has already been performed. Although we will focus the discussion on ground state CCSD and equation of motion CCSD (EOM-CCSD)<sup>29</sup> for excited states with the canonical form of the restricted and unrestricted Hartree–Fock (RHF/UHF) wave functions as a reference, our approach is general and can be applied to other truncated CC wave functions.

This approach is useful when multiple calculations with various basis sets are required, such as basis set extrapolation methods and model chemistry protocols.<sup>15–19</sup> Another example is excited state studies when exploratory calculations on many states are performed with a smaller basis that can provide semiquantitative results, and then the calculation for some specific state is refined with a larger basis set. We provide examples of these type of applications in this paper. Although no example is reported here, the application of such methodology to the Brueckner doubles (BD)<sup>30,31</sup> variant of CCSD (or to geometry optimizations) provides large savings in computational time as the orbitals are rotated at each macroiteration step. It can even be used in the reverse order, i.e., passing from a larger to a smaller basis, for instance, for investigating higher excited states with a less demanding basis set after converging lower ones.

This paper is organized as follows: The theory is presented in section 2, the results of the test calculations are reported in section 3, and concluding remarks are discussed in section 4.

**Received:** November 7, 2010

**Published:** February 28, 2011

## 2. THEORY

The usual guess used for ground state CCSD is to set to zero the singles amplitudes,  $t_i^a$  (when the canonical orbitals are used and the Fock matrix is diagonal), and use the first order wave function for the doubles amplitudes:

$$t_{ab}^{ij(1)} = \frac{\langle ij||ab \rangle}{\Delta_{ij}^{ab}} \quad (1)$$

where  $i, j$  and  $a, b$  represent occupied and virtual orbitals, respectively,  $\langle ij||ab \rangle$  are the antisymmetrized two-electron integrals, and  $\Delta_{ij}^{ab} = f_{ii} + f_{jj} - f_{aa} - f_{bb}$  is the combination of the diagonal elements of the Fock matrix (orbital energies for canonical orbitals). This is a sensible way to define the guess, but it may require many cycles to reach convergence if the second order energy from the perturbation theory expansion is not a good approximation of the correlation energy. For the evaluation of molecular properties, the calculation of the electronic density is required. In the case of CC wave functions, another set of amplitudes is necessary, called  $\Lambda$ - or  $Z$ -vector amplitudes,<sup>32,33</sup> that can use the converged  $t$  amplitudes as an initial guess. For EOM-CCSD, two sets of amplitudes (for each excited state) are necessary for the calculation of transition energies and properties that correspond to the left and right eigenvectors of the EOM-CCSD Hamiltonian. It has been shown that the most efficient way to compute them is to solve for the two sets separately.<sup>26–28</sup> For the solution of the first set, usually the right eigenvectors, the converged configuration interaction singles (CIS) eigenvectors are used for the guess. Since the CIS eigenvectors only include single excitations, the double excitation component of the EOM-CCSD eigenvectors are initialized to zero. When the right eigenvectors are found, they are used as the guess for the left ones. Although the CIS vectors are a sensible choice given that EOM-CCSD mainly provides an accurate description of one-electron transitions where the contribution of single excitations is predominant, the convergence can be slow if double excitation contributions are significant.

We propose to project the converged amplitudes from a previous calculation with a certain basis set onto the space spanned by a new basis set and use the projected amplitudes as a starting guess. This requires determining which part of the space in the new basis expansion is spanned by the old basis. This can be achieved using *corresponding orbitals*.<sup>1,2</sup> In the following, we assume real orbitals. We first define the overlap matrices  $\mathbf{S}$  between the occupied molecular orbitals (MOs) and  $\bar{\mathbf{S}}$  between the virtual MOs in the two basis sets:

$$\begin{aligned} S_{ij} &= \langle \tilde{i} | j \rangle = \sum_{\mu, \nu} c_{\mu i} \tilde{c}_{\mu} \langle \tilde{\mu} | \nu \rangle c_{\nu j} \\ \bar{S}_{ab} &= \langle \tilde{a} | b \rangle = \sum_{\mu, \nu} c_{\mu a} \tilde{c}_{\mu} \langle \tilde{\mu} | \nu \rangle c_{\nu b} \end{aligned} \quad (2)$$

where the indexes with the symbol  $\sim$  refer to the *old* basis set,  $c$  represents the MO coefficients, and  $\mu$  and  $\nu$  are the atomic orbitals (AOs). We assume the MOs to be orthonormalized within each set. The matrices  $\mathbf{S}$  and  $\bar{\mathbf{S}}$  are then decomposed using the single value decomposition (SVD) technique:<sup>34</sup>

$$\begin{aligned} \mathbf{S} &= \mathbf{U} \sigma \mathbf{V}^+ \\ \bar{\mathbf{S}} &= \bar{\mathbf{U}} \bar{\sigma} \bar{\mathbf{V}}^+ \end{aligned} \quad (3)$$

where  $\mathbf{U}$  and  $\mathbf{V}^+$  and  $\bar{\mathbf{U}}$  and  $\bar{\mathbf{V}}^+$  are unitary transformations of the *old* and *new* occupied and virtual MOs, and  $\sigma$  and  $\bar{\sigma}$  are diagonal matrices. This does not affect the wave function, as it is invariant

under unitary transformations of the occupied–occupied and virtual–virtual blocks of MOs. The values of  $\sigma$  and  $\bar{\sigma}$  are  $0 \leq \sigma_{ii} \bar{\sigma}_{aa} \leq 1$  (the number of nonzero values is equal to the size of the smaller of the MO sets). The transformations  $\mathbf{U}$ ,  $\mathbf{V}^+$ ,  $\bar{\mathbf{U}}$ , and  $\bar{\mathbf{V}}^+$  can be interpreted as the rotations of the two basis sets into two sets that most correspond to each other. If the values of  $\sigma$  and  $\bar{\sigma}$  are organized in descending order, we can define the projection matrices  $\mathbf{P}$  and  $\bar{\mathbf{P}}$  and the extraction matrices  $\mathbf{E}$  and  $\bar{\mathbf{E}}$  as

$$\begin{aligned} P_{ij} &= \sum_k^{\text{corr}} U_{ik} V_{jk} \\ \bar{P}_{ab} &= \sum_c^{\text{corr}} \bar{U}_{ac} \bar{V}_{bc} \end{aligned} \quad (4)$$

$$\begin{aligned} E_{ij} &= \sum_k^{\text{corr}} V_{ik} V_{jk} \\ \bar{E}_{ab} &= \sum_c^{\text{corr}} \bar{V}_{ac} \bar{V}_{bc} \end{aligned} \quad (5)$$

where the indexes  $k$  and  $c$  run over the *corresponding orbitals* with  $\sigma_{kk}$  and  $\bar{\sigma}_{cc} \geq 0.9$ . The projection matrices are used to map the amplitudes in the old basis set onto the new basis. The purpose of the matrices  $\mathbf{E}$  and  $\bar{\mathbf{E}}$ , when applied onto an orbital  $i$  or  $a$ , is to extract the contribution of such an orbital to the selected *corresponding orbitals*. Two sets of projection and extraction matrices are defined separately for electrons with spin  $\alpha$  and spin  $\beta$  for an unrestricted reference function. The new guess amplitudes can be calculated as

$$t_i^a = \sum_{\bar{i}\bar{a}} \bar{P}_{\bar{a}\bar{a}} t_{\bar{i}}^{\bar{a}} P_{\bar{i}\bar{i}} \quad (6)$$

$$t_{ij}^{ab} = t_{ij}^{ab(1)} - \sum_{klcd} \bar{E}_{ca} \bar{E}_{db} t_{kl}^{cd(1)} E_{ki} E_{lj} + \sum_{\bar{i}\bar{j}\bar{a}\bar{b}} \bar{P}_{\bar{a}\bar{a}} \bar{P}_{\bar{b}\bar{b}} t_{\bar{i}\bar{j}}^{\bar{a}\bar{b}} P_{\bar{i}\bar{i}} P_{\bar{j}\bar{j}} \quad (7)$$

where  $t_{ij}^{ab(1)}$  is the first order wave function as in eq 1. When the new guess for the amplitudes  $t_{ij}^{ab}$  in eq 7 is built, only the part of the first order amplitudes in eq 1 that gives no contribution to the *corresponding orbitals* in the old basis set is retained, by applying  $\mathbf{E}$  and  $\bar{\mathbf{E}}$ . The contraction in eq 7 scales as  $O(N^5)$ , where  $N$  is the number of basis functions, since the indexes are contracted one at a time. The same transformations in eqs 6 and 7 are employed for the  $\mathbf{Z}$  vector and the EOM-CCSD eigenvector amplitudes. We used the expression in eq 1 to define the first order term for the projected  $\mathbf{Z}$  vector double amplitudes, whereas this is set to zero for EOM-CCSD. This procedure has been implemented in the Gaussian 09 suite of programs.<sup>35</sup>

## 3. RESULTS

All of the data in this section is reported as the ratio of the number of iterative steps with the projected guess compared to the calculations with the normal guess. The absolute number of steps is reported in the Supporting Information. The extrapolation technique used to solve the ground state  $t$  and  $\mathbf{Z}$  vector equations is the RLE scheme of Purvis and Bartlett,<sup>24</sup> while an extended Davidson algorithm is used for the EOM-CCSD equations.<sup>27,28</sup> The details of these algorithms are described in the literature and need not be repeated here. The choice of these algorithms is due to their efficiency for the solution of the CC equations compared to other options, as demonstrated in refs 24, 25, 27, and 28.

We start with two examples where this methodology is applied in production-like calculations using compound energy models. Section 3.1 reports the application of the W1U theory<sup>36</sup> on the phenyl and phenol radicals with and without the CC amplitude projection. The projected guess is also tested in ground state calculations in section 3.2. The test molecules include formaldehyde, butadiene, ethene, and acetone for the RHF reference and the vinyl radical for the UHF reference. The geometries are taken from ref 37 for the closed shell species and from ref 38 for the radical. The same systems and basis sets are used for the excited states in section 3.3. We consider one and three states per irrep (thus, a total of 4 and 12 states for formaldehyde, butadiene, and acetone; 2 and 6 for the vinyl radical; and 8 and 24 for ethene). All of these molecules have been thoroughly studied experimentally and theoretically.<sup>5,37–39</sup> The values for the transition energies and oscillator strengths for all of the basis sets considered in this work can be found in the Supporting Information.

The convergence criteria for ground and excited state calculations are based on changes in both energy and wave function (the same criteria are used for all the basis sets). In particular, for the W1U case, convergence is achieved for changes  $<10^{-7}$  au in the energy and  $<10^{-5}$  in the wave function ( $t$  amplitudes). For the other ground state cases, the thresholds are  $10^{-8}$  au for the energy and  $10^{-6}$  for the wave function ( $t$  and  $Z$  amplitudes). For the EOM-CCSD calculations, the thresholds are  $10^{-7}$  au for the excitation energy and  $10^{-5}$  for the wave function (right and left eigenvectors).

Various sequences of basis sets are examined where each calculation uses the projected guess from the previous basis. This

sequences could be used, for instance, in basis set extrapolation techniques. We do not consider all of the possible combinations of basis sets as guesses for the others, as this would make the interpretation of the results confusing. The goal for these tests is to evaluate when a basis set may be a good starting point for projecting the amplitudes onto a larger basis. Therefore, we also include choices that may be not recommendable in production calculations (i.e., basis sets without diffuse functions for excited states). In these test cases, all the excited states are considered for all the basis sets, although it is more likely that in production calculations only a few states may be refined with a larger basis set after some exploratory calculations with a smaller basis.

**3.1. W1U Tests.** W1U<sup>36</sup> is the variation of the Weizmann-1 theory (W1)<sup>18,19</sup> that uses the unrestricted CCSD(T)<sup>40</sup> method as the highest level of theory and is used to obtain highly accurate thermochemical data for open shell systems (average accuracy within 1 kcal/mol). The method consists of seven subcalculations:

- Step 1: Geometry optimization at the B3LYP level with the cc-pVTZ+d basis set
- Step 2: Frequency at the B3LYP level with the cc-pVTZ+d basis set
- Step 3: Energy calculation at the CCSD(T) level with the augh-cc-pVDZ+2df basis set
- Step 4: Energy calculation at the CCSD(T) level with the augh-cc-pVTZ+2df basis set
- Step 5: Energy calculation at the CCSD level with the augh-cc-pVQZ+2df basis set

**Table 3. Ground State, Second Sequence of Basis Sets<sup>a</sup>**

	$t$ amplitudes		$Z$ vector	
	aug-cc-VDZ <sup>b</sup>	d-aug-cc-pVDZ	aug-cc-VDZ <sup>b</sup>	d-aug-cc-pVDZ
formaldehyde	0.93	0.73	1.00	0.85
ethene	0.85	0.69	0.85	0.90
butadiene	0.76	0.50	1.00	0.77
acetone	0.88	0.71	0.88	0.75
vinyl radical	0.89	0.63	0.94	0.80

<sup>a</sup> Ratio of the number of cycles with the projected amplitude guess with respect to the standard guess. <sup>b</sup> Initial basis set: cc-pVDZ. Amplitudes are projected from the previous basis in the sequence.

**Table 1. W1U Ratio of the Number of Cycles (For the Last Four Steps) and of the Total CPU Time with the Projected Amplitude Guess with Respect to the Standard Guess**

	$C_6H_5^*$	$C_6H_5O^*$
step 4	0.68	0.70
step 5	0.63	0.65
step 6	0.63	0.78
step 7	0.68	0.67
total time	0.79	0.74

**Table 2. Ground State, First Sequence of Basis Sets<sup>a</sup>**

	6-31+G <sup>b</sup>	6-31++G <sup>**</sup>	aug-cc-pVDZ	aug-cc-pVTZ	aug-cc-pVQZ
$t$ amplitudes					
formaldehyde	0.87	0.80	0.87	0.80	0.73
ethene	0.85	0.85	0.85	0.85	0.77
butadiene	0.76	0.76	0.82	0.82	
acetone	0.94	0.76	0.82	0.82	
vinyl radical	0.84	0.78	0.89	0.75	
$Z$ vector					
formaldehyde	1.08	0.85	0.92	0.92	0.85
ethene	0.85	0.92	0.85	0.85	0.71
butadiene	0.76	0.81	1.08	1.08	
acetone	0.94	0.76	0.82	0.93	
vinyl radical	1.07	1.00	0.94	0.82	

<sup>a</sup> Ratio of the number of cycles with the projected amplitude guess with respect to the standard guess. <sup>b</sup> Initial basis set: 6-31G\*. Amplitudes are projected from the previous basis in the sequence.

- Step 6: Energy calculation at the CCSD(T) level with the MTSmall basis set
- Step 7: Energy calculation at the CCSD(T) level with the

MTSmall basis set with correlation of the core electrons followed by an energy extrapolation. MTSmall consists of a completely decontracted cc-pVTZ basis set with two tight d and one tight f function added. We tested two radicals,  $\text{C}_6\text{H}_5^\bullet$  and  $\text{C}_6\text{H}_5\text{O}^\bullet$ , with and without the projection scheme presented in section 2. The geometries and the energy information are reported in the Supporting Information as well as the total number of iterations for the CC subcalculations. The ratios for

**Table 4. Ground State, Third Sequence of Basis Sets<sup>a</sup>**

	<i>t</i> amplitudes		Z vector	
	aug-cc-VTZ <sup>b</sup>	d-aug-cc-pVTZ	aug-cc-VTZ <sup>b</sup>	d-aug-cc-pVTZ
formaldehyde	0.87	0.67	0.92	0.71
ethene	0.85	0.69	0.85	0.82
vinyl radical	0.90	0.60	1.00	0.73

<sup>a</sup> Ratio of the number of cycles with the projected amplitude guess with respect to the standard guess. <sup>b</sup> Initial basis set: cc-pVTZ. Amplitudes are projected from the previous basis in the sequence.

the iterations for the steps 4–7 and for the total time are reported in Table 1. The savings in total time is smaller than for the iterative solution of the CCSD equations in the individual steps. This is because the perturbative triples correction to the CCSD energy in steps 4, 6, and 7 is not iterative, and therefore it does not benefit from the improved guess. Nonetheless, the savings in total CPU time is 20–25%, which is remarkable considering the length of this type of calculation. Hence, these test cases show a considerable reduction in the number of iterations and in the total calculation time when the projection scheme is used to define the guess for the next subcalculation. In fact, for open shell species with large spin contamination in the reference function, CCSD requires many iterations to remove the contribution of the first spin contaminant from the energy.<sup>41,42</sup> Thus, starting with projected *t* amplitudes from another basis significantly reduces the number of iterations.

**3.2. Ground State.** The results for the ground state calculations are reported in Tables 2–4. The general trends show a significant reduction of the computational cost for the *t* amplitudes, in the range of 10–30%, with peaks of 40–50% in Tables 3 and 4. The savings is smaller for the Z vector, which indicates that the converged *t* amplitudes are a good starting point. The results

**Table 5. Excited State, First Sequence of Basis Sets<sup>a</sup>**

	Right Eigenvector				
	6-31+G <sup>*b</sup>	6-31++G <sup>**</sup>	aug-cc-pVDZ	aug-cc-pVTZ	aug-cc-pVQZ
three states/irrep					
formaldehyde	0.97	0.90	0.89	0.91	0.86
ethene	0.96	0.88	0.93	0.92	0.86
butadiene	1.00	0.86	0.87	0.87	
acetone	1.01	0.91	0.86	0.85	
vinyl radical	1.03	0.89	0.84	0.80	
one state/irrep					
formaldehyde	0.97	0.87	0.89	0.84	0.72
ethene	1.07	0.95	0.94	0.88	0.83
butadiene	1.12	0.89	0.87	0.81	
acetone	1.00	0.93	0.88	0.76	
vinyl radical	0.87	0.84	0.75	0.78	
	Left Eigenvector				
	6-31+G <sup>*b</sup>	6-31++G <sup>**</sup>	aug-cc-pVDZ	aug-cc-pVTZ	aug-cc-pVQZ
three states/irrep					
formaldehyde	0.97	1.02	1.00	0.99	1.02
ethene	1.02	1.04	1.03	1.03	1.04
butadiene	0.99	1.05	1.01	1.09	
acetone	1.02	1.02	1.05	1.06	
vinyl radical	0.96	1.01	1.00	0.97	
one state/irrep					
formaldehyde	1.00	0.98	1.02	1.11	0.97
ethene	1.04	1.12	1.09	1.05	1.03
butadiene	0.97	1.05	1.06	0.95	
acetone	0.93	1.03	1.11	1.05	
vinyl radical	1.07	1.00	1.04	1.17	

<sup>a</sup> Ratio of the number of cycles with the projected amplitude guess with respect to the standard guess. <sup>b</sup> Initial basis set: 6-31G\*. Amplitudes are projected from the previous basis in the sequence.



Table 6. Excited State, Second Sequence of Basis Sets<sup>a</sup>

	Right Eigenvector			
	three states/irrep		one state/irrep	
	aug-cc-pVDZ <sup>b</sup>	d-aug-cc-pVDZ	aug-cc-pVDZ <sup>b</sup>	d-aug-cc-pVDZ
formaldehyde	1.03	0.90	0.94	0.85
ethene	0.99	0.93	1.07	0.96
butadiene	0.99	0.81	1.00	0.92
acetone	1.02	0.84	0.92	0.79
vinyl radical	1.02	0.79	0.94	0.67

  

	Left Eigenvector			
	three states/irrep		one state/irrep	
	aug-cc-pVDZ <sup>b</sup>	d-aug-cc-pVDZ	aug-cc-pVDZ <sup>b</sup>	d-aug-cc-pVDZ
formaldehyde	0.99	0.93	1.00	0.93
ethene	1.01	0.94	1.05	0.99
butadiene	1.00	0.95	0.95	1.07
acetone	1.00	0.95	1.06	1.05
vinyl radical	0.97	1.07	1.04	0.86

<sup>a</sup> Ratio of the number of cycles with the projected amplitude guess with respect to the standard guess. <sup>b</sup> Initial basis set: cc-pVDZ. Amplitudes are projected from the previous basis in the sequence.

in all of the sequences also show that projecting the amplitudes from a basis set without diffuse functions to one with diffuse functions may not be very efficient. On the other hand, a calculation with a basis set with “more” diffuse functions considerably benefits from the projected guess compared to the standard one.

**3.3. Excited States.** Excited state results are reported in Tables 5–7. Good performance is obtained for the right eigenvectors, although the savings is less than for the ground state *t* amplitudes. The results are better when one state per irrep is considered, while the savings is somewhat less when three states per irrep are computed. This is not surprising, as higher states are more diffuse, and increasing the basis set may lead to larger changes in the wave function. More than for the ground state, the presence of diffuse functions in the basis set used for the projection is important for excited states. Indeed, the CIS eigenvectors solved with the diffuse basis functions are a better guess than projected EOM-CCSD amplitudes without diffuse functions (6-31G\* → 6-31+G\*, cc-pVDZ → aug-cc-pVDZ, or cc-pVTZ → aug-cc-pVTZ). When a basis is augmented with an additional set of diffuse functions (aug-cc-pVDZ → d-aug-cc-pVDZ or aug-cc-pVTZ → d-aug-cc-pVTZ), the projected guess approach is better than the CIS guess.

For the left eigenvectors, using the converged right eigenvectors (with the same basis set) as the guess is almost consistently better than projecting the left eigenvectors from another basis set, similar to what was found for the ground state *Z* vector.

## 4. CONCLUSIONS

In this paper, we present a projection technique to use converged CC amplitudes as the initial guess for a new calculation with a different basis set. This technique is based on the corresponding orbitals approach and can be used for ground and excited state calculations. We test a variety of basis sets to define the guess for the next (ground and excited state) calculation on a

Table 7. Excited State, Third Sequence of Basis Sets<sup>a</sup>

	Right Eigenvector			
	three states/irrep		one state/irrep	
	aug-cc-pVTZ <sup>b</sup>	d-aug-cc-pVTZ	aug-cc-pVTZ <sup>b</sup>	d-aug-cc-pVTZ
formaldehyde	0.98	0.84	0.91	0.79
ethene	0.96	0.92	1.03	0.87
vinyl radical	1.01	0.80	0.91	0.58

  

	Left Eigenvector			
	three states/irrep		one state/irrep	
	aug-cc-pVTZ <sup>b</sup>	d-aug-cc-pVTZ	aug-cc-pVTZ <sup>b</sup>	d-aug-cc-pVTZ
formaldehyde	0.99	0.93	1.02	0.99
ethene	0.99	0.95	0.98	0.98
vinyl radical	0.94	0.86	1.21	0.87

<sup>a</sup> Ratio of the number of cycles with the projected amplitude guess with respect to the standard guess. <sup>b</sup> Initial basis set: cc-pVTZ. Amplitudes are projected from the previous basis in the sequence.

series of small closed and open shell molecules. We also report two examples of calculations with the WIU theory as representative of production level calculations.

The results in this work indicate that the projection strategy is not generally useful to generate an initial guess from another basis set when such result is not already available. In fact, the combined CPU time of a calculation with a smaller basis set and that with a larger basis with the projected guess is always larger than directly performing the calculation with the standard guess. Nevertheless, the projected guess is certainly useful when the calculation with a smaller basis set is already available, which is the case with basis set extrapolation techniques and model chemistry protocols such as W1. In these cases, the projected guess reduces the number of iterations for the solution of the CC equations by 25–30% without adding any significant computational cost. This projection is also very useful in geometry optimizations, especially when approaching the local minimum. For excited states, this methodology is suitable when the interest is in refining the calculation of a particular state by increasing the basis set with more diffuse functions (which are often necessary for accurate results<sup>37</sup>).

For ground state calculations, the *t* amplitudes benefit from the projected guess, and we find a reduction in the number of cycles from 10% to 50%, while the *Z* amplitudes do not benefit as much. For excited state calculations, the projected guess for the right eigenvectors is considerably better to use than the CIS eigenvectors when additional diffuse functions are included in the basis set. In such a case, the number of iterations for the solution of the EOM-CCSD equations is reduced by 10–20%. On the other hand, for the left eigenvectors, the guess that uses the converged right eigenvectors is very often a better choice than projecting left eigenvectors from a smaller basis set.

We also point out that amplitudes computed from other approximations can be projected as well, for example, projecting CCD or CISD amplitudes as a guess for CCSD, and that looser thresholds for the convergence of the calculation with a smaller basis set can be set if the interest is only in creating the guess. This strategy can even be used in the reverse order, i.e., projecting from a larger to a smaller basis, for example, investigating higher excited states with a less demanding basis set after converging lower ones.

## ■ ASSOCIATED CONTENT

**S Supporting Information.** Total number of iterative cycles for the solution of the CCSD and EOM-CCSD equations for all the molecules for each basis set, with the standard and the projected guess. Excitation energies and oscillator strengths for the EOM-CCSD calculations for each basis set. Geometries and energetics for the W1U test molecules. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

## Corresponding Author

\*E-mail: [marco@gaussian.com](mailto:marco@gaussian.com).

## ■ REFERENCES

- (1) Amos, A. T.; Hall, G. G. *Proc. R. Soc. London* **1961**, 263, 483–493.
- (2) King, H. F.; Stanton, R. E.; Kim, H.; Wyatt, R. E.; Parr, R. G. *J. Chem. Phys.* **1967**, 47, 1936–1941.
- (3) Bartlett, R. J.; Musial, M. *Rev. Mod. Phys.* **2007**, 79, 291–352.
- (4) Bartlett, R. J.; Purvis, G. D. *Int. J. Quantum Chem.* **1978**, 14, 561–581.
- (5) Wiberg, K. B.; de Oliveira, A. E.; Trucks, G. W. *J. Phys. Chem. A* **2002**, 106, 4192–4199.
- (6) Christiansen, O.; Koch, H.; Jorgensen, P. *Chem. Phys. Lett.* **1995**, 243, 409–418.
- (7) Koch, H.; Christiansen, O.; Jorgensen, P.; Olsen, J. *Chem. Phys. Lett.* **1995**, 244, 75–82.
- (8) Schutz, M.; Werner, H. *Chem. Phys. Lett.* **2000**, 318, 370–378.
- (9) Osted, A.; Kongsted, J.; Mikkelsen, K. V.; Christiansen, O. *J. Phys. Chem. A* **2004**, 108, 8646–8658.
- (10) Kowalski, K.; Valiev, M. *J. Phys. Chem. A* **2006**, 110, 13106–13111.
- (11) Caricato, M.; Vreven, T.; Trucks, G. W.; Frisch, M. J.; Wiberg, K. B. *J. Chem. Phys.* **2009**, 131, 134105.
- (12) Shiozaki, T.; Kamiya, M.; Hirata, S.; Valeev, E. F. *J. Chem. Phys.* **2008**, 129, 071101.
- (13) Valeev, E. F.; Crawford, T. D. *J. Chem. Phys.* **2008**, 128, 244113.
- (14) Tew, D. P.; Klopper, W.; Neiss, C.; Haettig, C. *Phys. Chem. Chem. Phys.* **2007**, 9, 1921–1930.
- (15) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, 94, 7221–7230.
- (16) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, 109, 7764–7776.
- (17) Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, 110, 2822–2827.
- (18) Martin, J. M. L.; de Oliveira, G. *J. Chem. Phys.* **1999**, 111, 1843–1856.
- (19) Parthiban, S.; Martin, J. M. L. *J. Chem. Phys.* **2001**, 114, 6014–6029.
- (20) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, 121, 11599–11613.
- (21) Pulay, P. *Chem. Phys. Lett.* **1980**, 73, 393–398.
- (22) Pulay, P. *J. Comput. Chem.* **1982**, 3, 556–560.
- (23) Scuseria, G. E.; Lee, T. J.; Schaefer, H. F. *Chem. Phys. Lett.* **1986**, 130, 236–239.
- (24) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1981**, 75, 1284–1292.
- (25) Trucks, G. W.; Noga, J.; Bartlett, R. J. *Chem. Phys. Lett.* **1988**, 145, 548–554.
- (26) Davidson, E. R. *J. Comput. Phys.* **1975**, 17, 87–94.
- (27) Hirao, K.; Nakatsuji, H. *J. Comput. Phys.* **1982**, 45, 246–254.
- (28) Caricato, M.; Trucks, G. W.; Frisch, M. J. *J. Chem. Theory Comput.* **2010**, 6, 1966–1970.
- (29) Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1993**, 98, 7029–7039.
- (30) Handy, N. C.; Pople, J. A.; Head-Gordon, M.; Raghavachari, K.; Trucks, G. W. *Chem. Phys. Lett.* **1989**, 164, 185–192.
- (31) Kobayashi, R.; Handy, N. C.; Amos, R. D.; Trucks, G. W.; Frisch, M. J.; Pople, J. A. *J. Chem. Phys.* **1991**, 95, 6723–6733.
- (32) Salter, E. A.; Trucks, G. W.; Bartlett, R. J. *J. Chem. Phys.* **1989**, 90, 1752–1766.
- (33) Gauss, J.; Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1991**, 95, 2623–2638.
- (34) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. In *Numerical Recipes in Fortran*, 2nd ed.; Cambridge University Press: Cambridge, U.K., 1992; pp 51–63.
- (35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Norm, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
- (36) Barnes, E. C.; Petersson, G. A.; Montgomery, J. A., Jr.; Frisch, M. J.; Martin, J. M. L. *J. Chem. Theory Comput.* **2009**, 5, 2687–2693.
- (37) Caricato, M.; Trucks, G. W.; Frisch, M. J.; Wiberg, K. B. *J. Chem. Theory Comput.* **2010**, 6, 370–383.
- (38) Koziol, L.; Levchenko, S. V.; Krylov, A. I. *J. Phys. Chem. A* **2006**, 110, 2746–2758.
- (39) Barone, V.; Bloino, J.; Biczysko, M. *Phys. Chem. Chem. Phys.* **2010**, 12, 1092–1101.
- (40) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, 157, 479–483.
- (41) Chen, W.; Schlegel, H. B. *J. Chem. Phys.* **1994**, 101, 5957–5968.
- (42) Stanton, J. F. *J. Chem. Phys.* **1994**, 101, 371–374.