

High-Multiplicity Natural Orbitals in Multireference Configuration Interaction for Excited States

Zhen Lu and Spiridoula Matsika*

Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, United States

S Supporting Information

ABSTRACT: Multireference configuration interaction (MRCI) is a very useful tool in studying excited states, dissociation of molecules, and chemical systems with multireference character. In many cases however, poor computational scaling limits its use to small systems. In the past, several different approaches have been taken in order to make MRCI and other multireference methods more accessible when dealing with larger systems. Here we propose the use of high-multiplicity natural orbitals (HMNO) in order to improve several aspects of the MRCI calculation. Natural orbitals, derived from a configuration interaction with single and double excitations calculation on a high-multiplicity reference state, were used in lieu of the standard complete active space self-consistent field (CASSCF) canonical molecular orbitals. This work examines the ability of the MRCI/HMNO approach to reliably reproduce vertical excitation energies for singlet states as well as energies for conical intersections. It is found that the MRCI/HMNO approach reliably reproduces vertical excitation energies obtained from a standard MRCI using CASSCF orbitals, with an average deviation of 0.12 and 0.15 eV for the uncorrected and Davidson corrected energies, respectively. We also explore some of the computational savings that the new method affords via systematic truncation of the virtual space. Overall, it is found that the HMNO approach is a reliable method for computing MRCI excited-state energies in a fraction of the time it would take to run a standard MRCI calculation.

1. INTRODUCTION

The multireference configuration interaction (MRCI) method is useful for describing many systems, including those that involve bond breaking, multireference ground states, excited states, particularly those that are doubly excited, and conical intersections (ci).^{1–3} Despite being such versatile tool, it has some disadvantages that limit its applicability. One of them is that it is not size extensive,⁴ and approaches have been developed to correct for this.^{5–12} Another important disadvantage is its poor scaling and inherently high computational cost that limits its use to small systems. Therefore, in many cases the usage of standard multireference methods is prohibited, even if the advantages are desired. In these cases, there are two logical approaches to be explored. One can simply try to augment a single reference method in order to adapt it to the problem at hand. Alternatively, one can attempt to reduce the computational cost of the multireference method while preserving the inherent quality of the desired results.

The first approach is currently being pursued by several researchers. For example, there are attempts to improve the ability of time-dependent density functional theory (TDDFT) to describe doubly excited states¹³ or to describe conical intersections with spin-flip techniques.^{14–16} Single-reference perturbation theory approaches have also been extended to account for near degeneracies and conical intersections.^{17,18}

Although single-reference methods may be partially adapted to deal with multireference problems, multireference methods are still the natural way to address the numerous systems stated above. It would therefore be very useful to be able to perform them with reduced cost. Reduced scaling MRCI approaches have been developed before, and in some cases linear scaling MRCI has been achieved.^{11,19–23} However, implementing these

approaches may be complicated, and a simpler approach would be attractive.

A general method of reducing the computational cost of a given method is to truncate a portion of the virtual space that has little contribution to the total correlation and, thereby, reduce the number of determinants or configuration state functions (CSFs) generated. Unlike canonical molecular orbitals, natural orbitals (NOs) lend themselves very well to this idea. Natural orbitals, first introduced in 1955 by Löwdin,^{24–26} are the eigenvectors of the one-electron density matrix. They have been used extensively in the past for obtaining ground-state energies.^{27–41} Generally their usage involves using moderately correlated methods in order to generate orbitals that can be used in subsequent calculations with higher levels of theory. For example, second-order Møller–Plesset perturbation theory (MP2) can be used to generate orbitals that are later used in coupled cluster theory.

Despite the obvious advantages of using cheaply generated NOs, their usage has been mainly restricted to the ground state because it is difficult to obtain a set of well-defined virtual orbitals for excited states. For example, NOs obtained from a ground-state MP2 calculation will adequately describe the closed shell ground state and will therefore yield respectable ground-state energies if used. However, the excited states will be destabilized because the virtual orbitals are obtained from a ground-state calculation and are thus optimum for the ground state. There have only been a few studies where NOs have been applied to excited states. In 1995, average NOs from a previous MRCI calculation for several states were employed to describe

Received: November 18, 2011

Published: January 18, 2012



the neutral and ionized states of ethylene.⁴² Neese has also developed a scheme called spectroscopy oriented configuration interaction, which generates and uses average approximate NOs in an iterative procedure to describe excited states.⁴³ More recently, Krylov et al. have successfully applied frozen natural orbitals (FNOs) to cation states using the coupled cluster method.⁴⁴ The FNO method is a well-known method to generate a set of NOs for truncation. It involves freezing the occupied Hartree–Fock orbitals and diagonalizing the virtual orbital block independently.⁴⁵ A FNO approach has also been used within the Cholesky decomposition-based CASPT2 approach by selecting a minimal number of virtual orbitals from a set of approximate NOs obtained as eigenvectors of a MP2-type density matrix. These orbitals were retained in the CASPT2, and the method was used to obtain excited states.⁴⁰

Multiconfigurational self-consistent field (MCSCF) is the standard method for optimizing orbitals to be used in MRCI to study systems that have multireference character; however, it can be very expensive, especially when large active spaces are used. Furthermore, there are several other problems associated with it, including those of convergence and the existence of multiple solutions. Finally, and quite importantly, there is no easy way to truncate virtual orbitals, as the ordering of the orbitals is not necessarily associated with their importance. There have been attempts in the past to find cheaper alternatives. For example, Freed et al. developed the improved virtual orbital-complete active space configuration interaction (IVO-CASCI) method in order to replace the MCSCF procedure with a more efficient calculation that still gives reasonable results.^{46–48} It was shown that the IVO-CASCI method is able to tackle problems that single reference methods cannot address. Previous approaches have also tried to use NOs as an initial guess for MCSCF^{31,32} and even replace MCSCF orbitals completely with NOs.^{29,30,34} These latter approaches are similar to our approach in the sense that they replace the MCSCF orbitals, but they were only applied to ground-state potential energy surfaces (PESs).

The logical thing to do in order to obtain high-quality NOs for excited states would be to generate NOs using a correlated excited-state calculation, but doing so defeats the purpose of using NOs in the first place. Therefore, the main question is how to obtain NOs suitable for excited states without doing an expensive excited-state calculation, quite a catch-22. However, if this question can be answered, the benefit of using NOs will be two-fold. First, the NO generation step eliminates the requirement of running an expensive MCSCF. Second, NOs are very useful when truncating the virtual space in order to dramatically reduce the number of CSFs that the MRCI will be required to generate.

In this work we will explore the idea of utilizing NOs derived from high-multiplicity states to describe singlet excited states. The idea of using higher multiplicity states has been used in the past, especially in spin-flip techniques.^{14,15} When eliminating the MCSCF step, it is not obvious whether the NOs will be sufficient to describe excited-state character and yield energies comparable to those from a standard MCSCF calculation. We will therefore explore the performance of high-multiplicity NOs in conjunction with the MRCI method, focusing on their efficacy when applied to excited states and conical intersections. We will show that the NOs obtained from high-multiplicity states prove to be sufficient for describing many low-lying excited states for a plethora of systems.

The paper is organized as follows: The basic theory will be described in Section 2, the exact method of execution in Section 3, and its performance in various molecules in Section 4. We first address vertical excitations and the effects of virtual space truncation, then discuss how the method performs in *ci* regions.

2. THEORY

Natural orbitals obtained from a correlated *ab initio* calculation are a perfect candidate to be used in calculations in which it is desirable to truncate the virtual space. Since they are defined as the eigenvectors of the one-electron density matrix, their eigenvalues are the occupation numbers, unlike canonical molecular orbitals, which have energy eigenvalues. Truncating molecular orbitals from the top of the virtual manifold gives very irregular results, as the energy of a given orbital does not have a direct correlation to the importance it carries in describing the system. Since NOs are ordered by occupation instead of energy, they are able to be truncated via ascending contribution to the correlation of the system. In the work discussed here, NOs are obtained from a configuration interaction with single and double excitations (CISD) calculation. Traditionally, it has been popular to obtain NOs from MP2, which scales as N^5 , while CISD scales as N^6 . Abrams and Sherill found that CISD NOs perform better than MP2 NOs in certain cases when used in MRCI calculations of PESs, although the results depended on whether restricted or unrestricted MP2 was used.³⁴ Here we chose to proceed with CISD NOs, but in the future, we plan to investigate the performance of MP2 or other types of NOs as well.

The field of using NOs to study ground-state systems is quite mature, and the ability of NOs to describe the closed shell ground state has been sufficiently explored. However, as stated before, the set of NOs obtained from a ground-state singlet calculation is generally not suitable for describing excited states. In order to remedy this, one can simply perform a correlated excited-state calculation to produce higher quality orbitals, but this option is against the spirit of attempting to reduce the computational cost of the calculation.

Alternatively, one can choose to obtain orbitals from a state of higher multiplicity. A simple way to understand the motivation for this choice is by considering an excited state described as a highest occupied molecular orbital (HOMO) → lowest unoccupied molecular orbital (LUMO) excitation. A singlet and a triplet state can be formed from this configuration. Although we are interested in describing the singlet excited state, the triplet state is easier to describe since it is the ground state in the triplet manifold, and therefore, ground-state methods are more directly applicable. Natural orbitals can then be obtained from this state and used for the description of the singlet excited state. This procedure seems reasonable for a number of excited states and molecules, but it will not be adequate for all cases. For many molecules the excited states may be more complicated, or we may be interested in more than one excited state. In this case we need to find a way to extend the approach. One way to do so is by using a state with multiplicity higher than three. We have tested this idea by increasing the multiplicity from three to five, and it seems to work much better in most systems studied here. The results are of course highly dependent on the nature of the triplet or quintet state. In most cases, we simply use the ground state of a given multiplicity, but there may be instances where this is not appropriate. We will elaborate more on this in the discussion of

Table 1. Computational Details for All Calculations Done with the MRCI/HMNO Approach^a

molecule	geometry	basis set	reference	active space	symmetry	no. frozen core	excitation level
ethylene	S ₀ min	aug-cc-pVDZ	[π , π^*]	(2,2)	D _{2h}	2	2
	S ₀ –S ₁ ci	aug-cc-pVDZ		(2,2)	C ₁	2	2
formaldehyde	S ₀ min	cc-pVDZ/TZ/QZ	[π , n, π^* , σ^*]	(4,3)	C _{2v}	2	2
hexatriene	S ₀ min	cc-pVDZ	[π , π , π^* , π^*]	(6,6)	D _{2h}	6	2
benzene	S ₀ min	cc-pVDZ	[π , π , π^* , π^*]	(6,6)	D _{2h}	6	2
phenol	S ₀ min	cc-pVDZ	[π , π , π^* , π^*]	(8,7)	C _s	21	2
uracil	S ₀ min	cc-pVDZ	[n, π , π^* , π^*]	(12,9)	C _s	23	1
	S ₁ –S ₂ ci	cc-pVDZ		(12,9)	C ₁	23	1

^aThe reference information applies to the generation step, while the active space, no. of frozen core, and excitation level apply to the MRCI step. The generation step is always done with single and double excitations from all occupied orbitals with only the core 1s orbitals frozen.

uracil below. In short, the rationale is to create a ground-state calculation that will generate orbitals that are similar to those used by the excited singlet state of the same system.

The applicability of the method also depends on the nature of the excited states, as one would expect. If we are interested in describing several excited states that have very different character, the choice of the high-multiplicity state that creates the NOs may be more complicated. For example, if we are interested in describing both valence and Rydberg excited states in a molecule, we have to make sure we use the appropriate NOs. In this case the high-multiplicity state should include both Rydberg and valence orbitals in order to be able to describe both types of states. Alternatively, it may be necessary to use different NOs for the different types of excited states. This idea is also used when MCSCF orbitals are used, where orbitals from a state-specific MCSCF are used for each state separately. Here we only focus on valence-type excited states, although we do consider several states in a molecule, often of different character. As will be seen below, the results are very encouraging even when many states are calculated, so we are optimistic that this approach can be generalized to more diverse states.

3. COMPUTATIONAL DETAILS

The systems chosen to test the theory are ethylene, formaldehyde, hexatriene, phenol, benzene, and uracil. Their structures are given in Supporting Information. Equilibrium geometries were obtained by optimization at the MP2/cc-pVDZ level. The equilibrium geometry and the ci geometry for uracil were obtained from calculations previously done by one of us⁴⁹ at the MRCI1/cc-pVDZ level. (MRCI1 denotes MRCI with single excitations from the complete active space.) The ethylene ci and equilibrium geometries were taken from Lischka's study on ethylene using MRCI.⁵⁰

The high-multiplicity natural orbital (HMNO) approach used in this paper entails a two step process, which we will call the generation step and the production step. First, the NOs were generated at the single reference CISD level using a high-multiplicity reference state (generation step). The starting orbitals were obtained from a restricted open shell Hartree–Fock of the same high multiplicity. The NOs were then used as an ansatz for an MRCI calculation, where the singlet ground and excited states were calculated with various degrees of virtual space truncation (production step). Aside from the choice of the high-multiplicity reference, the NO generation step is always the same; it is always a single reference CISD with only the 1s core orbitals on the heavy atoms frozen. Therefore, a MRCI/HMNO calculation is characterized by two things: the set of orbitals that define the high-multiplicity reference and the

parameters of the subsequent MRCI calculation. Four orbitals (two) define the quintet (triplet) high-multiplicity reference for the CISD generating step which are the singly occupied orbitals, while all other occupied orbitals are doubly occupied. This orbital set will be designated [x_1 , x_2 , x_3 , x_4] ([x_1 , x_2]), where x_i designates orbital character.

Pertinent information for the standard calculations performed can be found in Table 1. The same basis set and symmetry were used in both the generating step and the production step. The active space, frozen core, and excitation levels are given for the production step only. In all of our calculations a complete active space was used for the reference space in MRCI. The choice of the active space in the MRCI step depends on the molecule and the excited states we want to explore. For the conjugated molecules, all the π and π^* orbitals were included in the active space. For formaldehyde, the lone pair on oxygen is also included since this is involved in the first excited state. Similarly, for uracil the lone pair on one of the oxygen atoms (O7, see structure and labeling of atoms in Supporting Information) is included since it contributes to the first excited S₁ state. The lone pair on the second oxygen (O8) is not included since it has been found not to affect the excitation energies of the first two excited states.⁴⁹

For ethylene, the aug-cc-pVDZ basis set was used so that our results could be compared with previous MRCI calculations.⁵⁰ An investigation of the basis set effect was carried out for formaldehyde, where the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets were used. For the remaining molecules the smaller cc-pVDZ basis set was used.

In addition to the calculations laid out in Table 1, a few truncation studies were also performed for the uracil, formaldehyde, and ethylene molecules. These involved studying the effect of virtual space truncation on the vertical excitation energies of these molecules. Truncation calculations for ethylene and formaldehyde were performed with the cc-pVQZ basis set, while cc-pVDZ was used for uracil. All other aspects of the calculations remained the same. Virtual orbitals were truncated in accordance to their occupational contribution to the virtual space, as demonstrated by Krylov et al.,⁴⁴ which will be described below.

Standard MRCI calculations were performed with state-averaged complete active space self-consistent field (CASSCF) molecular orbitals for comparison to the analogous MRCI/HMNO calculations. The CASSCF calculations were set up with the same active space as their MRCI counterparts.

A slightly modified COLUMBUS computational package was used for all MRCI calculations.^{51–54} In some cases, the MOLCAS module seward⁵⁵ was used to generate integrals instead of the COLUMBUS module Dalton through the

Table 2. Vertical Excitation Energies (in eV) Taken from MRCI/HMNO and MRCI/CASSCF and the Difference (Δ) between the Two Methods^a

	symmetry	excited-state configuration	MRCI/CASSCF	MRCI/HMNO	Δ
formaldehyde	$^1A''$	$n^1\pi^{*1}$	4.15 (3.91)	4.18 (3.31)	+0.03 (−0.60)
ethylene	$^1B_{1u}$	$\pi^1\pi^{*1}$	8.08 (7.88)	8.56 (7.85)	+0.48 (−0.03)
	1A_g	$\pi^0\pi^{*2}$	13.12 (12.16)	13.16 (11.41)	−0.04 (−0.75)
hexatriene	1B_u	$\pi^1\pi^{*1}$	5.43 (5.40)	5.43 (5.41)	0.00 (+0.01)
	1A_g	$\pi^0\pi^{*2}$	6.26 (5.80)	6.35 (5.81)	+0.09 (+0.01)
	1B_u	$\pi^1\pi^{*1}$	6.60 (6.57)	6.89 (6.67)	+0.29 (+0.10)
benzene	$^1B_{2u}$	$e_g^1e_u^{*1}$	5.00 (5.04)	4.97 (5.03)	−0.03 (−0.01)
	$^1B_{1u}$	$e_g^1e_u^{*1}$	7.11 (6.82)	7.15 (6.82)	+0.04 (0.00)
	$^1E_{1u}$	$e_g^1e_u^{*1}$	8.22 (7.73)	8.24 (7.71)	+0.02 (−0.02)
phenol	$^1A'$	$\pi^1\pi^{*1}$	4.97 (4.93)	4.92 (4.93)	−0.05 (0.00)
uracil	$^1A''$	$n^1\pi^{*1}$	5.44 (5.41)	5.19 (5.27)	−0.25 (−0.14)
	$^1A'$	$\pi^1\pi^{*1}$	6.24 (6.13)	6.41 (6.24)	+0.17 (+0.11)

^aThe corresponding Davidson corrected energies are given in parentheses.

COLUMBUS–MOLCAS interface. All MP2 optimizations were performed using GAMESS.^{56,57} Molecular orbitals were viewed using MOLDEN⁵⁸ and MacMolPlt⁵⁹ for COLUMBUS and GAMESS calculations, respectively.

4. RESULTS AND DISCUSSION

4.1. Vertical Excitation Energies Using HMNO. The choice of orbitals, especially active orbitals, is very important in MRCI, so it is not clear that by substituting MCSCF orbitals with HMNO, the MRCI excitation energies remain accurate. Therefore it is important to compare the MRCI/HMNO results to the standard method. Here we will compare how well the MRCI/HMNO approach compares to the standard MRCI using CASSCF orbitals, both with and without the Davidson correction. The Davidson correction is the simplest way to account for the size-extensivity problem in MRCI.^{5,6} Table 2 shows the vertical excitation energies obtained using MRCI/HMNO for the list of molecules we tested. The MRCI energies using the CASSCF orbitals are also shown for comparison as well as the difference in the energy between the two methods.

Our list consists mainly of typical closed shell organic molecules with double bonds. The excited states that we consider involve mainly π orbitals and in some cases lone pairs on oxygen. Conjugated polyenes belong to a class of molecules that is fairly difficult to describe because of the ubiquitous doubly excited state that each of them possesses. Doubly excited states cannot be described by the common single reference methods available, so a multireference method must be used. Ethylene is the simplest conjugated system which is studied here. Natural orbitals derived from a triplet reference are sufficient for ethylene, which has a low-lying HOMO LUMO $\pi^1\pi^{*1}$ transition that is described by the same orbitals as the $\pi^1\pi^{*1}$ triplet state. In fact, attempting to use a quintet state yields energies that are considerably worse than those obtained from the triplet state. The difference between the MRCI/CASSCF and the MRCI/HMNO approach, however, is the largest we see in our calculations, 0.48 eV. This state is difficult to describe in general since correlation effects and mixing with Rydberg character are important. The doubly excited state in ethylene has two electrons in the LUMO orbital. The excitation energy for this state is described very well with HMNO, much better than the $\pi^1\pi^{*1}$ state, with only 0.03 eV difference from the analogous MRCI/CASSCF energy.

More complicated systems require more active orbitals to describe the low lying excited states, and therefore a quintet state must be used to extend occupation to these orbitals. For the rest of the molecules studied, a quintet reference was used. We studied two conjugated molecules with six carbons, *trans*-hexatriene and benzene. Both of these molecules have six π orbitals that are included in the active space. The quintet state that is used for the NOs is built from four of these π orbitals. Hexatriene, like ethylene, has a doubly excited S_2 state, which the MRCI/HMNO approach reliably replicates. Benzene excited states are formed by exciting an electron from the doubly degenerate e_g orbitals to the doubly excited e_u virtual orbitals.⁶⁰ This leads to four excited states, two of which are degenerate. All of these states are described remarkably well with the MRCI/HMNO approach. Including a substituent in the benzene ring, as in phenol, did not create any problems. Using again four π orbitals in the reference and seven π orbitals in the active space (six from the ring and one from the OH substituent) gave very good results.

Formaldehyde is included in this study since it is the simplest molecule which has a $n^1\pi^{*1}$ excited state that is lower in energy than the $\pi^1\pi^{*1}$ state. Several basis sets were used, the effect of which will be discussed later. The values reported in Table 2 are obtained with the cc-pVQZ basis set. In order to obtain reasonable results, the high-multiplicity state must include the lone pair orbital. Using a triplet $n^1\pi^{*1}$ state for the NOs gave the excitation energy of 3.60 eV for the singlet $n^1\pi^{*1}$ state. This value is rather low compared to the analogous MRCI/CASSCF value which is 4.11 eV. The NOs from the triplet state overstabilize the excited state compared to the ground state, resulting in a low excitation energy. When NOs from a quintet state are used, the agreement between MRCI/HMNO and MRCI/CASSCF is really good. This is because, in this case, the ground and excited state energies are destabilized by the same amount when compared to the MRCI/CASSCF energies, so the overall excitation energy is not affected much. It should be noted that the error in formaldehyde when the triplet state orbitals are used is similar to the error in ethylene with NOs from the triplet state, but in ethylene the quintet NOs does not perform better.

It is important to ascertain the effectiveness of the method on more complicated molecules, so we chose uracil as an additional test molecule. This molecule is an important biological chromophore, and extensive theoretical work has

been done for its excited states. Obtaining accurate excited states for uracil is actually very difficult, such that an extended basis set and inclusion of excitations beyond double is essential.⁶¹ Here we will only test the ability of MRCI/HMNO to match the equivalent MRCI/CASSCF and not to obtain the experimental excitation energy. Unlike for the other molecules, the quintet for uracil must be carefully chosen. The first two excited states have $n^1\pi^{*1}$ and $\pi^1\pi^{*1}$ character, so both the n and π orbitals are essential and should be included in the reference quintet. There are two quintet states $[\pi, n, \pi^*, \pi^*]$ or $[\pi, \pi, \pi^*, \pi^*]$ which have CISD energies of -413.09841 and -413.07394 au, respectively. The desirable quintet is then the ground state in the quintet manifold, but the $[\pi, \pi, \pi^*, \pi^*]$ is not much higher in energy so one has to be careful. If the desirable quintet was not the ground state, then we would have to choose an excited quintet state as the reference, complicating the situation. Even in the present case, it is sometimes difficult to converge to the right quintet. In this case, the calculations can be done using symmetry, and since the two states have different overall symmetry, it is easy to distinguish them. As expected, MRCI energies using the quintet containing the lone pair orbitals were much more favorable than those derived from the purely π quintet. Actually the energy of the $n^1\pi^{*1}$ state was several eV too high when NOs from the $[\pi, \pi, \pi^*, \pi^*]$ quintet were used. It is therefore important to note that the excitation energies can vary wildly when using different high-multiplicity references.

In general, the comparison of vertical excitation energies between the MRCI/CASSCF energies and the MRCI/HMNO energies shows that MRCI/HMNO does very well in reproducing vertical excitation energies. Although the average absolute difference between the two methods is 0.12 eV, most of the energies differ by much a smaller amount (0.00–0.04 eV), with a few stark outliers. Interestingly, the largest deviation of 0.48 eV is shown in ethylene, the simplest molecule studied. The third excited state in hexatriene also shows a large deviation of 0.29 eV and, finally, the uracil excited states show deviations of 0.25 and 0.17 eV for the first and second excited states, respectively. In uracil, since the active space in the MRCI is large, including nine orbitals, one would expect that the HMNOs will perform less accurately than for the smaller molecules. Quite interestingly though, for the smallest molecules studied, it was also not as straightforward to reproduce the MRCI/CASSCF energies with MRCI/HMNO. This is probably related to the fact that in the limit of a full CI, the energies are independent of the choice of the orbitals, and here, being on the opposite extreme, the very small active space does not allow for enough flexibility.

The Davidson correction provides a very easy way to include a correction for the size extensivity problem in truncated configuration interactions methods. This correction only depends on the weight of the reference wave functions of the MRCI expansion.^{5,6} It is seen that including the correction in most cases does not change the difference between MRCI/HMNO and MRCI/CASSCF much, with an average absolute difference between the two methods of 0.15 eV. Actually, in most cases, the Davidson-corrected MRCI/HMNO energies have better agreement with their MRCI/CASSCF counterparts, except for two cases where the Davidson correction makes the agreement a lot worse.

Although the generation of NOs via CISD replaces the CASSCF orbital optimization step, there may still be some effort involved in creating the correct reference. So, this method

is not necessarily a black box approach. It is trivial to obtain NOs that adequately describe the ground state, as there is no uncertainty in which reference to use. However, since the objective is to obtain NOs that can describe excited states, the high-multiplicity reference must be chosen carefully. One question is whether a triplet or a quintet state should be used. In our tests, only ethylene was described better with NOs from a triplet state rather than from the quintet state. Since most molecules are more complicated than ethylene, it is probably safer to use a quintet rather than a triplet state. The ground state within the quintet manifold is appropriate most of the times. If there is symmetry in the system, one may utilize the symmetry restrictions to obtain orbitals appropriate for a specific symmetry. Once the appropriate reference is found, the HMNO approach performs outstandingly. The energies obtained using high-multiplicity NOs compare quite well to their CASSCF counterparts, only varying by an average of 0.12 eV as seen above; this error is smaller than that of the MRCI itself. The computational savings are also tremendous. The CISD step is much simpler than the CASSCF step to obtain orbitals, but the most important savings can be obtained by virtual space truncation, as will be seen in the next few sections.

4.2. Virtual Space Truncation. The most important benefit of using NOs is the ability to systematically truncate the virtual space in a subsequent correlated calculation without loss of accuracy. Although this benefit is well-known for ground-state calculations, here we explore how well it applies to excited states. The savings allowed by virtual space truncation are illustrated in Figure 1, which shows the dependence of the

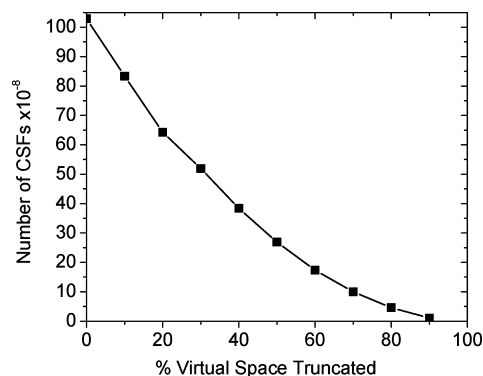


Figure 1. Correlation between number of CSFs generated and number of orbitals truncated. Taken from MRCI calculation with single and double excitations from all occupied and active orbitals (except the 1s core) for uracil with cc-pVDZ basis set and (12,9) active space.

number of CSFs generated on the number of virtual orbitals truncated for a large MRCI calculation including single and double excitations. In order for truncation methods to be used effectively, the virtual space must be truncated in a standard and logical fashion. The simplest way to do this is by simply truncating an arbitrary number of orbitals, for example, truncating 10 orbitals at a time and observing the impact on the energy of the excited states. This is not very useful as 10 is an arbitrary number and has correlation neither to the total number of virtual orbitals nor to the system of interest. Truncating 10 orbitals in a virtual manifold of 50 total orbitals is completely different from truncating 10 orbitals in a manifold of a 100. In order to better tailor the choice of which virtual orbitals to truncate to the system of interest, orbitals can be selected in terms of percentage of the complete virtual space.

This is a standardized scale in which to compare different systems as well as the effect of different basis sets on the same system. Still, simply truncating orbitals based on their number alone is not a very smart choice, as this weights the importance of each virtual orbital equally, whereas all orbitals are not created equally. A more sophisticated way has been proposed by Krylov et al.⁴⁴ which uses the occupation threshold (OCCT) cutoff criterion. The OCCT criterion involves virtual space truncation by selecting a percentage of occupation that is to be preserved and then truncating all of the orbitals above this threshold. Note that the virtual space here is defined by the set of unoccupied orbitals in the orbital-generating CISD calculation.

The correlation between the percentage of occupation preserved and the percentage of the total virtual space that is truncated is shown in Figure 2 for formaldehyde, ethylene, and

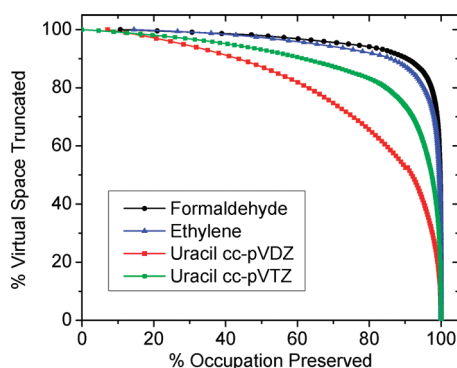


Figure 2. Percent of virtual space truncated as a function of occupation preserved in the virtual space. Data for ethylene and formaldehyde molecules were obtained from calculations done with the cc-pVQZ basis set, while uracil data are shown for two basis sets, cc-pVDZ and cc-pVTZ.

uracil. It is apparent that the dependence is not the same for all molecules, but in general, a large number of virtual orbitals can be truncated while still retaining a high amount of occupation and ostensibly a low amount of error. The curvature shows the distribution of the occupancy. A gradual decline is less desirable because it means that occupation is being lost gradually with the truncation of orbitals, whereas a flat curve followed by a steep decline denotes almost no occupation loss until a certain threshold. The slope is worst for uracil using the cc-pVDZ basis set but very favorable for ethylene and formaldehyde with uracil/cc-pVTZ in between the two cases. The effect of the basis set is very obvious in uracil, where going from the cc-pVDZ to the cc-pVTZ basis set the % occupation preserved is much higher for a given truncation. The benefits of virtual space truncation are more important for larger basis sets as we will see in the next section.

Figure 3 demonstrates the general trend the excitation energies follow when orbitals are truncated by occupation number for several molecules. The figure shows the difference of the computed energy with a truncated virtual space compared to the original energy with the full virtual space. The most important thing to note is that the energy changes are on a very small scale, even as many orbitals are truncated. For example, in formaldehyde, preserving 99.9% occupation corresponds to truncating 30% of the virtual space, yet the energy only changes by 0.005 eV. This small change in energy is negligible compared to many other errors in this calculation.

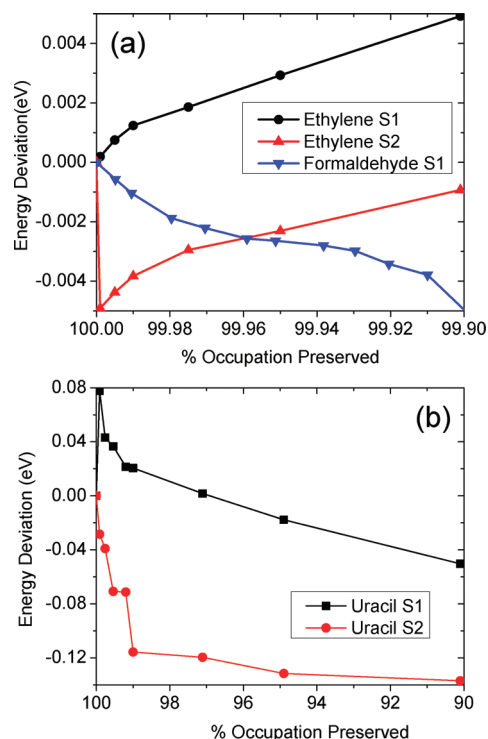


Figure 3. Energy deviation in eV as virtual orbitals are truncated by occupation contribution. The zero is the energy for each molecule when there is no truncation of the virtual space. (a) Results for ethylene and formaldehyde for % occupation preserved between 100 and 99.9 using the cc-pVQZ basis set. (b) Results for uracil for % occupation preserved between 100 and 90 using the cc-pVDZ basis set.

Truncating 81% of the virtual space preserves 95% occupation and only changes the excitation energy by 0.027 eV. Even for the uracil molecule, for which we use a smaller basis set, one can keep 90% of the occupation by truncating 52% of the virtuals, changing the excitation energy by only 0.04 eV for the S_1 state, while S_2 changes by 0.14 eV.

As opposed to truncation by percentage of orbitals, we expected that using the OCCT truncation criterion allows for a predictable energy trend. It has been found in the past that the energy change is largely linear in nature⁴⁴ when truncating orbitals by occupation. However, given the erratic behavior of the energy in some systems using HMNO, it does not seem that extrapolation is currently an option. It is however important to realize that one can truncate a large percentage of the virtual space and still the scale of the error in most cases is on the order of 0.01 eV. Even without extrapolation, truncation works extremely well.

4.3. Basis Set Effect. Although we have already demonstrated that successive truncation of the virtual space has little impact on excitation energies, we have not discussed in detail comparisons between calculations done with different basis sets. Figure 2 implies an important effect. The choice of which basis set to use in any given calculation is an important one. The choice for the HMNO calculation is no different, but it is made easier by studying the effect of using different basis sets. Logistical information for the formaldehyde basis set study is shown in Table 3. Three different basis sets were compared, the cc-pVDZ, cc-pVTZ and cc-pVQZ. In each case calculations were done without truncating any virtual orbitals and with truncating enough orbitals so that up to 99% of the occupation

Table 3. Comparisons of Recorded Run Times for MRCI on Formaldehyde Using Different Basis Sets and Truncation of Different Numbers of Virtual Orbitals^a

	% occ. preserved	no. orbs. truncated	# CSFs	wall time	S ₁ energy (eV)
cc- pVDZ	100	0	103094	0:00:53	4.36
	99.75	2	89533	0:00:47	4.34
	99	4	76928	0:00:29	4.33
cc- pVTZ	100	0	752819	0:11:15	4.26
	99.75	18	450086	0:04:31	4.25
	99.5	24	366383	0:03:07	4.25
	99	30	291284	0:02:06	4.25
cc- pVQZ	100	0	3111836	2:17:41	4.18
	99.75	65	1109836	0:21:41	4.18
	99.5	81	771908	0:12:34	4.18
	99	97	495164	0:04:29	4.18

^aWall time is given for the MRCI step in hours. Vertical excitation energies are given in eV with respect to the ground state at equilibrium geometry.

is preserved. The first striking difference is that, for cc-pVDZ, removing 1% of the occupation corresponds to truncating only 4 virtual orbitals, while, for cc-pVQZ, it corresponds to truncating 97 virtual orbitals. The savings obviously are really important for the larger basis set. This is evident in the number of CSFs and in the wall time for the calculations. In all cases the excitation energy stays almost unaffected by virtual space truncation. On the other hand, the excitation depends a lot on the basis set, so that calculations carried out with larger basis sets with a substantial portion of the virtual space truncated still have much better excitation energies than those done with smaller basis sets with no truncated virtual orbitals. For example, a calculation can be run with a cc-pVTZ basis set with no truncated virtual orbitals in 11 min, yielding 4.26 eV for S₁. Alternatively, the same calculation can be run with cc-pVQZ with 99% occupation preserved (97 of 159 virtual orbitals truncated) in 4 min, yielding 4.18 eV for S₁. The experimental value for the excitation energy in formaldehyde is 4.07 eV.⁶² If no orbitals are truncated in the cc-pVQZ, the calculation takes 2 h and 17 min, instead of 4 min (34 times speed-up!), yet the result is practically the same. Judging by these results, it is a prudent decision to use larger basis sets when possible, even when it is required that a larger portion of the virtual space be truncated to keep the calculation to a manageable size.

4.4. Conical Intersections. Multireference methods are particularly useful in describing nonadiabatic regions of the PESs of molecules and conical intersections. So it is important to ascertain whether the current approach is able to describe the PESs at these regions as well as at the Franck–Condon region. We will examine the performance of MRCI/HMNO in regions of conical intersections for two molecules here. As this is only a cursory examination, more work will have to be done to examine the ability of this method to address conical intersections. There is also a related, but more general, question of how well this approach can predict smooth PESs, but we will not examine it in detail here. Previous work has tested NO approaches and found that they can indeed predict smooth PESs for dissociation of molecules in the ground state.³⁴ The smoothness of excited state PESs with the current approach will have to be tested in more detail.

We will focus on two different types of conical intersections. The simplest example of a ci is found in ethylene, which has a ci

between the ground state and a $\pi^1\pi^{*1}$ excited state.^{50,63} The geometry and MRCI/CASSCF energies for the ci in ethylene were obtained by Lischka and co-workers.⁵⁰ This ci is formed when the double bond twists, forcing the $\pi^1\pi^{*1}$ state to stabilize and the ground state to destabilize until they eventually cross. The two orbitals that are very important in this ci are the initial two π orbitals which, having been twisted, now look more like atomic p orbitals on each carbon. The energies for S₀ and S₁ at this geometry are shown in Table 4 as well as the energies

Table 4. Conical Intersection Energies for the S₀/S₁ ci in Ethylene and the S₁/S₂ ci in Uracil Calculated with MRCI/HMNO and Compared To MRCI/CASSCF Values^a

	MRCI/HMNO	MRCI/CASSCF
Ethylene		
S ₀	4.43	4.83 ^b
S ₁	5.01	4.83 ^b
S ₂	10.01	9.75 ^b
Uracil		
S ₀	2.15	2.15 ^c
S ₁	5.70	5.37 ^c
S ₂	6.22	5.37 ^c

^aVertical excitation energies are given in eV with respect to the S₀ state at the equilibrium geometry. ^bTaken from ref 50. ^cTaken from ref 49.

obtained by Lischka and co-workers using MRCI/CASSCF. NOs were derived from the triplet state of ethylene as it was done for the vertical excitations. The excited-state energies given by MRCI/HMNO do fairly well in describing the ci. The average energy of the first two states is 4.72 eV, very similar to the degenerate energy (4.83 eV) given by the standard MRCI. The splitting however is 0.6 eV. Although it is best when it is small, a splitting is always expected when a different method is used for the energies than the method used for the optimization. This is always observed in the most common protocol for studying conical intersections where a CASSCF wave function is used to locate the ci and CASPT2 is used to obtain the energies at these geometries. Optimization at the same level of theory would find the geometry where the splitting is zero for the new approach. Currently we do not have the ability to calculate gradients with the HMNO approach, so we cannot locate the ci at this level and, therefore, cannot know how that geometry would differ from the current geometry. The splitting however is within the limits of what we have seen when using CASSCF gradients and then refining with correlated methods, so the current finding indicates that the method is promising.

The second example we studied is a ci between two excited states. This is a ci between the S₁ and S₂ states in uracil, where S₁ has $n^1\pi^{*1}$ character and S₂ has $\pi^1\pi^{*1}$ character.^{49,64,65} The geometry at this ci is not nearly as distorted from the equilibrium geometry as the previous ci presented. By using similar quintet reference as at the equilibrium geometry, we obtain energies that are reasonable and comparable with those obtained from the standard MRCI. The splitting is again approximately 0.5 eV, while the average energy is 5.96 eV, 0.6 eV higher than the MRCI/CASSCF degenerate energy. The error here is higher than in ethylene, but this may be reasonable since uracil is a much more complicated molecule than ethylene.

Overall, the deviation of the MRCI/HMNO energies compared to the MRCI/CASSCF energies at ci geometries is

somewhat higher than the deviation for vertical excitation energies observed at equilibrium geometries. Nevertheless, the results are promising and within deviations seen in usual protocols followed for CI searches. The idea of truncating the virtual space also carries to applications on conical intersections. We tested different truncations and found that the energy change from truncating a large percentage of the virtual space is also negligible here.

5. CONCLUSIONS

We have tested a new approach to obtain excited-state energies for systems where a multireference wave function is required. In this approach we use an MRCI calculation where the orbitals are not obtained from the usual CASSCF procedure, but they are the NOs obtained from a single reference CISD calculation on a triplet or quintet reference state. For most systems, the vertical excitation energies for the singlet states obtained from this MRCI/HMNO approach are very similar to the analogous MRCI calculation done with CASSCF orbitals; usually with a difference of less than 0.1 eV. Despite the relatively high accuracy relative to its counterpart, the MRCI/HMNO approach is much faster computationally, as it is not necessary to run the CASSCF step to generate the orbitals, and it is possible to truncate a large percentage of the virtual space, making the calculation even faster without sacrificing significant accuracy. We have also found that the benefits are greater when a large basis set is used. It has been generally found that using the largest basis set possible with at least 99% occupation preserved gives excellent results for the computational effort that would otherwise be associated with a smaller basis set. Although with the HMNO approach we can bypass the CASSCF step with all its choices, the method by no means is a black box approach. There are still many decisions to be made about the active space for the MRCI, and one has to be careful about the choice of the multiplet state used for the NOs. The ability of the approach to tackle the problem of complicated PESs and conical intersections seems promising, pending further study.

■ ASSOCIATED CONTENT

Supporting Information

A figure of the structures of all molecules studied here. This information is available free of charge via the Internet at <http://pubs.acs.org>

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: smatsika@temple.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge support from the National Science Foundation under grant number CHE-0911474.

■ REFERENCES

- (1) Shavitt, I. Modern Theoretical Chemistry. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum Press: New York, 1977; Vol. 4; pp 189–275.
- (2) *Recent advances in multireference methods*; Hirao, K., Ed.; World Scientific: River Edge, NJ, 1999.
- (3) Paldus, J. *Methods of Computational Molecular Physics*; Plenum Press: New York, 1992; p 99.
- (4) Bartlett, R. J. *J. Phys. Chem.* **1989**, 93, 1697–1708.
- (5) Langhoff, S. R.; Davidson, E. R. *Int. J. Quantum Chem.* **1974**, 8, 61.
- (6) Duch, W.; Dierksen, G. H. F. *J. Chem. Phys.* **1994**, 101, 3018.
- (7) Gdanitz, R.; Ahlrichs, R. *Chem. Phys. Lett.* **1988**, 143, 413.
- (8) Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, 214, 418–488.
- (9) Szalay, P. G.; Bartlett, R. J. *J. Chem. Phys.* **1995**, 103, 3600.
- (10) Füsti-Molnár, L.; Szalay, P. G. *J. Phys. Chem.* **1996**, 100, 6288–6297.
- (11) Venkatnathan, A.; Szilva, A. B.; Walter, D.; Gdanitz, R. J.; Carter, E. A. *J. Chem. Phys.* **2004**, 120, 1693.
- (12) Amor, N. B.; Maynau, D.; Malrieu, J.-P.; Monari, A. *J. Chem. Phys.* **2008**, 129, 064112.
- (13) Maitra, N.; Zhang, F.; Cave, R.; Burke, K. *J. Chem. Phys.* **2004**, 120, 5932–5937.
- (14) Krylov, A. I. *Chem. Phys. Lett.* **2001**, 338, 375.
- (15) Krylov, A. I. *Acc. Chem. Res.* **2006**, 39, 83–91.
- (16) Minezawa, N.; Gordon, M. S. *J. Chem. Phys.* **2009**, 113, 12749–12753.
- (17) Head-Gordon, M.; Oumi, M.; Maurice, D. *Mol. Phys.* **1999**, 96, 593.
- (18) Laikov, D.; Matsika, S. *Chem. Phys. Lett.* **2007**, 448, 132–137.
- (19) Angeli, C.; Calzado, C.; Cimiraglia, R.; Evangelisti, S.; Guihery, N.; Leininger, T.; Malrieu, J.; Maynau, D.; Ruiz, J.; Sparta, M. *Mol. Phys.* **2003**, 101, 1389–1398.
- (20) Amor, N. B.; Bessac, F.; Hoyau, S.; Maynau, D. *J. Chem. Phys.* **2011**, 135, 014101.
- (21) Walter, D.; Szilva, A. B.; Niedfeldt, K.; Carter, E. A. *J. Chem. Phys.* **2002**, 117, 1982.
- (22) Walter, D.; Venkatnathan, A.; Carter, E. A. *J. Chem. Phys.* **2003**, 118, 8127–8139.
- (23) Chwee, T. S.; Szilva, A. B.; Lindh, R.; Carter, E. A. *J. Chem. Phys.* **2008**, 128, 224106.
- (24) Löwdin, P. *Phys. Rev.* **1955**, 97, 1474–1489.
- (25) McWeeny, R. *Rev. Mod. Phys.* **1960**, 32, 335–369.
- (26) Davidson, E. *Rev. Mod. Phys.* **1972**, 44, 451–464.
- (27) Thunemann, K. H.; Röhmelt, J.; Peyerimhoff, S. D.; Buenker, R. J. *Int. J. Quantum Chem.* **1977**, 11, 743–752.
- (28) Cave, R. J.; Xantheas, S. S.; Feller, D. *Theor. Chim. Acta* **1992**, 83, 31.
- (29) Bofill, J. M.; Pulay, P. *J. Chem. Phys.* **1989**, 90, 3637–3646.
- (30) Grev, R. S.; Schaefer, H. F. III. *J. Chem. Phys.* **1992**, 96, 6850–6856.
- (31) Pulay, P.; Hamilton, T. P. *J. Chem. Phys.* **1988**, 88, 4926.
- (32) Jorgen, H.; Jensen, A.; Jorgensen, P.; Ågren, H.; Olsen, J. *J. Chem. Phys.* **1988**, 88, 3834.
- (33) Gordon, M. S.; Schmidt, M. W.; Chaban, G. M.; Glaesemann, K. R.; Stevens, W. J.; Gonzalez, C. *J. Chem. Phys.* **1999**, 110, 4199.
- (34) Abrams, M. L.; Sherrill, C. D. *Chem. Phys. Lett.* **2004**, 395, 227–232.
- (35) Sherrill, C. D.; Schaefer, H. F. III. *Adv. Quantum Chem.* **1999**, 34, 143.
- (36) Ivanic, J.; Ruedenberg, K. *Theor. Chim. Acta* **2002**, 107, 220.
- (37) Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2004**, 121, 10852.
- (38) Taube, A. G.; Bartlett, R. J. *Collect. Czech. Chem. Commun.* **2005**, 70, 837–850.
- (39) Taube, A. G.; Bartlett, R. J. *J. Chem. Phys.* **2008**, 128, 164101.
- (40) Aquilante, F.; Todorova, T. K.; Gagliardi, L.; Pedersen, T. B.; Roos, B. O. *J. Chem. Phys.* **2009**, 131, 034113.
- (41) Klopper, W.; Noga, J.; Koch, H.; Helgaker, T. *Theor. Chem. Acc.* **1997**, 97, 164–176.
- (42) Desjardins, S. J.; Bawagan, A. D. O.; Liu, Z. F.; Tan, K. H.; Wang, Y.; Davidson, E. R. *J. Chem. Phys.* **1995**, 102, 6385.
- (43) Neese, F. *J. Chem. Phys.* **2003**, 119, 9428.
- (44) Landau, A.; Khistyayev, K.; Dolgikh, S.; Krylov, A. I. *J. Chem. Phys.* **2010**, 132, 014109.
- (45) Barr, T. L.; Davidson, E. R. *Phys. Rev. A* **1970**, 1, 644–658.
- (46) Potts, D. M.; Taylor, C. M.; Chaudhuri, R. K.; Freed, K. F. *J. Chem. Phys.* **2001**, 114, 2592.

- (47) Chaudhuri, R. K.; Chattopadhyay, S.; Mahapatra, U. S.; Freed, K. F. *J. Chem. Phys.* **2010**, *132*, 034105.
- (48) Chattopadhyay, S.; Chaudhuri, R. K.; Freed, K. F. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7514–7523.
- (49) Matsika, S. *J. Phys. Chem. A* **2004**, *108*, 7584.
- (50) Barbatti, M.; Paier, J.; Lischka, H. *J. Chem. Phys.* **2004**, *121*, 11614.
- (51) Lischka, H.; Shepard, R.; Brown, F. B.; Shavitt, I. *Int. J. Quantum Chem.* **1981**, *15*, 91–100.
- (52) Shepard, R.; Shavitt, I.; Pitzer, R. M.; Comeau, D. C.; Pepper, M.; Lischka, H.; Szalay, P. G.; Ahlrichs, R.; Brown, F. B.; Zhao, J. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1988**, *22*, 149.
- (53) Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Müller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; Zhang, Z. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664–673.
- (54) Lischka, H. et al. *COLUMBUS, an ab initio electronic structure program*, release 5.9.2; University of Vienna: Vienna, Austria, 2008.
- (55) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222.
- (56) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
- (57) Gordon, M. S.; Schmidt, M. W. *Theory and Applications of Computational Chemistry: the first forty years*; Elsevier: Amsterdam, The Netherlands, 2005; pp 1167–1189.
- (58) Schaftenaar, G.; Noordik, J. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 123.
- (59) Bode, B. M.; Gordon, M. S. *J. Mol. Graphics Modell.* **1998**, *16*, 133–138.
- (60) Hashimoto, T.; Nakano, H.; Hirao, K. *J. Chem. Phys.* **1996**, *104*, 6244–6258.
- (61) Epifanovsky, E.; Kowalski, K.; Fan, P. D.; Valiev, M.; Matsika, S.; Krylov, A. I. *J. Phys. Chem. A* **2008**, *112*, 9983–9992.
- (62) Robin, M. *Higher Excited States of Polyatomic Molecules*; Academic Press: New York, 1985; p 285.
- (63) Levine, B. G.; Martinez, T. J. *Annu. Rev. Phys. Chem.* **2007**, *58*, 613–634.
- (64) Barbatti, M.; Aquino, A. J. A.; Szymczak, J. J.; Nachtigallova, D.; Hobza, P.; Lischka, H. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21453–21458.
- (65) Matsika, S.; Krause, P. *Annu. Rev. Phys. Chem.* **2011**, *62*, 621–643.