This is an open access article published under a Creative Commons Attribution (CC-BY) License, which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

Cite This: J. Phys. Chem. A XXXX, XXX, XXX−XXX

pubs.acs.org/JPCA

# Molecular Vertical Excitation Energies Studied with First-Order RASSCF (RAS[1,1]): Balancing Covalent and Ionic Excited States

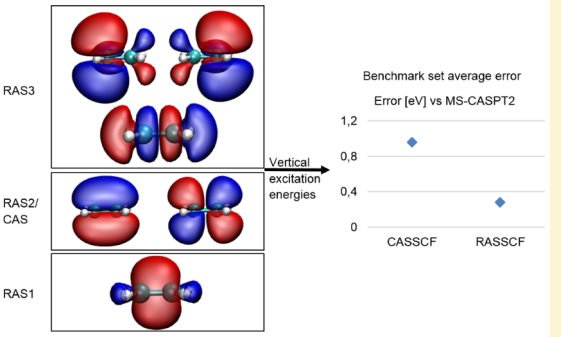
*Published as part of The Journal of Physical Chemistry virtual special issue "Leo Radom Festschrift".*

Thierry Tran, Javier Segarra-Martí, Michael J. Bearpark, and Michael A. Robb*

Department of Chemistry, Molecular Sciences Research Hub, Imperial College London, White City Campus, 80 Wood Lane, London W12 0BZ, United Kingdom

Supporting Information

**ABSTRACT:** RASSCF calculations of vertical excitation energies were carried out on a benchmark set of 19 organic molecules studied by Thiel and co-workers [*J. Chem. Phys.* **2008**, *128*, 134110]. The best results, in comparison with the MS-CASPT2 results of Thiel, were obtained using a RASSCF space that contains at most one hole and one particle in the RAS1 and RAS3 spaces, respectively, which we denote as RAS[1,1]. This subset of configurations recovers mainly the effect of polarization and semi-internal electronic correlation that is only included in CASSCF in an averaged way. Adding all-external correlation by allowing double excitations from RAS1 and RAS2 into RAS3 did not improve the results, and indeed, they were slightly worse. The accuracy of the first-order RASSCF computations is demonstrated to be a function of whether the state of interest can be classified as covalent or ionic in the space of configurations built from orbitals localized onto atomic sites. For covalent states, polarization and semi-internal correlation effects are negligible (RAS[1,1]), while for ionic states, these effects are large (because of inherent diffusiveness of these states compared to the covalent states) and, thus, an acceptable agreement with MS-CASPT2 can be obtained using first-order RASSCF with the extra basis set involving 3p orbitals in most cases. However, for those ionic states that are quasi-degenerate with a Rydberg state or for nonlocal n$\pi^*$ states, there remains a significant error resulting from all external correlation effects.




## 1. INTRODUCTION

The theoretical study of excited state reactivity using nonadiabatic molecular dynamics represents a challenge for electronic structure methods because they must account for a changing importance of open versus closed shell configurations and covalent versus ionic character in the wave function.[1] This balance involves electronic correlation effects. Thus, one needs a balanced description of excited states (i.e., comparable excitation energies and ordering of states) of different character with the same relative accuracy. The challenge is to include electron correlation effects via theoretical methods that are sufficiently efficient to perform the very many energy evaluations required in nonadiabatic dynamics. The essential feature here is to include only the part of the electron correlation effect that is different for each electronic state and omit the part that is the same for each excited state. Furthermore, the theoretical method used needs to permit the computation of first and second derivatives analytically (as opposed to finite difference computation) as well as the necessary derivative couplings.

For the treatment of static (or internal) correlation,[2−4] a multireference method such as complete active space self-consistent field (CASSCF)[5] can be used as the starting point. The addition of complete active space second-order perturbation theory (CASPT2)[6−8] in which dynamic correlation is included using second-order multireference perturbation theory on top of CASSCF results is the accepted standard (for recent reviews of the computation of excited states including dynamics, see the work of Lischka et al.,[9] González et al.,[10] and Dreuw et al.[11]). However, this method includes all electron correlation effects, including those that do not change between the excited states under consideration. Furthermore, while gradients and derivative couplings have recently been made available for CASPT2, analytic Hessians are not yet available.[12] In the quantum mechanical treatment of nuclear motion one needs the Hessian, as well as gradients, for the evaluation of the matrix elements.[13]

The restricted active space self-consistent field (RASSCF)[14−17] is a computational method that can include those differential contributions to electron correlation that are required to describe the different correlation effects in excited state chemistry. Furthermore, RASSCF is a fully variational

Received: April 20, 2019
Revised: May 31, 2019
Published: May 31, 2019

Downloaded by IMPERIAL COLLEGE LONDON at 02:50:40:716 on June 21, 2019
from https://pubs.acs.org/doi/10.1021/acs.jpca.9b03715.

approach, where full orbital and CI coefficient optimization is used; thus, analytical derivatives are available. The only problem is the choice of the active orbitals to be included. This choice is, in turn, dictated by the need to include the different contributions of electron correlation in a controlled way by restricting the number of configuration state functions (CSF) generated using RASSCF to yield a balanced description of excited states.

Dynamic correlation involves double excitations of the inactive or active orbitals into the virtual space. We refer to this as all external correlation. In CASSCF, single excitations from the inactive orbitals (holes) to the virtual orbitals (particles) vanish only in an average way. Indeed, the set of configurations (the first-order multiconfiguration wave function, (MR-FOCI)[18,19]) involving the creation of at most a single "hole" in the inactive space and a single "particle" in the virtual space with respect to a CASSCF reference make the most important correction to the RASSCF wave function. Here the key physical idea involves the concept of semi-internal correlation introduced by Sinanoğlu.[20,21] Semi-internal correlation involves a "double" excitation which consists of the simultaneous excitation of an active electron within the active space together with an excitation of an inactive electron to the virtual space.

The main aim of this work is to demonstrate the accuracy of the RASSCF approach for the computation of valence singlet excited states and, thus, the potential suitability of the method for investigating potential energy surfaces relevant for photochemistry using on-the-fly *ab initio* nonadiabatic dynamics that require analytical derivatives. The other aim is to demonstrate that first-order RASSCF (which recovers semi-internal correlation and polarization) dominates the differential electronic correlation between excited states with very different characters (ionic or covalent). The first-order RASSCF effects vanish on average in CASSCF yet are small for covalent states and large for ionic states. The assessment is done on a set of small to medium-sized organic molecules based on Thiel's benchmark set,[22] where vertical excitation energies computed with MS-CASPT2 method are available for direct comparison. Due to the scaling of the current RASSCF approach, only a subset of the original benchmark set has been investigated. Of course, the RASSCF scheme, used in this paper, only tests the error in the correlation effect in the vicinity of the Franck–Condon region. However, as we will discuss subsequently, it turns out that the largest errors are encountered for specific types of electronic states only. Accordingly, the results may have considerable generality away from the Franck–Condon region.

The accuracy of RASSCF has been assessed using MS-CASPT2/TZVP results of Schreiber et al.[22] as reference values in order to compare two levels of theory keeping the basis set fixed. Better results for MS-CASPT2 can be found in the work of Silva-Junior et al.[23] where a diffuse basis set was employed. This work will be discussed after we present our approach, which is focused on the improvement of CASSCF results using a RASSCF approach.

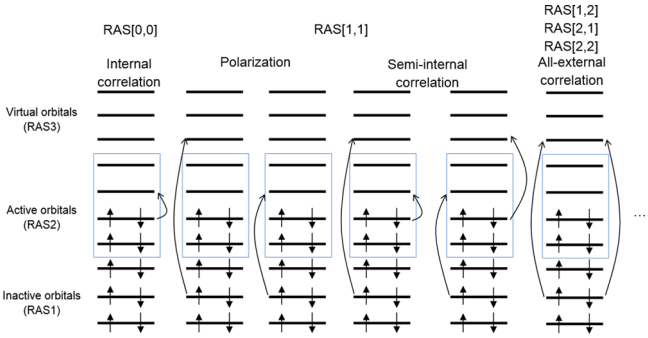
## 2. THEORETICAL DETAILS

RASSCF differs from CASSCF in that the active space is divided into three subspaces: RAS1 where at most a few "holes" (usually 1 or 2) are allowed, RAS3 where at most a few "particles" (usually 1 or 2) are allowed, and RAS2 where all possible occupancies are allowed. Here we use the term "hole" to denote a vacancy (i.e., excitation) in inactive orbitals/electrons and "particle" to denote an occupancy in a virtual orbital that is unoccupied in all reference configurations. The RAS2 set of orbitals is usually the same space as CASSCF when used to include correlation. Other strategies for choosing the RAS space partition can be found in the literature.[24−28] For example, the method is also used for treating systems where CASSCF becomes impractical such as in transition metal complexes.[29,30] In this work, RASSCF is merely employed as a form of multireference CI within a subset (i.e., a window) of optimized closed shell and virtual orbitals. The original first-order multiconfiguration wave function (MR-FOCI)[18,19] method was just a such a special case where one has at most 1 hole in the RAS1 and 1 particle in the RAS3.

In previous work (Santolini et al.[31]), we demonstrated that an "initial" RASSCF active space can be constructed from natural bond orbitals (NBO).[32] The NBO are constructed by combining natural atomic orbitals (NAO).[33,34] The choice of which NBO orbitals to include is guided by the concept of correlating orbital pairs or oscillator orbitals.[35,36] The order (first-order or second-order) of the RASSCF configuration interaction expansion is associated with the number of holes and particles in RAS1 and RAS3 which is, in turn, guided by the concepts of semi-internal and dynamic or all-external correlation.[20,21] We now give some insight into this aspect.

We can illustrate the preceding concepts with an example. We suppose that we have a molecule where the excited states involve different occupancies of $\pi$ orbitals. Of course, the total correlation energy for any excited state is dominated by the correlation energy of $\sigma$ electrons (computed by double excitations of the $\sigma$ electrons) into the virtual orbitals and the active orbitals. But this is the same for all excited states and can be neglected when our main focus is the energy difference between states. Thus, for the excited states of the $\pi$ electrons, the correlation contributions can be separated into three types (see Scheme 1) according to the excitation pattern: (1) the internal correlation (i.e., static correlation), (2) polarization plus semi-internal effects, and (3) all-external (dynamic) correlation.[20,21] The last, (3), as we have just discussed, is approximately the same for all the valence excited states. In Scheme 1, one uses the term inactive orbitals to denote the doubly occupied orbitals in CASSCF, a subset of which become the RAS1 orbitals of RASSCF in the way we use it here. The effect of polarization and semi-internal correlation on the energies of excited states with different bond character (e.g., covalent versus ionic) is strongly structure and state dependent. Thus, the magnitude of the correlation of a $\sigma$ electron and a $\pi$ electron will depend on the nature of the electronic state (covalent vs zwitterionic).[31,37−40] Thus, we

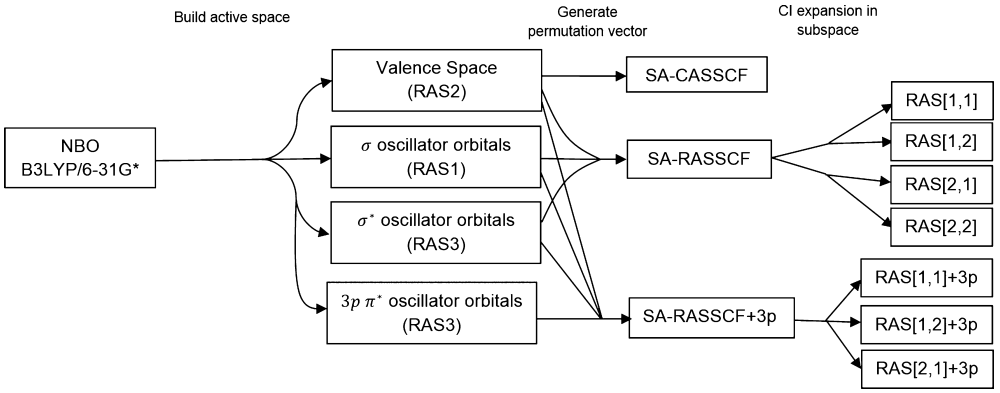
**Scheme 1. Classification of RASSCF Configuations**





DOI: 10.1021/acs.jpca.9b03715
J. Phys. Chem. A XXXX, XXX, XXX−XXX

**Scheme 2. Summary of RASSCF Procedure Used**




expect the most important differential contribution to the relative energies of the $\pi$ orbital excited states will come from type (2) above, which is a RASSCF computation with at most 1 hole in the RAS1 space and 1 particle in the RAS3 space. We shall refer to this level as first-order RASSCF, which we denote by RAS[1,1]. Within RASSCF it is straightforward to add configurations with 2 holes and 2 particles in the RAS1 and RAS3 spaces (RAS[1,2], RAS[2,1], and RAS[2,2]). This all-external contribution should be very weakly structure/state dependent.[1,20,21]

Oscillator orbitals, a concept introduced by Foster and Boys,[35,36] are a set of virtual orbitals constructed from an initial set of occupied orbitals so that they possess an additional node in the orbital, yet have the same spatial extent. Thus, using localized NBO as a starting guess for the RASSCF computation allows the explicit construction of an active space by placing correlating orbitals localized in the same region of space as the occupied orbitals used in a standard CASSCF active space. The pair of oscillator orbitals is formed by including the strongly occupied NBO in the closed shell manifold RAS1 and the weakly occupied NBO in the RAS3 subspace.

Using the same model as an example, we now discuss the nature of the orbitals that need to be included. To describe the $\pi\pi^*$ and $n\pi^*$ transitions involved in the lower valence electronic states, the most important orbitals (2p $\pi$ and 2p n) are included in the main active space (RAS2). The $\sigma$ correlation is included by adding the $\sigma$ orbitals of the frame of the molecules in the RAS1 subspace and their antibonding equivalent ($\sigma^*$) in RAS3 as oscillator orbitals. The correlation of the $\pi$ system can be recovered by adding an extra set of 3p orbitals (in RAS3) that will act as oscillator orbitals for the $\pi$ system.[37] Different scaling factors of these 3p orbitals are needed to avoid collapsing to Rydberg states while using a diffuse basis set.

## 3. COMPUTATIONAL DETAILS

The geometries for the benchmark set of 19 organic molecules were taken from the Supporting Information of the work by Schreiber et al.[22] where the geometry optimizations had been carried out at the MP2/6-31G* level of theory. The vertical excitation energies here are computed using CASSCF and RASSCF following a systematic approach used previously by Santolini et al.[31]

All computations are performed with Gaussian 16.[41] The CASSCF/RASSCF computations were done using a state-average (SA) over the three lowest-lying valence states with equal weight unless stated otherwise (see the Supporting Information for details). The NBO were symmetry adapted to the appropriate abelian symmetry point group (see the Supporting Information). Two types of atomic orbital basis sets were used for the computation of vertical excitation energies: a standard Pople 6-31G* basis set and an extended basis set derived from the 6-31G* by adding a set of 3p atomic orbitals to carbon, oxygen and nitrogen where the 3p functions are taken from the 6-31G basis set of silicon, sulfur, and phosphorus, respectively. The notation "+3p" is used to indicate the use of an extended basis set and the scaling factor of the 3p atomic orbitals is shown in parentheses (either 0.5 or 1.0). The protocol used can be summarized as follows:

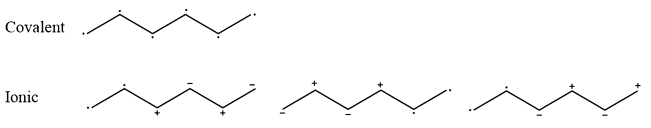
1. The initial orbital guesses for the RASSCF are generated by selecting a set of NBO at the B3LYP/6-31G* level of theory.
2. The CASSCF energies were computed using a standard active space (see the Supporting Information) for each molecule.
3. The RASSCF energies are computed by using a standard active space for the RAS2 subspace and by correlating different types of molecular orbitals (MO) added inside the RAS1 and RAS3 subspaces. The number of holes ($h$) and particles ($p$) allowed in the subspace is given by the following notation "RAS[$h$,$p$]".
    i. The $\sigma$ and $\sigma^*$ MOs (oscillator orbital pair for each 2p $\pi$ orbital) are added respectively into RAS1 and RAS3 with the following settings for the configuration interaction expansion.
        - 1 hole and 1 electron allowed in RAS1 and RAS3 (RAS[1,1])
        - 1 hole and 2 electrons allowed in RAS1 and RAS3 (RAS[1,2])
        - 2 holes and 1 electron allowed in RAS1 and RAS3 (RAS[2,1])
        - 2 holes and 2 electrons allowed in RAS1 and RAS3 (RAS[2,2])
    ii. The $\sigma$ MOs are added into RAS1 and the $\sigma^*$ and 3p $\pi^*$ MO are put into the RAS3 subspace with the following settings for the configuration interaction expansion.
        - 1 hole and 1 electron allowed in RAS1 and RAS3 (RAS[1,1]+3p)
        - 1 hole and 2 electrons allowed in RAS1 and RAS3 (RAS[1,2]+3p)
        - 2 holes and 1 electron allowed in RAS1 and RAS3 (RAS[2,1]+3p)


DOI: 10.1021/acs.jpca.9b03715
J. Phys. Chem. A XXXX, XXX, XXX−XXX

A summary of the method employed can be found in Scheme 2.

For the discussion of the contribution of various types of electron correlation, it turns out to be convenient to classify states as ionic or covalent. The ionic character of a $\pi\pi^*$ excited state was determined by carrying out a CI analysis in the space of the RAS2 orbitals after localization using the Boys localization method.[36] This localization of the orbitals has the property of localizing them onto atomic sites. The CI eigenvector now corresponds to an orthogonal VB expansion. In this expansion, there are two types of configurations: covalent, where each atom-localized orbital has an occupancy of one, and ionic, where some orbitals have an occupancy of zero or two. In Scheme 3 below, we show an example of a covalent configuration on line one and, on line two, we show the three dominant ionic configurations in a hexatriene ionic state. The $n\pi^*$ excited states can also be classified using the covalent/ionic terminology by assessing the localized character of the transition using the natural orbitals obtained from the RASSCF (see examples for pyridine, pyrazine, pyrimidine, and pyridazine in the Supporting Information where we have plotted these orbitals).

**Scheme 3. Examples of Covalent and Ionic VB Structures for Hexatriene**




## 4. RESULTS AND DISCUSSION

A set of 19 small to medium size organic molecules (see Figure 1), based on the data set of Schreiber et al.[22] has been used for the comparison of vertical excitation energies at the RASSCF level.

The states of interest, in the molecules selected for study for the evaluation of the vertical excitation energies in this work, were the two lowest-lying singlet valence excited states of $\pi\pi^*$ and $n\pi^*$ character. (A larger number of excited states can be found in the work of Schreiber et al.[22] For example, they also included lowest-lying excited states of $\sigma\pi^*$ type for some molecules such as cyclopropene, formaldehyde, and acetone. This type of excited state was not included in this work and would have required a slightly different RAS2 active space, which included $\sigma$ MO.) All our results are compared to MS-CASPT2 data provided by the work of Schreiber et al.[22] Our objective was to determine the efficiency of adding different types of electron correlation using the RASSCF variational approach. (Note that Schreiber et al.[22] use mostly CASPT2 values obtained from the literature and their own computations for singlet states as theoretical best estimates. Here we only use MS-CASPT2 results as a reference due to their obvious connections to SA-CASSCF/RASSCF.)

The results for the test set of 19 molecules are summarized in Figure 2. Here we show the error of the vertical excitation

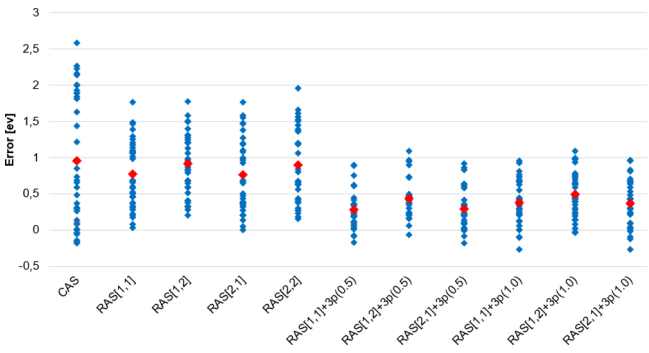



**Figure 2.** Scatter plot of error (in blue) and its average (in red) on vertical excitation of all investigated excited states computed at CASSCF/RASSCF level compared MS-CASPT2 results of Schreiber et al.[22] Rydberg states are not included.

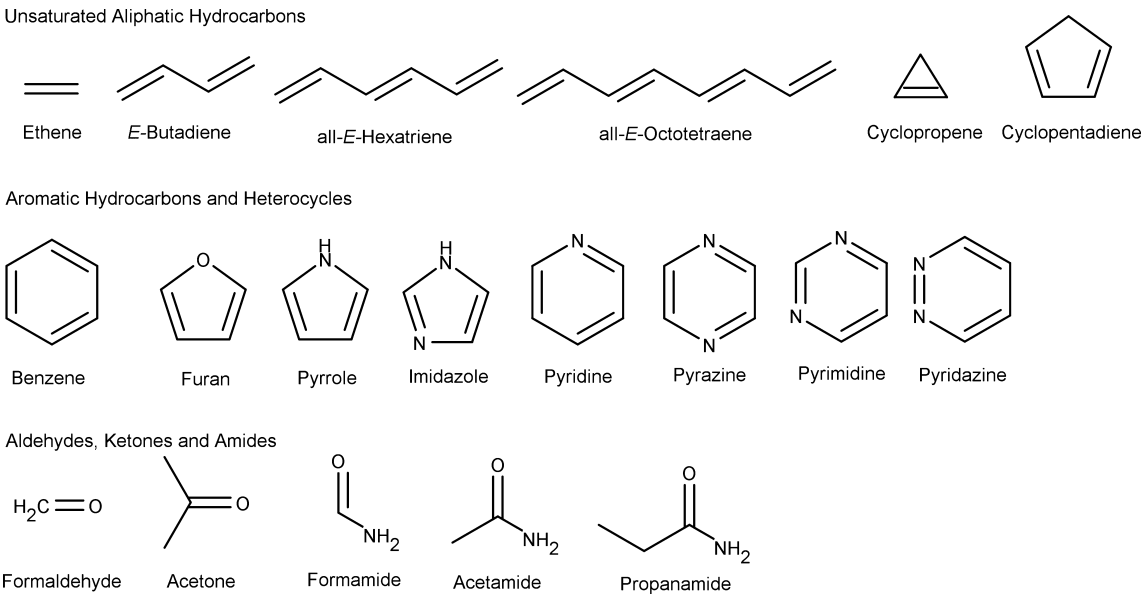



**Figure 1.** List of molecules investigated for the benchmark of vertical excitation energies using CASSCF/RASSCF based the data set of Schreiber et al.[22] A subset of the Thiel benchmark set is used rather than the full set due to unfavorable scaling of the RASSCF active space with the current approach.



DOI: 10.1021/acs.jpca.9b03715
J. Phys. Chem. A XXXX, XXX, XXX−XXX

(relative to the MS-CASPT2/TZVP results of Schreiber et al.[22]), as well as the average error (in red) produced by adding different electron correlation effects (see the Supporting Information for detailed vertical excitation energies of each molecule and a comparison of the error between molecules).

From Figure 2, one can see that RAS[1,1]+3p(0.5) gives the most accurate results, as the average error is the lowest and the deviations are smaller overall. In addition, the inclusion of the 3p $\pi^*$ MO in the RAS3 space is essential, irrespective of the order of the correlation effect. Notice also that there are still some points with an error exceeding 0.5 eV. These "outliers" correspond to a few examples where the model we are using fails. Further, it turns out that the classification of the $\pi\pi^*$ states into covalent or ionic (or Rydberg in a few cases) and of the n$\pi^*$ states as local (covalent) or nonlocal (ionic) gives some physical insights. We now discuss these points in more detail.

Overall, the CASSCF and RASSCF results tend to overestimate the vertical excitation energies compared to the reference values. Starting with RAS[1,1] with the addition of the $\sigma$ and $\sigma^*$ MO in RAS1 and RAS3 (first four columns of Figure 2), the overall error in the vertical excitation energies for the two lowest-lying excited states of n$\pi^*$ and $\pi\pi^*$ type remains comparable to CASSCF results. Only with the addition of the 3p $\pi^*$ MO in the RAS3 subspace does one reduce the average error to less than 0.5 eV. Given that the RAS[1,1] contribution needs these orbitals (cf column 2 versus column 6), it is clear that the semi-internal correlation effect is in fact dominated by these 3p $\pi^*$ MO.

In Figure 2, it can be seen that the differential effect of adding double excitations into the virtual space (RAS3) with RAS [1,2] and RAS[2,2] compared to RAS[1,1] and RAS[2,1] is rather small and indeed tends to make the results slightly worse. Moreover, allowing double excitation from the occupied orbitals (RAS1) with RAS[2,1] shows almost no change compared to RAS[1,1]. These contributions are part of the all-external contribution to the correlation energy. This part of the correlation energy is slowly convergent, and including a subset of the orbitals does not appear to be strategically useful. The double excitations to the RAS3 subspace should allow a partial recovery of dynamic correlation but the state benefiting the most from this enhanced correlation is the ground state, thus blue-shifting (overestimating) the vertical excitation energy with respect to the reference.

As we have just observed, the addition of the 3p $\pi^*$ MOs in the RAS3 subspace is essential for recovering the $\pi$ contribution to first-order RASSCF (RAS[1,1]). Here we also tested two types of scaling factor for the extra basis set but obtained very similar results for each one. On one hand, using a scaling factor of 1.0 makes the results slightly worse. On the other hand, using a value of 0.5 for the scaling can bias convergence toward Rydberg states rather than the target valence states.

RAS[1,1] (i.e., including polarization and semi-internal correlation) with the 3p $\pi^*$ MO thus seems to provide the optimum strategy for calculating the vertical excitation energies of the excited states selected here. This idea has already been shown in the past by the early work of Schaefer et al.[18] in the method called first-order configuration interaction (FOCI). Of course, here we show that the dominant effect can be recovered with the compact RASSCF approach with only a subset of optimized orbitals, as opposed to the full set of inactive and virtual orbitals in multireference CI. Thus, the main differential electronic correlation is captured using a first-order correction, RAS[1,1]+3p, to CASSCF with RASSCF.

All the excited states do not benefit in the same way from the addition of the different contributions of electron correlation. Previous work has shown that the improvement to the description of valence excited states depends on the nature of such excited states, and thus, the contribution of electron correlation is rather different for ionic versus covalent states.[31,37,40] Indeed, the vertical excitation energies of certain types of state can be very similar to MS-CASPT2 even at the CASSCF level. We now give some discussion of the role of various types of correlation effect for the different types of electronic state studied in this work.

In Figure 3, we show the data, presented earlier, in Figure 2 broken down into contributions obtained by classifying the

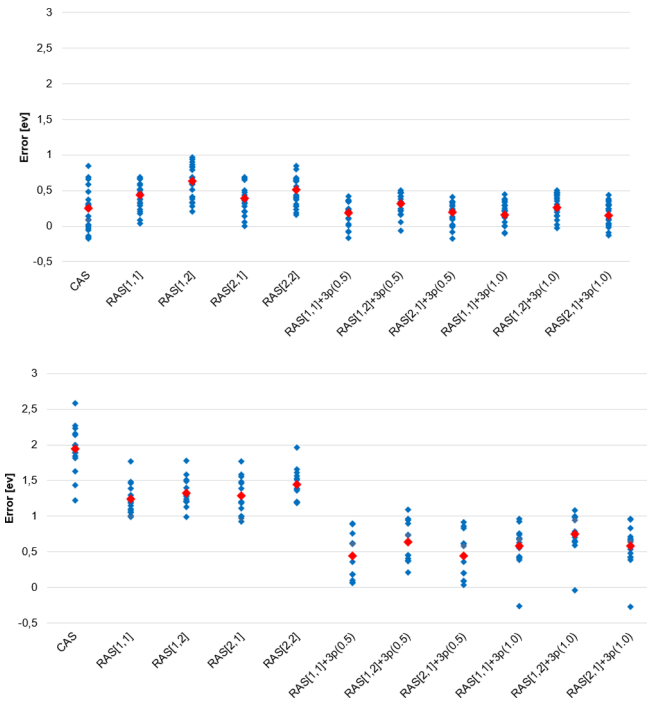



**Figure 3.** Scatter plot of the error (in blue) and its average (in red) for the vertical excitation energies for covalent (top) and ionic (bottom) excited states. The comparison is done against MS-CASPT2 results of Schreiber et al.[22] The Rydberg states are not included. (The three outliers in the last three columns correspond to the ethene V state which is a Rydberg−ionic mixture.)

different $\pi\pi^*$ excited states as either covalent or ionic. In addition, the n$\pi^*$ transitions can also be classified as either covalent or ionic by determining whether the excitation is local or nonlocal in character. The pyrazine (Supporting Information Table S12 and Figure S3) molecule illustrates this idea. Here the $1^1B_{3u}$ state is classified as local because the excitation is localized on the N atoms. In contrast, for the $1^1A_u$ state, the excitation (Supporting Information Figure S3) involves a charge transfer from the N atom lone pairs to the C atoms thus inducing a zwitterionic charge separation. Thus, the n$\pi^*$ transitions that are "local" are included in Figure 3 as covalent, while those that are nonlocal are included as ionic.

From Figure 3 one can observe that the RAS[1,1]+3p level computations for the covalent $\pi\pi^*$ states hardly change the results from the CASSCF level, which are already very similar to MS-CASPT2 results. The same observation can be made for


DOI: 10.1021/acs.jpca.9b03715
J. Phys. Chem. A XXXX, XXX, XXX−XXX

the n$\pi^*$ excited states with local character (in imidazole, pyridine, pyrimidine, and pyridazine). In contrast, in Figure 3, one can observe that at the CASSCF level the vertical excitation energies of zwitterionic states are significantly overestimated. Upon addition of polarization and semi-internal correlation, via RAS[1,1]+3p, the description of ionic states is significantly improved in general, although there is some dispersion and some outliers that we will discuss subsequently. The further addition of external correlation via (RAS[1,2], RAS[2,1], and RAS[2,2]) actually leads to slightly worse results for the covalent and ionic states because it leads to an additional stabilization of the ground state.

In summary, the addition of semi-internal correlation at the RAS[1,1]+3p level significantly improves the description of ionic states. At the same time, the covalent states description at the RAS[1,1]+3p level remains comparable to CASSCF results, indicating that the semi-internal correlation effect is very small for covalent states.

The inherent diffusiveness of the orbital description of ionic states explains the improvement obtained by augmenting the set of valence 2p $\pi$ with the more diffuse 3p $\pi$ orbitals. This idea was first proposed by Roos and co-workers[37] and it is also documented in other work.[42−45] Allowing the excited electron to populate the 3p $\pi$ orbitals leads to a better stabilization of ionic states by allowing a delocalization of the charge in the more diffuse orbitals. Moreover, ionic states are more sensitive to the diffusiveness of the basis set compared to covalent states. Using a more contracted description of 3p orbitals (i.e., using a scaling factor of 1.0 for the extra basis) leads to larger overestimation of vertical excitation energies compared to the more diffuse basis set.

From Figure 3, one can observe that the dispersion of the error around the average is larger for ionic states in spite of the fact that RAS[1,1] with the 3p orbitals improves the results overall. Thus, there are still a few ionic states where a large overestimation of the excitation energy remains at the RAS[1,1]+3p level. Examples include, the low-lying $\pi\pi^*$ ionic states of ethene ($B_{1u}$) (shown as the three outliers in Figure 3), hexatriene ($B_u$), octatetraene ($B_u$), benzene ($B_{1u}$), and pyrrole ($B_2$) and the nonlocal n$\pi^*$ transition of pyrazine ($A_u$). In all of these molecules, except for pyrazine, there is a low-lying Rydberg state of similar symmetry that can mix with these different ionic states.[46−48]

In these examples, using the 3p (more diffuse) basis set leads to convergence to a Rydberg state in the computations on ethene and pyrrole. In the other cases, a Rydberg−ionic mixed state is obtained.

In the case of pyrazine, the n$\pi^*$ excited state with nonlocal character (i.e., ionic) is only slightly improved with the RAS[1,1]+3p computation and using a different scaling factor for the extra basis (different diffusiveness) barely changes the results for the $A_u$ state. This zwitterionic state has the charges well separated and probably requires a dispersion energy correction.

Thus, while the RAS[1,1]+3p approach improves the description of ionic states compared to CASSCF (Figure 3), there remain large errors for particular isolated $\pi\pi^*$ ionic states and nonlocal n$\pi^*$ transitions. Indeed, in some of these cases, the ionic state and covalent states are ordered incorrectly (see the Supporting Information for details). For $\pi\pi^*$ ionic states this arises from the fact that there is a significant differential effect of all-external correlation of the ionic state as opposed to the Rydberg state. In contrast, the description of $\pi\pi^*$ ionic states in molecules such as butadiene ($B_u$), cyclopentadiene ($B_2$), furan ($B_2$), formamide (A′), acetamide (A′), and propanamide (A′) are comparable to the MS-CASPT2 results of Schreiber et al.[22]

The focus of the work presented in this paper was to investigate the accuracy of first-order RASSCF where additional electronic correlation has been included within a variational method. While we are using MS-CASPT2/TZVP results as reference values for direct comparison, it is known that the (Coupled Cluster 3) CC3 approach yields more accurate results than CASPT2 for low-lying valence states with no strong double-excitation character.[50] Furthermore, the effect of a larger basis set was not investigated in this current work.

In Figure 4, we compare of our current results against other values computed with different methods found in the literature[22,23,49−52] (see Tables S20−S22 in the Supporting Information), using as a reference, the theoretical best estimates (i.e., TBE-2) of Silva-Junior et al.[23] Compared to these reference values, RASSCF[1,1]/6-31G*+3p tends to consistently overestimate the vertical excitation energies as shown by the overlapping value of the mean absolute error and mean standard error. The overall accuracy of RASSCF[1,1]/6-31G*+3p is similar to CC2/TZVP results. Our RASSCF results, obtained at the 6-31G*+3p level, will certainly improve with a larger basis set.

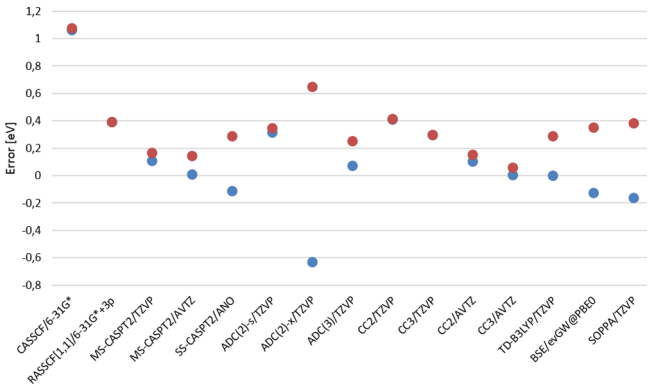



**Figure 4.** Comparison of the mean standard error (in blue) and mean absolute error (in red) for the vertical excitation energies of the 19 molecules computed with different electronic structure method (see Tables S20−S22 in the Supporting Information for details). The values are compared to the theoretical best estimated (i.e., TBE-2) in the work of Silva-Junior et al.[23] The statistics are done over 36 excited states for all methods, except for CC2/AVTZ and CC3/AVTZ where only 29 and 6 values were available for the current set of molecules, respectively.

## 5. CONCLUSION

In this work, we have investigated the efficacy of using first-order RASSCF RAS[1,1] for the treatment of valence excited states with different characters. Physically, RAS[1,1] can be understood using the description of electron correlation of Sinanoğlu,[20,21] which separates the electron correlation effect for a multireference system into internal, polarization plus semi-internal, and all-external correlation energy. In our computations, the initial active space was constructed on the principle of pairs of oscillator orbitals, first described by Foster and Boys.[36] We implemented these using natural bond orbitals as described by Weinhold.[32] The results were rationalized by


DOI: 10.1021/acs.jpca.9b03715
J. Phys. Chem. A XXXX, XXX, XXX−XXX

classifying the different excited states as covalent or ionic (or as local and nonlocal for the nπ* state). The RAS[1,1]+3p computations that include mainly the polarization and semi-internal correlation effects were able to correct the error in the CASSCF results for ionic states in most cases. In general, the role of all-external correlation (RAS[1,2], RAS[2,1], RAS-[2,2]) did not change the results significantly and even made them slightly worse. The examples with strong ionic−Rydberg mixing or near degeneracies could not be properly described at the RAS[1,1]+3p level of theory.

In general, the use of RAS[1,1]+3p seems to recover the state-dependent part of the electron correlation energy, except in cases of strong ionic−Rydberg mixing or clear charge separation. Further, both first and second derivatives with respect to nuclear motion can be computed analytically. Thus, the method proposed here serves as an alternative for treating excited states, based on a variational approach for quantum molecular dynamics where many energies, gradient, Hessians, and nonadiabatic couplings are computed.

The assessment of the efficacy of RASSCF has been done in the Franck−Condon region. Nevertheless, the failures of RAS[1,1]+3p for ionic states seems to be associated mainly with quasi-degenerate states so the results may be more generally applicable. Thus, the semi-internal correlation should be structure dependent as well as state dependent as shown here. However, this conjecture remains to be investigated.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.9b03715.

Vertical excitation energies and covalent/ionic character for the lowest-lying valence excited state, active space employed for the CASSCF and RASSCF computation, molecular orbitals for the nπ* transition, data from literature for comparison with other electronic structure method, error comparison between the 19 molecules at CASSCF and RAS[1,1]+3p (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

*E-mail: mike.robb@imperial.ac.uk.

### ORCID

Michael A. Robb: 0000-0003-0478-8301


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We acknowledge the Imperial College Research Computing Service (DOI: 10.14469/hpc/2232) where all computations were carried out. This work formed part a project carried out for the Master project in Molecular and Biological Chemistry at Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. This work received partial financial support from Gaussian Inc.

## ■ REFERENCES

(1) Robb, M. A. *Theoretical Chemistry for Electronic Excited States; Theoretical and Computational Chemistry Series*; The Royal Society of Chemistry, 2018. ( DOI: 10.1039/9781788013642)

(2) Löwdin, P. O. Quantum Theory of Many-Particle Systems. III. Extension of the Hartree-Fock Scheme to Include Degenerate Systems and Correlation Effects. *Phys. Rev.* **1955**, *97*, 1509−1520.

(3) Clotet, A.; Daudey, J. P.; Malrieu, J. P.; Rubio, J.; Spiegelmann, F. The Effect of Dynamical Correlation on the Valence Wavefunction of Molecules: Dressed Complete Active Space Self-Consistent Field Calculations. *Chem. Phys.* **1990**, *147*, 293−307.

(4) Karafiloglou, P.; Malrieu, J. P. The Effect of Electronic Correlation on Molecular Wavefunctions. *Chem. Phys.* **1986**, *104*, 383−398.

(5) Roos, B. O.; Taylor, P. R. A Complete Active Space SCF Method (CASSCF) Using A Density Matrix Formulated Super-CI Approach. *Chem. Phys.* **1980**, *48*, 157−173.

(6) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-Order Perturbation Theory with a CASSCF Reference Function. *J. Phys. Chem.* **1990**, *94*, 5483−5488.

(7) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-Order Perturbation Theory with a Complete Active Space Self-Consistent Field Reference Function. *J. Chem. Phys.* **1992**, *96*, 1218.

(8) Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. The Multi-State CASPT2 Method. *Chem. Phys. Lett.* **1998**, *288*, 299−306.

(9) Lischka, H.; Nachtigallová, D.; Aquino, A. J. A.; Szalay, P. G.; Plasser, F.; MacHado, F. B. C.; Barbatti, M. Multireference Approaches for Excited States of Molecules. *Chem. Rev.* **2018**, *118*, 7293−7361.

(10) González, L.; Escudero, D.; Serrano-Andrés, L. Progress and Challenges in the Calculation of Electronic Excited States. *ChemPhysChem* **2012**, *13*, 28−51.

(11) Dreuw, A.; Head-Gordon, M. Single-Reference Ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105*, 4009−4037.

(12) Park, J. W.; Shiozaki, T. Analytical Derivative Coupling for Multistate CASPT2 Theory. *J. Chem. Theory Comput.* **2017**, *13*, 2561−2570.

(13) Worth, G. A.; Robb, M. A.; Lasorne, B. Solving the Time-Dependent Schrodinger Equation for Nuclear Motion in One Step: Direct Dynamics of Non-Adiabatic Systems. *Mol. Phys.* **2008**, *106*, 2077−2091.

(14) Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. A. Determinant Based Configuration Interaction Algorithms for Complete and Restricted Configuration Interaction Spaces. *J. Chem. Phys.* **1988**, *89*, 2185−2192.

(15) Malmqvist, P.-Å.; Rendell, A.; Roos, B. O. The Restricted Active Space Self-Consistent-Field Method, Implemented with a Split Graph Unitary Group Approach. *J. Phys. Chem.* **1990**, *94*, 5477−5482.

(16) Klene, M.; Robb, M. A.; Blancafort, L.; Frisch, M. J. A New Efficient Approach to the Direct Restricted Active Space Self-Consistent Field Method. *J. Chem. Phys.* **2003**, *119*, 713−728.

(17) Tomasello, G.; Garavelli, M.; Orlandi, G. Tracking the Stilbene Photoisomerization in the S1 State Using RASSCF. *Phys. Chem. Chem. Phys.* **2013**, *15*, 19763−19773.

(18) Schaefer, H. F.; Klemm, R. A.; Harris, F. E. First-Order Wavefunctions, Orbital Correlation Energies, and Electron Affinities of First-Row Atoms. *J. Chem. Phys.* **1969**, *51*, 4643−4650.

(19) Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Müller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; et al. High-Level Multireference Methods in the Quantum-Chemistry Program System COLUMBUS: Analytic MR-CISD and MR-AQCC Gradients and MR-AQCC-LRT for Excited States, GUGA Spin-Orbit CI and Parallel CI Density. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664−673.

(20) Öksüz, I.; Sinanoğlu, O. Theory of Atomic Structure Including Electron Correlation. I. Three Kinds of Correlation in Ground and Excited Configurations. *Phys. Rev.* **1969**, *181*, 42−53.

(21) Sinanoğlu, O. Many-Electron Theory of Atoms, Molecules and Their Interactions. In *Advances in Chemical Physics*; Prigogine, I., Ed.; John Wiley & Sons, Ltd.: New York, 1964; pp 315−412.

(22) Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. Benchmarks for Electronically Excited States: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys.* **2008**, *128*, 134110.
(23) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. Benchmarks of Electronically Excited States: Basis Set Effects on CASPT2 Results. *J. Chem. Phys.* **2010**, *133*, 174318.
(24) Tokmachev, A. M.; Boggio-Pasqua, M.; Mendive-Tapia, D.; Bearpark, M. J.; Robb, M. A. Fluorescence of the Perylene Radical Cation and an Inaccessible D0/D1 Conical Intersection: An MMVB, RASSCF, and TD-DFT Computational Study. *J. Chem. Phys.* **2010**, *132*, 044306.
(25) Boggio-Pasqua, M.; Robb, M. A.; Bearpark, M. J. Photostability via a Sloped Conical Intersection: A CASSCF and RASSCF Study of Pyracylene. *J. Phys. Chem. A* **2005**, *109*, 8849−8856.
(26) Bearpark, M. J.; Ogliaro, F.; Vreven, T.; Boggio-Pasqua, M.; Frisch, M. J.; Larkin, S. M.; Morrison, M.; Robb, M. A. CASSCF Calculations for Photoinduced Processes in Large Molecules: Choosing When to Use the RASSCF, ONIOM and MMVB Approximations. *J. Photochem. Photobiol., A* **2007**, *190*, 207−227.
(27) Krausbeck, F.; Mendive-Tapia, D.; Thom, A. J. W.; Bearpark, M. J. Choosing RASSCF Orbital Active Spaces for Multiple Electronic States. *Comput. Theor. Chem.* **2014**, *1040−1041*, 14−19.
(28) Veryazov, V.; Malmqvist, P.-Å.; Roos, B. O. How to Select Active Space for Multiconfigurational Quantum Chemistry? *Int. J. Quantum Chem.* **2011**, *111*, 3329−3338.
(29) Malmqvist, P.-Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The Restricted Active Space Followed by Second-Order Perturbation Theory Method: Theory and Application to the Study of $CuO_2$ and $Cu_2O_2$ Systems. *J. Chem. Phys.* **2008**, *128*, 204109.
(30) Sauri, V.; Serrano-Andrés, L.; Shahi, A. R. M.; Gagliardi, L.; Vancoillie, S.; Pierloot, K. Multiconfigurational Second-Order Perturbation Theory Restricted Active Space (RASPT2) Method for Electronic Excited States: A Benchmark Study. *J. Chem. Theory Comput.* **2011**, *7*, 153−168.
(31) Santolini, V.; Malhado, J. P.; Robb, M. A.; Garavelli, M.; Bearpark, M. J. Photochemical Reaction Paths of Cis-Dienes Studied with RASSCF: The Changing Balance between Ionic and Covalent Excited States. *Mol. Phys.* **2015**, *113*, 1978−1990.
(32) Foster, J. P.; Weinhold, F. Natural Hybrid Orbitals. *J. Am. Chem. Soc.* **1980**, *102*, 7211−7218.
(33) McWeeny, R. Some Recent Advances in Density Matrix Theory. *Rev. Mod. Phys.* **1960**, *32*, 335−369.
(34) Rothenberg, S.; Davidson, E. R. Natural Orbitals for Hydrogen-Molecule Excited States. *J. Chem. Phys.* **1966**, *45*, 2560−2576.
(35) Foster, J. M.; Boys, S. F. A Quantum Variational Calculation for HCHO. *Rev. Mod. Phys.* **1960**, *32*, 303−304.
(36) Foster, J. M.; Boys, S. F. Canonical Configurational Interaction Procedure. *Rev. Mod. Phys.* **1960**, *32*, 300−302.
(37) Matos, J. M. O.; Roos, B. O.; Malmqvist, P. Å. A CASSCF-CCI Study of the Valence and Lower Excited States of the Benzene Molecule. *J. Chem. Phys.* **1987**, *86*, 1458−1466.
(38) Serrano-Andrés, L.; Roos, B. O.; Merchán, M. Theoretical Study of the Electronic Spectra of Cis-1,3,5-Hexatriene and Cis-1,3-Butadiene. *Theor. Chim. Acta* **1994**, *87*, 387−402.
(39) Serrano-Andrés, L.; Merchán, M.; Nebot-Gil, I.; Roos, B. O.; Fülscher, M. Theoretical Study of the Electronic Spectra of Cyclopentadiene, Pyrrole, and Furan. *J. Am. Chem. Soc.* **1993**, *115*, 6184−6197.
(40) Malmqvist, P.-Å.; Roos, B. O. Inclusion of Dynamic $\sigma$-$\pi$ Polarization in $\pi$-Electron Ab Initio Calculations. *Theor. Chim. Acta* **1992**, *83*, 191−199.
(41) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. *Gaussian 16*, Revision A.03; Gaussian Inc., 2016.
(42) Boggio-Pasqua, M.; Bearpark, M. J.; Klene, M.; Robb, M. A. A Computational Strategy for Geometry Optimization of Ionic and Covalent Excited States, Applied to Butadiene and Hexatriene. *J. Chem. Phys.* **2004**, *120*, 7849−7860.
(43) Cave, R. J.; Davidson, E. R. A. Theoretical Investigation of Some Low-Lying Singlet States of 1,3-Butadiene. *J. Phys. Chem.* **1987**, *91*, 4481−4490.
(44) Angeli, C. On the Nature of the $\pi \rightarrow \pi^*$ Ionic Excited States: The V State of Ethene as a Prototype. *J. Comput. Chem.* **2009**, *30*, 1319−1333.
(45) Segarra-Martí, J.; Mukamel, S.; Garavelli, M.; Nenov, A.; Rivalta, I. Towards Accurate Simulation of Two-Dimensional Electronic Spectroscopy. *Top. Curr. Chem.* **2018**, *376*, 24.
(46) Serrano-Andrés, L.; Merchán, M.; Nebot-Gil, I.; Lindh, R.; Roos, B. O. Towards an Accurate Molecular Orbital Theory for Excited States: Ethene, Butadiene, and Hexatriene. *J. Chem. Phys.* **1993**, *98*, 3151−3162.
(47) Hashimoto, T.; Nakano, H.; Hirao, K. Theoretical Study of Valence and Rydberg Excited States of Benzene Revisited. *J. Mol. Struct.: THEOCHEM* **1998**, *451*, 25−33.
(48) Christiansen, O.; Gauss, J.; Stanton, J. F.; Jørgensen, P. The Electronic Spectrum of Pyrrole. *J. Chem. Phys.* **1999**, *111*, 525−537.
(49) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. Benchmarks for Electronically Excited States: Time-Dependent Density Functional Theory and Density Functional Theory Based Multireference Configuration Interaction. *J. Chem. Phys.* **2008**, *129*, 104103.
(50) Harbach, P. H. P.; Wormit, M.; Dreuw, A. The Third-Order Algebraic Diagrammatic Construction Method (ADC(3)) for the Polarization Propagator for Closed-Shell Molecules: Efficient Implementation and Benchmarking. *J. Chem. Phys.* **2014**, *141*, 064113.
(51) Jacquemin, D.; Duchemin, I.; Blase, X. Benchmarking the Bethe-Salpeter Formalism on a Standard Organic Molecular Set. *J. Chem. Theory Comput.* **2015**, *11*, 3290−3304.
(52) Sauer, S. P. A.; Pitzner-Frydendahl, H. F.; Buse, M.; Jensen, H. J. A.; Thiel, W. Performance of SOPPA-Based Methods in the Calculation of Vertical Excitation Energies and Oscillator Strengths. *Mol. Phys.* **2015**, *113*, 2026−2045.


DOI: 10.1021/acs.jpca.9b03715
*J. Phys. Chem. A* XXXX, XXX, XXX−XXX