

© <2019>. This manuscript version is made available under the CC-BY-NC-ND 4.0 license  
<http://creativecommons.org/licenses/by-nc-nd/4.0/>  
The definitive publisher version is available online at <https://doi.org/10.1016/j.cplett.2019.136803>

# Multistate density functional theory applied with 3 unpaired electrons in 3 orbitals: the singdoublet and tripdoublet states of the ethylene cation

Likun Yang<sup>a</sup>, Adam Grofe<sup>b</sup>, Jeffrey R. Reimers<sup>ac</sup>, Jiali Gao<sup>def</sup>

<sup>a</sup> International Centre for Quantum and Molecular Structures and the Department of Physics, Shanghai University, Shanghai, 200444, China

<sup>b</sup> Laboratory of Theoretical and Computational Chemistry, Institute of Theoretical Chemistry, Jilin University, Changchun, Jilin Province 130023, China

<sup>c</sup> School of Mathematical and Physical Sciences, University of technology Sydney, NSW, 2007, Australia

<sup>d</sup> Shenzhen Bay Laboratory, Shenzhen 518055, China

<sup>e</sup> Peking University Shenzhen Graduate School, Shenzhen 518055, China

<sup>f</sup> Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455, United States.

Email: [Jeffrey.Reimers@uts.edu.au](mailto:Jeffrey.Reimers@uts.edu.au), [jiali@jialigao.org](mailto:jiali@jialigao.org)

## Abstract

Multistate density functional theory (MSDFT) is extended to facilitate treatment of situations involving more than two open-shell electrons. The method is applied to determine energies for the two doublet state (tripdoublet and singdoublet) and the quartet-state components that arise when two electrons of one spin type and one electron of the other singly occupy three orbitals. A test system, the  $(\pi, \pi^*)$  excitation of the ethylene cation, is utilized, with MSDFT delivering energies that are numerically superior to those from time-dependent density-functional theory (TD-DFT) and states free from spin contamination.

Kohn-Sham [1] density functional theory (DFT) of electronic structure is the most widely used theory in quantum chemistry. Its remaining weaknesses often involve the effects of strong electron correlation; of concern herein is static

correlation where more than one determinant is required to provide the simplest description of the electron density [2]. For instance, currently no known functional in DFT is able to simultaneously describe the dissociation of the ground state of  $\text{H}_2^+$  and  $\text{H}_2$  [2-5], with time-dependent DFT (TDDFT) not able to describe excited states of  $\text{H}_2$  past the ground-state singlet-triplet instability [6]. However, Wang and Ziegler were able to treat the dissociation limit of  $\text{H}_2$  using spin-flip TDDFT with a non-collinear exchange correlation potential, but at the cost of poorly representing the bonding region [7]. For some problems, an improved exchange-correlation functional could help [2, 8], but a more general strategy is to consider multi-reference approaches.

Several DFT methods that in some way embody multi-reference character have been recently presented [8-17]. The best-known of these methods is constrained DFT (CDFT), introduced by Dederichs, Błgel, Zeller and Akai [18]. In CDFT, one can set the total charge or spin to a particular value in certain regions of space. Thus, by using CDFT, one can construct different configurations, followed by a configuration interaction (CI) calculation, making this a multi-reference DFT (MRDFT) calculation. This procedure describes well the dissociation of  $\text{H}_2^+$ ,  $\text{H}_2$  and LiF, also it provides a natural way for the description of electron-transfer process [8, 19-21]. A shortcoming of CDFT is that it does not work well when the fragments are strongly overlapping as the constraint formulas are less well defined.

Very recently Plaisance, Van Santen and Reuter [9] reported a method called Constrained-Orbital Density Functional Theory (CO-DFT). In contrast to CDFT, in CO-DFT certain Kohn-Sham orbitals are confined to single atoms. A plane-wave basis set was used in this method, and it was implemented in the Vienna Ab-initio Simulation Package (VASP). They also showed that multi-configurational Kohn-Sham calculations can be performed using CO-DFT to construct configuration states.

Another well-known multi-reference DFT is multi-configuration pair-density functional theory (MC-PDFT). Manni et al. [22] developed it by combining multi-configuration wavefunction theory and density functional theory. Put simply, the

authors compute a multiconfigurational wave function using standard methods. Then the density and pair density are translated into variables that conventional DFT functionals use, and these are used to compute the remaining correlation energy. By including the DFT contribution after optimization, this method is inherently non-variational. They showed that MC-PDFT can be as accurate as CASPT2 for electronic excitations [23].

Gao, Grofe, Ren and Bao [11] described a multistate density functional theory (MSDFT), in which block-localized Kohn-Sham orbitals are optimized to construct basis states for non-orthogonal configuration interaction calculations [24, 25]. Similar to CDFT, block localized Kohn-Sham configurations can limit the effect of self-interaction error, but remains well defined for strongly overlapping units. MSDFT can be used to study electron transfer processes and other chemical reactions as diabatic states can be easily represented using it [10, 26, 27]. Furthermore, MSDFT has been used to model the spin multiplet components of atoms and molecules to yield strict degeneracy between the high and low spin components [28]. Very recent applications of MSDFT include Lewis-base catalysed hydrogen-atom transfer by Asgari et al. [29], and C-H bond activation in a non-heme iron oxo complex [30].

In this work, we develop new applications for DFT approaches to systems involving strongly overlapping orbitals and coupled spin-multiplet components. In particular, this approach has been shown to deliver mean absolute errors of 0.045 eV for the singlet-triplet splittings of main-group divalent radicals [28], much lower than similar errors reported for MC-PDFT [31, 32].

The details of MSDFT have been presented previously [10-12, 28]. For completeness, we briefly outline the method. MSDFT treats the multireference problem by taking a linear combination of configurations and uses Kohn-Sham density functional theory to variationally optimize a set of determinant configurations. The density for state  $I$  is given as a weighted combination of densities ( $\rho_A$ ) for basis configurations and transition densities ( $\rho_{AB}$ ):

$$\rho_I = \sum_A^{N^D} |C_{AI}|^2 \rho_A + \sum_{A \neq B}^{N^D} C_{AI} C_{BI} \rho_{AB} \quad (1)$$

where  $C_{AI}$  is the configuration coefficient for MSDFT state  $I$  with  $N^D$  basis determinants,  $A$ . The configuration coefficients can then be determined by solving the generalized secular equation

$$\mathbf{H}^{MS} \mathbf{C} = \mathbf{E} \mathbf{S} \mathbf{C} \quad (2)$$

where  $\mathbf{H}^{MS}$  is the MSDFT Hamiltonian,  $\mathbf{S}$  is the overlap matrix and  $\mathbf{E}$  is the diagonal matrix of eigenvalues (also the MSDFT energies). The elements of the MSDFT Hamiltonian are

$$H_{AA}^{MS} = E^{KS}[\rho_A] \quad (3)$$

$$H_{AB}^{MS} = E^{TDF}[\rho_A, \rho_A, \rho_{AB}] \quad (4)$$

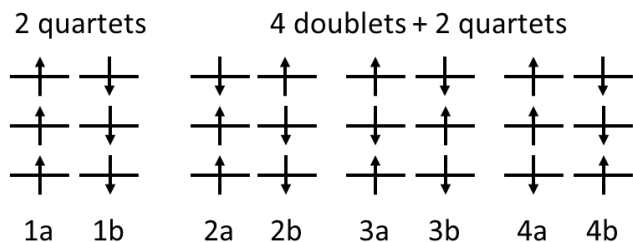
where  $E^{KS}[\rho_A]$  is the energy of Kohn-Sham DFT from a single-determinant for state  $A$  and  $E^{TDF}$  is a new class of density functional of the transition density and those of the interacting states [11, 26, 27]. As for the Kohn-Sham density functional approximation to the exchange-correlation functional, the exact transition density functional in MSDFT is not known. Therefore, an approximate form of the transition density functional has been used:

$$H_{AB}^{MS} = H_{AB}^{WFT} + \frac{1}{2} S_{AB} (E^{corr}[\rho_A] + E^{corr}[\rho_B]) , \quad (5)$$

where the first term is the wave function theory (WFT) off-diagonal element, and the second term is the overlap weighted average of the correlation energies. Alternatively, a correlation energy weighted approximation has been used [11, 26, 27]. Here, the correlation energy is defined as the energy difference using DFT and WFT with the same set of orbitals.

In this work, we continue [28] an extension of MSDFT to provide a general form of the transition density functional for spin-coupled determinants. The simplest such class is doublet states made with two electrons of one type (e.g., spin-up,  $\alpha$ ) and one of the other (spin-down,  $\beta$ ). In total, when 3 unpaired electrons are placed in 3 orbitals,  $2^3 = 8$  spin combinations are possible, as illustrated in Fig. 1. In general, not all configurations in Fig. 1 are eigenstates of the  $S^2$  operator, which depicts the total spin of

the system. Three of the eight spin components have  $m_z = \frac{1}{2}$ , and their interaction leads to one quartet state and two doublet states, with corresponding spin-flipped states for  $m_z = -\frac{1}{2}$ . Two of the four components of the quartet state are specified simply in terms of all  $\alpha$  (1a) or all- $\beta$  (1b) electrons. For each, the Gunnarsson-Lundqvist theorem [3] applies, meaning that standard DFT is well-defined to compute the energy of the quartet state; as the components are degenerate, using either 1a or 1b produces the same results. However, the other 6 determinants shown in Fig 1 interact with each other, and the Gunnarsson-Lundqvist theorem [3] does not apply, offering no physical meaning (i.e., spin contaminated states) to the results coming from DFT calculations that set the electronic occupancy to any of 2a, 2b, 3a, 3b, 4a, or 4b. Application of MSDFT allows the energies of the two doublet states and the quartet state to be determined, processing either the determinant sets 2a, 3a, and 4a or else 2b, 3b, and 4b.



**Fig. 1.** The 8 spin components possible for 3 unpaired electrons in 3 orbitals interact with each other to produce 4 degenerate components of a quartet state and two sets of two degenerate components of doublet states (the tripdoublet and the singdoublet).

The two doublet states from combinations of the three configurations are usually termed the *tripdoublet* state and the *singdoublet* state, based on the forms of the wavefunctions generated in a Hartree-Fock based multi-configurational approach. The meaning of these names becomes clear through consideration of the interaction of a closed-shell molecule with a molecule in a doublet state [33]. The closed-shell molecule can display singlet and triplet spectroscopic transitions, and, if the other molecule is infinitely separated, the transition energies (e.g., for fluorescence and

phosphorescence) will occur at the standard gas-phase values. The singlet fluorescence is spin-allowed and may be bright, whereas the triplet phosphorescence is spin-forbidden and will be weak. As the radical is brought close to the molecule, the combined states of the dimer become the tripdoublet and the singdoublet states. The singdoublet corresponds to the asymptotic singlet state whilst the tripdoublet corresponds to the asymptotic triplet state. Bringing the two molecules into van der Waals contact perturbs the energies of each transition only slightly, but now the once-forbidden tripdoublet emission can become very intense, owing to intermolecular interactions. Both singdoublet and tripdoublet transitions are formally spin-allowed. Tripdoublet transitions can be very important in molecular materials carrying charge, for example the special-pair radical cation in photosynthesis [34, 35].

The energies of the spin projections of quartet states involving multiple determinants should exactly equal the energies determined from single-determinants 1a or 1b. In WFT, this is achieved as determinants 2a through 4b are coupled through the exchange operator due to the Slater-Condon rules. By analogy with WFT, we derive effective spin-coupling terms using DFT that guarantee this basic result. In WFT, the total energy for the determinants in Figure 1 is [36]

$$E^{WFT} = \sum_i^3 h_{ii} + \sum_{j<i}^3 J_{ij} - \sum_{j<i}^3 K_{ij} \delta(s_i - s_j) \quad (6)$$

where  $h_{ii}$  are the one-electron energies of spin orbital  $i$ ,  $J_{ij}$  and  $K_{ij}$  are the electrostatic repulsion and exchange integrals between orbitals  $i$  and  $j$ , respectively, and  $s_i$  is the spin coordinate.  $\delta(s_i - s_j)$  is the Dirac delta function, which ensures that the exchange interaction is only included for same spin-electrons. Then the energy difference between determinants 2a and 1a is

$$\Delta E(2a, 1a) = E^{WFT}(2a) - E^{WFT}(1a) = - \sum_{j<i}^3 K_{ij} \left( \delta(s_i^{2a} - s_j^{2a}) - \delta(s_i^{1a} - s_j^{1a}) \right) \quad (7)$$

where we have included a superscript to denote the determinant from which the spin orbital originates. Thus, the energy difference is only a linear combination of exchange integrals. Here, we use the minimal number of electrons to simplify the

formulas for presentation. However, this result includes systems that have closed shell core electrons as these terms would be included in both determinants, and would therefore cancel out. Considering only determinants 2a, 3a and 4a, we find that there are three energy differences using determinant 1a as a unified reference and three exchange integrals. Thus, we can define a linear system of equations

$$\Delta E^{WFT} = -\boldsymbol{\varepsilon} \mathbf{K}^{WFT} \quad (8)$$

where  $\Delta E^{WFT}$  and  $\mathbf{K}^{WFT}$  are 3 dimensional vectors, and  $\boldsymbol{\varepsilon}$  is a 3 x 3 matrix of 1's and 0's.  $\varepsilon_{ij}$  has a value of 0 when both determinants have the same exchange interaction and 1 when the determinants do not. It should be noted that  $\Delta E^{WFT}$  indexes over the determinants, and  $\mathbf{K}^{WFT}$  indexes over molecular orbital integrals. The order of the exchange integrals in  $\mathbf{K}^{WFT}$  and the order of the determinants in  $\Delta E^{WFT}$  determine the structure of  $\boldsymbol{\varepsilon}$ . Equation 8 is an exact result from WFT, but it is not necessarily true in a multistate density functional theory because of contributions due to dynamic correlation. This property of linearity of the exchange operator is a fundamental reason as to why there is an energy degeneracy between the high spin and low spin components because there is an exact addition and cancellation of terms. KS-DFT approximations often use nonlinear terms with respect to the density to compute the exchange and correlation. Thus, there is no guarantee that the same energy degeneracy would result from a simple spin coupling scheme in DFT. Using Equation 8, it is possible to impose the correct symmetry in DFT by replacing the WFT terms with a density functional approximation. Then, multiplying on the left by  $\boldsymbol{\varepsilon}^{-1}$  yields the MSDFT spin-couplings,

$$\mathbf{K}^{DFT} = -\boldsymbol{\varepsilon}^{-1} \Delta E^{DFT} \quad (9)$$

where  $\mathbf{K}^{DFT}$  is an effective exchange-correlation integral of functionals of two determinants. The form of an a priori theoretical definition of the density dependence of  $\mathbf{K}^{DFT}$  is not yet known and requires further research. However, we can define its value using Eqn. 9 on the basis of energy constraints of spin projections. Thus, the spin-couplings can be derived from simple energy differences in the high spin and low spin determinants. Fundamentally, because they were derived from



DFT energies, the spin-couplings include dynamic correlation. Then the MSDFT coupling between determinants 2a and 4a as an example is

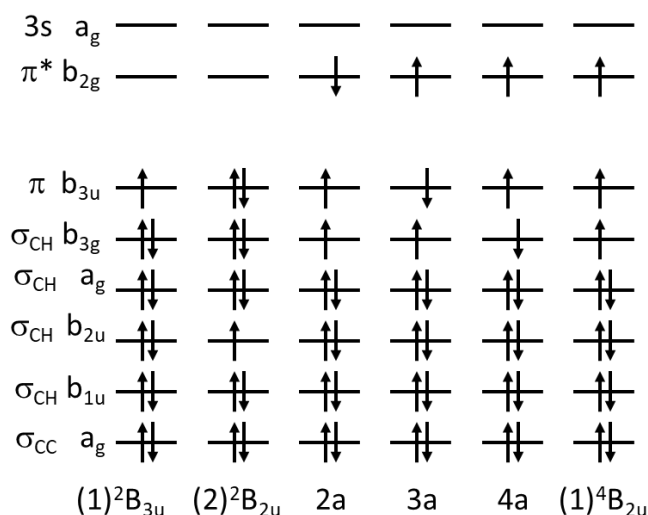
$$H_{2a,4a}^{MS} = - \sum_{j<i}^3 K_{ij}^{DFT} \left[ \left( 1 - \delta(s_i^{2a} - s_i^{4a}) \right) \left( 1 - \delta(s_j^{2a} - s_j^{4a}) \right) \right] = -K_{12}^{DFT} \quad (10)$$

where only a single exchange integral remains. Equations 9 and 10 constitute a general method for computing the spin coupling in MSDFT. For simplicity, we presented the equations using three spin-coupled determinants, but the algorithm cannot be easily extended to the  $N^D$  determinant case. The total number of unique configurations is the  $2^{N^{SC}} - 2$  where  $N^{SC}$  is the number of spin coupled orbitals, which will always be larger than the number of couplings  $\left( \frac{1}{2} N^{SC} (N^{SC} - 1) \right)$ . The antiparallel high spin configuration (e.g.  $m_z = -\frac{3}{2}$ ) cannot be used as a separate configuration in Eqn. 9 as this would result in  $\epsilon$  being a singular matrix. Therefore, it is possible that the system of equations is solvable for the  $N$  dimensional case. Equation 10 does not enforce the Slater-Condon rules, but they can easily be enforced using the corresponding orbitals method [37]. In the three-determinant case, this is not a problem. Additionally, Eqn. 9 recovers similar formulae to those presented by Grofe et al., but does not require the computation of Hartree-Fock energies for computing the transition density functional [28].

As a simple example system, we consider the  $(\pi, \pi^*)$  excitation of the ethylene cation. Key determinants involved are indicated in Fig. 2. The ground state of the cation,  $(1)^2B_{3u}$ , involves oxidation from the  $b_{3u}$   $\pi$  orbital. The states of primary interest, corresponding to those in Fig. 1, involve  $(\sigma_{CH}, \pi^*)$   $b_{3g} \rightarrow b_{2g}$  excitation, making either the  $(1)^4B_{2u}$  quartet state or the  $^2B_{2u}$  singdoublet or tripdoublet states. Another state of relevance is a low-lying state named  $(2)^2B_{2u}$  of  $(\sigma_{CH}, \pi)$  type.

MSDFT calculations required the specification of input orbitals depicting the determinants to be included in the calculations, and the method has been implemented into a modified version of the GAMESS(US) programme [10, 12, 38], using orbitals generated by Gaussian16 [39]. Results are not strongly dependent of the choice of

initial orbitals, with those for the spin-restricted quartet state used herein. Use of spin-restricted calculations is essential to ensure that different spin-state components evaluate to the same energy. We report two different MSDFT calculations; the first is called “MSDFT(3)” that includes excited states 2a, 3a, and 4a only, and the other is “MSDFT(4)” with the additional  $(2)^2B_{2u}$  state.



**Fig. 2.** Determinants used in constructing the  $(1)^2B_{3u}$  ground-state,  $(1)^4B_{2u}$  quartet state,  $(2)^2B_{2u}$  state, and the  $^2B_{2u}$  singdoublet and tripdoublet states of the ethylene cation.

MSDFT calculations are performed using the PBE0 density functional [40] and the aug-cc-pVDZ [41] basis set. As Fig. 2 shows, the next-highest orbital above the  $\pi^*$  LUMO is the C 3s Rydberg bonding orbital, and hence details of the results and technical issues concerning the feasibility with which the valence states of interest can be obtained depend on basis set size. Table 1 compares MSDFT results for this basis set, as well as for cc-pVDZ and cc-pVTZ [41], showing only minor changes. Larger basis sets proved impractical; results obtained using the minimal STO-3G basis vary by less than 1 eV from the aug-cc-pVDZ ones. The PBE0/aug-cc-pVDZ optimized geometry of ethylene is used in all calculations.

**Table 1.** Basis set dependence of PBE0 MSDFT energies, in eV, for the ethylene cation relative to its  $(1)^2B_{3u}$  ground state.

State	cc-pVDZ	aug-cc-pVDZ	cc-pVTZ
$(1)^4B_{2u}$	6.62	6.61	6.61
$^2B_{2u}$ tripdoublet	7.27	7.26	7.22
$^2B_{2u}$ singdoublet	10.60	10.57	10.53

Table 2 compares the aug-cc-pVDZ MSDFT results to data obtained using complete-active-space self-consistent field (CASSCF) calculations, multi-reference configuration interaction (MRCI), EOM-CCSD, TD-DFT, and full configuration-interaction (FCI) calculations using the semi-empirical intermediate neglect of differential overlap (INDO-FCI) [42, 43] method, truncated at states of energy 80 eV. The CASSCF and MRCI calculations use a CAS(11,7) active space in which all valence electrons are distributed amongst all valence orbitals of lower energy than the first Rydberg orbital.

**Table 2.** Comparison of excited-state energies  $\Delta E$ , for the ethylene cation, and spin eigenvalues  $S^2$  for methods embodying spin contamination (aug-cc-pvDZ basis, PBE0 used in MSDFT).

State	$\Delta E / \text{eV}$							$S^2$	
	CAS(11,7)	MRCI(11,7)	MSDFT(3)	MSDFT(4)	EOM-CCSD	TD-DFT	INDO-FCI	HF/CIS <sup>c</sup>	TD-DFT
$(1)^2B_{3u}$	[0]	[0]	[0]	[0]	[0]	[0]	[0]	0.755	0.752
$(1)^4B_{2u}$	7.56	7.25	6.61 <sup>a</sup>	6.61 <sup>a</sup>	7.17		7.81		
$(2)^2B_{2u}$	6.16	5.85	-	3.85 <sup>b</sup>	5.86	5.57	8.61	1.50	1.08
tripdoublet	8.52	8.01	7.26	7.30	7.56	6.94	6.97	1.32	1.73
singdoublet	13.55	12.07	10.57	12.68	8.59	8.93	11.49	2.75	1.45
splitting	5.03	4.06	3.31	5.38	1.03 <sup>d</sup>	1.99 <sup>d</sup>	4.52	-	-

a: the same result is obtained by DFT evaluation of the energy of 1a.

b: 5.81 eV by DFT evaluation of the energy of the single-determinants.

c: For doublet states without spin contamination,  $S^2 = 3/4$ ; EOM-CCSD are based on a Hartree-Fock (HF) reference, initiated by single-excitation singles (CIS) calculations, only for which  $S^2$  values are available; EOM-CCSD values will show improvements upon these.

d: 0.63 eV (EOM-CCSD) and 2.65 eV (TD-DFT) using the STO-3G basis set hence the low values are not caused by interactions with Rydberg states.

The most accurate transition energies in Table 2 are believed to be the MRCI ones, listing 5.85 eV for  $(2)^2B_{2u}$ , 7.25 eV for  $(1)^4B_{2u}$ , 8.01 eV for the tripdoublet state and 12.07 eV for the singdoublet. Of these results, the only sizable change from the CASSCF results is -1.48 eV for the singdoublet, and it would therefore appear that most states are accurately evaluated. The EOM-CCSD results for  $(1)^4B_{2u}$  and  $(2)^2B_{2u}$  are very similar to the MRCI ones, but for the singdoublet there is a significant difference of 4.96 eV. The calculated singdoublet-tripdoublet energy difference (Table 2) is 4.06 eV from MRCI but only 1.03 eV from EOM-CCSD. Using the minimal STO-3G basis set, the EOM-CCSD value is 0.63 eV and hence the unexpectedly low result does not arise from interactions with Rydberg states that may not properly included in the MRCI. It is possible that spin-flip EOM-CCSD would yield better values for this excitation, but we will leave that for a future work [44]. TD-DFT also gives an anomalously low singdoublet energy, with a splitting of just 1.99 eV. The poor results from EOM-CCSD and TD-DFT are most likely indicative of the difficulty in treating open-shell states with techniques that embody single-reference characteristics. As indicated in Table 2, these methods are associated with excited-state wavefunctions depicting significant spin contamination, questioning the reliability of non-energetic properties determined by them.

Using MSDFT(3) and MSDFT(4), the energy of the quartet state is predicted to be 6.61 eV, identical to the value obtained by direct computation of the density associated with determinant 1a. This is an important indicator suggesting that the MSDFT results for the doublet states accurately portray the properties of the density-functional used in the calculations. These values are 0.7 eV less than the MRCI value of 7.25 eV, however. Similarly, the MSDFT(3) and MSDFT(4) energies for the tripdoublet state are 0.7 eV less than the MRCI value. Most significantly, the MSDFT(4) energy for  $(2)^2B_{2u}$  is 2 eV less than the MRCI value, with that for the singdoublet being 0.6 eV higher (and 2.1 eV higher than the MSDFT(3) result). This energy lowering arises from strong interactions perceived by MSDFT between it and, in particular, the singdoublet state.

If this interaction is realistically represented, then an explanation for the observed error could be the absence of other states in the MSDFT that would similarly influence the ground state, a recognized problem with MCSCF calculations that inspired the development of CASSCF theory. Of course, computational results from MSDFT are dependent on the particular functionals used, which is PBE0 in the present study. Nevertheless, the singdoublet-tripdoublet splitting is 3.31 eV from MSDFT(3) and 5.38 eV from MSDFT(4), both closer to the MRCI value of 4.06 eV than are the EOM-CCSD and TD-DFT results.

INDO is an efficient semi-empirical wavefunction-based method that when implemented using MRCI would be competitive with MSDFT for many technological applications, much more so than either EOM-CCSD or TD-DFT. Here using a CAS(11,12) reference that is equivalent to FCI, it captures critical static electron correlation and therefore gives much better result of 4.52 eV for the singdoublet-tripdoublet splitting, but it also predicts  $(2)^2B_{2u}$  to be too high by 3.03 eV, indicating that this method lacks general reliability. Expanding the CI from CAS(5,4) to CAS(11,7) to FCI makes little change to the results, however.

The ethylene cation may be considered to be a worst-case scenario as the orbitals involved in the transitions are all strongly localized, making the energy scale quite large. Nevertheless, MSDFT is shown to make semi-quantitative predictions for its lowest-energy tripdoublet and singdoublet states. We found that energies are closer to values from MRCI than those from TD-DFT. It is perhaps the most generally reliable method that could routinely be applied to free radical systems of technological relevance. In addition to the choice of a Kohn-Sham density functional approximation for each determinant configuration, their orbitals could also be individually optimized, whereas a single set of orbitals are used in the present study. In general, it would appear necessary to ensure that MSDFT results are converged with respect to the number of states included in the calculations, and these are the minimum number of states for a given set of spin complement. This issue is also important for multi-reference applications of wavefunction theory based on Hartree-Fock reference states, leading to

the development of CASSCF approaches so as to minimise empiricism. Such effects are expected to be independent of the density-functional used as they address issues concerning the Kohn-Sham or its extension, the Gunnarsson-Lundqvist theorem [3], rather than density-functional design.

### **Data Availability**

Input scripts and sample output needed to run the MSDFT code in GAMESS, as well as the interface code to Gaussian16, are included in Data In Brief associated with this manuscript.

### **Acknowledgments**

We thank the National Natural Science Foundation of China for supporting this research under grants 11674212 and 91541124, and the Ministry of Science and Technology under grant 2018YFB0203400.

### **References**

- [1] W. Kohn, L.J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Physical Review* 140 (1965) A1133–A1138.
- [2] A.J. Cohen, P. Mori-Sánchez, W. Yang, Challenges for density functional theory, *Chemical Reviews* 112 (2012) 289–320.
- [3] O. Gunnarsson, B.I. Lundqvist, *Phys. Rev. B* 13 (1976) 4274.
- [4] A.M. Lee, N.C. Handy, Dissociation of hydrogen and nitrogen molecules studied using density functional theory, *Journal of the Chemical Society, Faraday Transactions* 89 (1993) 3999–4003.
- [5] R. Neumann, R.H. Nobes, N.C. Handy, Exchange functionals and potentials, *Molecular Physics* 87 (1996) 1–36.
- [6] Z.-L. Cai, J.R. Reimers, Application of time-dependent density-functional theory to the  $^3\Sigma^-$  first excited state of  $\text{H}_2$ , *J. Chem. Phys.* 112 (2000) 527.
- [7] F. Wang, T. Ziegler, Time-dependent Density Functional Theory Based on a Noncollinear Formulations of the Exchange-Correlation Potential, *J. Chem. Phys.* 121 (2004) 12191–12196.
- [8] B. Kaduk, T. Kowalczyk, T. Van Voorhis, Constrained density functional theory, *Chemical Reviews* 112 (2012) 321–370.

- [9] C.P. Plaisance, R. A. Van Santen, K. Reuter, Constrained-Orbital Density Functional Theory. Computational Method and Applications to Surface Chemical Processes, *J. Chem. Theory Comput.* 13 (2017) 3561-3574.
- [10] H. Ren, M.R. Provorse, P. Bao, Z. Qu, J. Gao, Multistate Density Functional Theory for Effective Diabatic Electronic Coupling, *J. Phys. Chem. Lett.* 7 (2016) 2286-2293.
- [11] J. Gao, A. Grofe, H. Ren, P. Bao, Beyond Kohn-Sham Approximation: Hybrid Multistate Wave Function and Density Functional Theory, *J. Phys. Chem. Lett.* 7 (2016) 5143-5149.
- [12] A. Cembran, L. Song, Y. Mo, J. Gao, Block-localized density functional theory (BLDFT), diabatic coupling, and their use in valence Bond theory for representing reactive potential energy surfaces, *J. Chem. Theory Comput.* 5 (2009) 2702-2716.
- [13] A.B.J. Parusel, S. Grimme, A Theoretical Study of the Excited States of Chlorophyll a and Pheophytin a, *J. Phys. Chem. B* 104 (2000) 5395-5398.
- [14] B.B. Miehlich, H. Stoll, A. Savin, A correlation-energy density functional for multideterminantal wavefunctions, *Mol. Phys.* 91 (1997) 527-536.
- [15] J. Gräfenstein, D. Cremer, The combination of density functional theory with multi-configuration methods - CAS-DFT, *Chem. Phys. Lett.* 316 (2000) 569-577.
- [16] J. Gräfenstein, D. Cremer, Can Density Functional Theory Describe Multi-reference Systems? Investigation of Carbenes and Organic Biradicals, *Phys. Chem. Chem. Phys.* 2 (2000) 2091-2103.
- [17] J. Gräfenstein, D. Cremer, The Self-interaction Error and the Description of Non-dynamic Electron Correlation in Density Functional Theory, *Theor. Chem. Acc.* 123 (2009) 171-182.
- [18] P.H. Dederichs, S. Blögel, R. Zeller, H. Akai, Ground states of constrained systems: Application to cerium impurities, *Phys. Rev. Lett.* 53 (1984) 2512-2515.
- [19] Q. Wu, C.L. Cheng, T. Van Voorhis, Configuration interaction based on constrained density functional theory: A multireference method, *J. Chem. Phys.* 127 (2007).
- [20] Q. Wu, B. Kaduk, T. Van Voorhis, Constrained density functional theory based configuration interaction improves the prediction of reaction barrier heights, *J. Chem. Phys.* 130 (2009).
- [21] Q. Wu, T. Van Voorhis, Constrained density functional theory and its application in long-range electron transfer, *J. Chem. Theory Comput.* 2 (2006) 765-774.
- [22] G.L. Manni, R.K. Carlson, S. Luo, D. Ma, J. Olsen, D.G. Truhlar, L. Gagliardi, Multiconfiguration pair-density functional theory, *J. Chem. Theory Comput.* 10 (2014) 3669-3680.
- [23] C.E. Hoyer, S. Ghosh, D.G. Truhlar, L. Gagliardi, Multiconfiguration Pair-Density Functional Theory Is as Accurate as CASPT2 for Electronic Excitation, *J. Phys. Chem. Lett.* 7 (2016) 586-591.
- [24] Y. Mo, L. Song, Y. Lin, Block-localized wavefunction (BLW) method at the density functional theory (DFT) level, *J. Phys. Chem. A* 111 (2007) 8291-8301.
- [25] Y. Mo, J. Gao, An ab initio molecular orbital-valence bond (MOVb) method for simulating chemical reactions in solution, *J. Phys. Chem. A* 104 (2000) 3012-3020.
- [26] X. Guo, Z. Qu, J. Gao, The charger transfer electronic coupling in diabatic perspective: A multi-state density functional theory study, *Chem. Phys. Lett.* 691 (2018) 91-97.

- [27] A. Grofe, Z. Qu, D.G. Truhlar, H. Li, J. Gao, Diabatic-At-Construction Method for Diabatic and Adiabatic Ground and Excited States Based on Multistate Density Functional Theory, *J. Chem. Theory Comput.* 13 (2017) 1176-1187.
- [28] A. Grofe, X. Chen, W. Liu, J. Gao, Spin-multiplet components and energy splittings by multistate density functional theory, *J. Phys. Chem. Lett.* 8 (2017) 4838-4845.
- [29] P. Asgari, Y. Hua, A. Bokka, C. Thiamsiri, W. Prasitwatcharakorn, A. Karedath, X. Chen, S. Sardar, K. Yum, G. Leem, B.S. Pierce, K. Nam, J. Gao, J. Jeon, Catalytic hydrogen atom transfer from hydrosilanes to vinylarenes for hydrosilylation and polymerization, *Nature Catalysis* 2 (2019) 164-173.
- [30] L. Yang, X. Chen, Z. Qu, J. Gao, Combined Multistate and Kohn-Sham Density Functional Theory Studies of the Elusive Mechanism of N-Dealkylation of N,N-Dimethylanilines Mediated by the Biomimetic Nonheme Oxidant FeIV(O) (N4Py) (ClO4)2, *Frontiers in Chemistry* 6 (2018).
- [31] J.L. Bao, A. Sand, L. Gagliardi, D.G. Truhlar, Correlated-Participating-Orbitals Pair-Density Functional Method and Application to Multiplet Energy Splittings of Main-Group Divalent Radicals, *J. Chem. Theory Comput.* 12 (2016) 4274-4283.
- [32] T. Scott, R. Nieman, A. Luxon, B. Zhang, H. Lischka, L. Gagliardi, C.A. Parish, A Multireference Ab Initio Study of the Diradical Isomers of Pyrazine, *J. Phys. Chem. A.* 123 (2019) 2049-2057.
- [33] J.R. Reimers, N.S. Hush, The influence of spin-forbidden monomer excitations on the spin-allowed electron-transfer and electron-localized states of mixed-valence and single-valence dimeric systems, *Inorg. Chim. Acta* 226 (1994) 33.
- [34] J.R. Reimers, N.S. Hush, Understanding the observed Stark spectra, midpoint potential versus degree of charge localization, and intervalence transition energies of the special-pair radical cation of *Rhodobacter sphaeroides* and its mutant strains, *J. Am. Chem. Soc.* 126 (2004) 4132-4144.
- [35] J.R. Reimers, M. Biczysko, D. Bruce, D.F. Coker, T.J. Frankcombe, H. Hashimoto, J. Hauer, R. Jankowiak, T. Kramer, J. Linnanto, F. Mamedov, F. Müh, M. Rätsep, T. Renger, S. Styring, J. Wan, Z. Wang, Z.-Y. Wang-Otomo, Y.-X. Weng, C. Yang, J.-P. Zhang, A. Freiberg, E. Krausz, Challenges facing an understanding of the nature of low-energy excited states in photosynthesis, *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1857 (2016) 1627-1640.
- [36] A. Szabo, N.S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications Inc. 1996.
- [37] H.F. King, R.E. Stanton, H. Kim, R.E. Wyatt, R.G. Parr, Corresponding Orbitals and the Nonorthogonality Problem in Molecular Quantum Mechanics, *J. Chem. Phys.* 47 (1967) 1936-1941.
- [38] M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, General Atomic and Molecular Electronic Structure System, *J. Comput. Chem* 14 (1993) 1347-1363.
- [39] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A.V. Marenich, J. Bloino, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J.V. Ortiz, A.F. Izmaylov, J.L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng,



A. Petrone, T. Henderson, D. Ranasinghe, V.G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M.J. Bearpark, J.J. Heyd, E.N. Brothers, K.N. Kudin, V.N. Staroverov, T.A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A.P. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, J.M. Millam, M. Klene, C. Adamo, R. Cammi, J.W. Ochterski, R.L. Martin, K. Morokuma, O. Farkas, J.B. Foresman, D.J. Fox, Gaussian 16, Gaussian Inc., Wallingford, CT, 2016.

[40] C. Adamo, V. Barone, *J. Chem. Phys.* 110 (1999) 6158.

[41] R.A. Kendall, T.H. Dunning, Jr., R.J. Harrison, Electron affinities of the first-row atoms revisited. Systematic basis sets and wavefunctions, *J. Chem. Phys.* 96 (1992) 6796.

[42] J. Zeng, N.S. Hush, J.R. Reimers, Solvent effects on molecular and ionic spectra. V. Development of a method for simulation of the liquid structure and solvatochromic shift of inorganic complexes such as pentaamminopyridylruthenium(II) in water, *J. Phys. Chem.* 99 (1995) 10459.

[43] B. Tejerina, J.R. Reimers, An INDO-MRCI program for nanotechnology, Nanohub, <http://www.nanohub.org/>, 2008.

[44] A.I. Krylov, Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space, *Annu. Rev. Phys. Chem.* 59 (2008) 433-462.