Seniority number description of potential energy surfaces: Symmetric dissociation of water, N_2, C_2, and Be_2

Laimutis Bytautas,[1] Gustavo E. Scuseria,[2,3] and Klaus Ruedenberg[4]
[1]*Department of Chemistry, Galveston College, 4015 Ave. Q, Galveston, Texas 77550, USA*
[2]*Department of Chemistry, Rice University, Houston, Texas 77005, USA*
[3]*Chemistry Department, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia*
[4]*Department of Chemistry, Iowa State University, Ames, Iowa 50011, USA*

(Received 20 April 2015; accepted 12 August 2015; published online 3 September 2015)

The present study further explores the concept of the seniority number (Ω) by examining different configuration interaction (CI) truncation strategies in generating compact wave functions in a systematic way. While the role of Ω in addressing static (strong) correlation problem has been addressed in numerous previous studies, the usefulness of seniority number in describing weak (dynamic) correlation has not been investigated in a systematic way. Thus, the overall objective in the present work is to investigate the role of Ω in addressing also dynamic electron correlation in addition to the static correlation. Two systematic CI truncation strategies are compared *beyond* minimal basis sets and full valence active spaces. One approach is based on the seniority number (defined as the total number of singly occupied orbitals in a determinant) and another is based on an excitation-level limitation. In addition, molecular orbitals are energy-optimized using multiconfigurational-self-consistent-field procedure for all these wave functions. The test cases include the symmetric dissociation of water (6-31G), N_2 (6-31G), C_2 (6-31G), and Be_2 (cc-pVTZ). We find that the potential energy profile for H_2O dissociation can be reasonably well described using only the $\Omega = 0$ sector of the CI wave function. For the Be_2 case, we show that the full CI potential energy curve (cc-pVTZ) is almost exactly reproduced using either Ω-based (including configurations having up to $\Omega = 2$ in the virtual-orbital-space) or excitation-based (up to single-plus-double-substitutions) selection methods, both out of a full-valence-reference function. Finally, in dissociation cases of N_2 and C_2, we shall also consider novel hybrid wave functions obtained by *a union* of a set of CI configurations representing the full valence space and a set of CI configurations where seniority-number restriction is imposed for a complete set (full-valence-space and virtual) of correlated molecular orbitals, simultaneously. We discuss the usefulness of the seniority number concept in addressing both static and dynamic electron correlation problems along dissociation paths. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4929904]

I. INTRODUCTION

In recent years, there have been a number of novel developments[1–52] in *ab initio* theory for reducing the computational cost of quantum chemistry calculations while maintaining a good quality description of molecular systems undergoing chemical reactions. An essential component of high-quality studies has to do with the accurate description of the electron correlation phenomenon. The electron correlation energy is typically defined as the deviation in energy from the single Hartree-Fock (HF) determinant compared to the exact electronic energy calculated at full configuration interaction (FCI) level in a given basis set. Ordinarily, there are two kinds of electron correlation that are recognized: strong (static or non-dynamic) correlation and weak (dynamic) correlation. Within this context, some reports[53–56] suggest that it is more informative to partition the electron correlation into *three categories*: static, nondynamic, and dynamic. For instance, static electron correlation corresponds to contributions of additional determinants to account for proper spin symmetries and their interactions, while nondynamic electron correlation reflects the contribution of additional determinants to allow a proper dissociation of a molecule into its fragments.[53–56] The dynamic electron correlation represents a contribution of determinants, typically with very small weights, in a wave function that reduces the probability of any two electrons being close to each other. The separation of electron correlation into dynamic and nondynamic components has been useful for exploring processes involving breaking or formation of chemical bonds.[56,57] While non-dynamic electron correlation is critical for, at least qualitatively, the correct description of bond-breaking processes and can be provided by very compact wave functions,[58–60] the inclusion of dynamic electron correlation is essential in cases when high-accuracy is necessary, e.g., when near-spectroscopic quality[8,11,12,14] of predicted vibrational energy levels is required or to precisely determine the location of a tiny barrier[9] along the dissociation coordinate in the $^1\Sigma_g^+$ ground-state of F_2 (see also Ref. 13 for a more recent study). Within the framework of configuration interaction (CI) wave functions,[1,2] the compactness of CI expansion is highly significant. Recent advances in this direction have been documented in the literature and can be

categorized as (1) CI configuration selection schemes based on energy-criterion (see, e.g., Refs. 18–22) or (2) configuration-selection schemes that focus on nature of orbital occupation in an determinant rather than its energy-contribution in a wave function.[23,39] Examples of CI methods of the second type include CI truncations based on the restriction of a highest excitation (substitution) level[2,61] from a reference wave function or a seniority-number(Ω)-based criterion, which restricts CI configurations based on the highest seniority number of a determinant. The seniority number is defined (see Fig. 1 of Ref. 39 for a graphical illustration) as the total number of singly occupied orbitals in a determinant.

In this study, we focus on the second-type of configuration selection schemes, namely, investigating the usefulness of seniority number in the case of the symmetric dissociation of H_2O, N_2, C_2, and Be_2. While the utility of the seniority number in CI wave functions has been investigated recently by several groups,[39–44] the primary focus was centered on the static correlation problem. Thus, in this work, we expand the size of the active space *beyond* full valence active space and minimal basis sets considerably broadening the scope of investigations focusing on the role of Ω as compared to earlier studies. Thus, here we explore the performance of CI configuration restriction schemes with full orbital (multiconfigurational-self-consistent-field (MCSCF)) optimization to address strong and weak (dynamic) electron correlation simultaneously and compare them with FCI results. We note that some of the earlier studies[42,44] have already considered double-zeta basis sets in exploring the symmetric dissociation problem in H_2O using $\Omega = 0$ wave functions. However, here we have a different focus in analyzing the potential energy profile for H_2O compared to what it was done in Refs. 42 and 44. Also, differently from Ref. 42, here we use full orbital optimization for the double-zeta basis set. Thus, we shall explore the effectiveness of wave functions restricted to configurations with low-seniority numbers up to $\Omega = 2$ for describing the profile of potential energy curves[62] for H_2O, N_2, C_2, and Be_2. To this end, we present potential energy curves for Ω-restricted wave functions as well as for CI expansions that are based on excitation-level restriction for comparison purposes. Furthermore, for the case of Be_2 dissociation (cc-pVTZ basis set), we apply a seniority-number restriction to molecular orbitals (MOs) *only in the virtual space*. This means that no Ω-constraint is used for the strongly occupied orbitals that correspond to the full valence active space representing the reference function (REF). Thus, a complete multireference CI wave function is used only for the Be_2 case where the reference function contains 4 electrons in 8 orbitals. We compare the performance of Ω-based schemes with the traditional CI method based on single and double substitutions out of the reference for the Be_2 case. In the latter case, only symmetry-adapted (optimized) MOs are used. Finally, in cases of N_2 and C_2 dissociations, we shall also consider novel hybrid wave functions obtained by *a union* of a set of CI configurations representing a full valence space (or a complete active space (CAS)), and a set of CI configurations where seniority-number restriction is imposed for a complete set of molecular orbitals (both strongly occupied and virtual MO sets), simultaneously. These hybrid wave functions are investigated in order to explore possible advantages of such schemes where both static and dynamic electron correlations play a significant role. The presence of CAS configurations in the hybrid CI expansions ensures the correct separability of these wave functions.

While CI expansions, even restricted to the lowest seniority, i.e., $\Omega = 0$ configurations are still associated with a relatively high computational cost,[44] they nevertheless offer a new pathway in developing novel and cost-effective *ab initio* methods. Such examples may include coupled cluster type models, e.g., pair-cluster-doubles (pCCD)[43,44] or an efficient ordering of configurations within the framework of energy renormalization group (RG) methods for bond dissociations.

II. COMPUTATIONAL DETAILS

Calculations for H_2O, N_2, C_2 (6-31G basis sets), and Be_2 (cc-pVTZ basis set) have been performed along ground-state (singlet-multiplicity) dissociation paths. These computations have been carried out by one of us (LB) using a pilot code for generating the CI determinants with selected seniority number restrictions. These CI expansions provided the input for the general CI code of Ivanic and Ruedenberg[20] (available in the GAMESS quantum chemistry package[63]) which was used to calculate the potential energy curves by one of us (LB). For all CI expansions above, the orbital-energy optimization has been performed using MCSCF-codes available in GAMESS. We choose larger basis set in the case of the beryllium dimer because it contains fewer electrons and the resulting CI wave functions are small enough to be handled by our pilot code. We use the symmetry-adapted molecular orbitals in the case of Be_2 in order to use symmetry to reduce the number of Slater determinants. The computations using the Ω-based restrictions on molecular orbitals described below can be performed using the generalized selected-CI code of Ivanic and Ruedenberg[20] available in GAMESS package.[63] The graphical display of molecular orbitals has been produced using the MacMolPlt software.[64]

III. BACKGROUND

A. Seniority number

For any given Slater determinant, the seniority number is defined as the number of singly occupied orbitals. In general, a CI wave function is represented by a linear combination of Slater determinants that can be grouped into different sectors, each labeled by a specific seniority number.[39] Although the seniority number concept has found much use[65] in the physics community, there have been relatively few applications[66] to quantum chemistry problems with only recent applications found in Refs. 39–44. It should be noted that perfect-pairing (seniority-zero) wave functions are conceptually similar to the ones based on geminals (two-electron basis functions).[67–72] For instance, the antisymmetrized geminal power (AGP) has been found useful in many recent variants of quantum chemical approaches, like, e.g., the AP1roG (antisymmetric product of 1-reference-orbital geminals) of Ayers and co-workers[45] or the Cluster-Jastrow-AGP ansatz of Neuscamman.[33]

B. Methods

First, we introduce the notation for CI wave functions that are used in this study. Based on the seniority number constraint, a set of configurations in a CI expansion can be specified as follows:

$$[n,M] = \{REF[n,m]/\Omega_{max}: m_{min}, M\}. \quad (1)$$

Here, n represents the number of correlated electrons, M represents the total number of molecular orbitals used for correlation (core orbitals are being kept doubly occupied at all times and are *not* included in M), and m represents the number of MOs used in the REF. The symbol REF[n,m] represents the active space of *the reference function* only. We shall use [n,M] to represent *the total* active space for CI expansion. The symbol Ω_{max} indicates the *highest* seniority number allowed in the set molecular orbitals, from m_{min} to M. The smallest seniority number is zero. Thus, Ω_{max} is the number of orbitals in the set $\{m_{min}, m_{min}+1, \ldots, M\}$ that are occupied by a single electron. Here, m_{min} is the first orbital from which the restriction based on Ω begins and proceeds in ascending fashion until it reaches the value M. In this study, we only consider two possible values of m_{min}. One case is with $m_{min} = 1$ (H_2O, N_2, and C_2 dissociations) and another case is $m_{min} = m + 1$ (Be_2 dissociation), where m is the size of the orbital space in the reference function. Namely, we shall consider the following Ω-restrictions:

$$[n,M] = \{REF[n,m]/\Omega_{max}: 1, M\}, \quad (2)$$

$$[n,M] = \{REF[n,m]/\Omega_{max}: m+1, M\}. \quad (3)$$

The excitation-based CI selection scheme has been used extensively in the past and can be represented in the following conventional format:

$$[n,M] = REF[n,m]/(EXC_{max}), \quad (4)$$

where EXC_{max} is the highest excitation level out of the reference function REF[n,m]. For example, if $EXC_{max} = 2$, then $\{REF[n,m]/(EXC_{max} = 2)\}$ represents the traditional MR-CISD, i.e., multireference CI singles and doubles. Note that if $\Omega_{max} = n$ or $EXC_{max} = n$, then [n,M] represents the FCI wave function.

Finally, for cases of N_2 and C_2 molecules, we shall also consider a *hybrid* CI wave function where we consider *a union* of a set of CI configurations representing the full valence space (or CAS) and a set of CI configurations where seniority-number Ω_{max} restriction is imposed for a set of molecular orbitals {1, M}, simultaneously. We shall denote such a wave function using the following notation:

$$HYBRID[n,M] : \{CAS + 1REF/\Omega_{max}:1, M\}. \quad (5)$$

Here, the symbol CAS represents the full valence (CAS) space and the symbol 1REF (single-reference) indicates that the seniority-number restriction for the second set of determinants applies to a set of all orbitals {1, M}, and not only to a set {m + 1, M}. This means that the wave function in Equation (5) does not represent a genuine (complete) multireference function, in contrast to the case specified by Equation (3). For example, it implies that in the wave function represented by Equation (5), besides determinants corresponding to the full valence space, no determinant has more than Ω_{max} singly occupied orbitals in *the complete correlated* MO space. Note that in making Ω_{max}-restricted determinants, the starting configuration *is a single determinant* (1*REF*) containing only *doubly occupied* MOs (for example, there are only 5 such orbitals in the case of N_2), thus, the symbol 1REF in Equation (5). Clearly, some of the configurations generated in such a way will match the ones in the CAS space (for example, 10 electrons in 8 orbitals in the case of N_2); however, most of the CI configurations will be distinct from the set in the CAS expansion.

We consider wave functions of the compact form given by Equation (5) instead of fully multi-reference CI expansions that would result in Equation (3) when REF[n,m] = CAS[n,m] and it corresponds to a multi-configurational expansion (e.g., CAS[10,8] in case of N_2) in order to maintain a small size of such CI wave functions. We intend to use the results of this study to develop novel computational methods, especially in combination with the RG procedure.[73,74]

IV. RESULTS AND DISCUSSION

A. H_2O symmetric dissociation

First, we shall focus on the potential energy curve for the symmetric dissociation of H_2O in the ground electronic state. The equilibrium geometry is R(OH) = 0.9572 Å, angle(HOH) = 104.52° deduced from experimental data (see, e.g., Ref. 75). This geometry has been frequently used for benchmark applications[75,76] in *ab initio* quantum chemistry. The basis sets used are 6-31G. The 1s core orbital on oxygen is kept doubly occupied at all times; thus, n = 8 (8 valence-electrons) and M = 12. The FCI wave function contains 61 441 Slater determinants (singlet-A_1 state) when symmetry-adapted MOs are used and 245 025 determinants (singlet-A state) in case of broken-symmetry (split-localized) molecular orbitals. Table I displays the relevant data concerning the length of CI expansions in terms of Slater determinants. Since the CI spaces generated using symmetry-adapted MOs on one hand, and CI spaces produced using symmetry-broken (e.g., localized) MOs on the other, *are not* invariant with respect to orbital rotations within the same orbital space it is preferable to use symmetry-broken (more flexible) MO set for the optimal results. Our earlier studies indicated advantages[39] of the symmetry-broken MOs for describing potential energy profile when using full valence spaces for H_8, N_2, and CO_2. However, since in this study we go beyond full valence spaces, we display the data for potential energy curves for H_2O using both, i.e., symmetry-adapted and symmetry-broken MOs (split-localized orbitals[77,78]) to make sure that the same trend is observed. Indeed, Figure 1 clearly demonstrates the *superiority of symmetry-broken* MOs (a graphical display of these orbitals will be provided later in the text) for representing the potential energy curve in H_2O using the [8, 12] = $\{REF[8,4]/\Omega_{max} = 0: 1, 12\}$ wave function while the length of both expansions consists of 495 Slater determinants. In fact, on the scale of Figure 1, the profile produced by the $\{REF[8,4]/\Omega_{max} = 0: 1, 12\}$ wave function using symmetry-broken (split-localized[78]) MOs appears almost parallel to the potential energy curve generated by FCI. In order to gain a more quantitative insight into the quality of this wave

TABLE I. Number of CI determinants in selected CI wave functions for the symmetric dissociation in H_2O and N_2 (6-31G basis sets). Core 1s orbital for oxygen is kept doubly occupied at all times. Core 1s orbitals for nitrogen are kept doubly occupied at all times.

Wave function	Number of determinants: Symmetry-adapted-MOs	Number of determinants: Symmetry-broken-MOs
H_2O		
CAS[8,6]	65	...
$[8,12]=\{REF[8,4]/\Omega_{max}=0:1,12\}$	495	495
$[8,12]=\{REF[8,4]/\Omega_{max}=2:1,12\}$	5 775	16 335
[8, 12] = Full-CI	61 441	245 025
$[8,12]=REF[8,4]/(EXC_{max}=2)$	409	...
$[8,12]=REF[8,4]/(EXC_{max}=4)$	13 751	...
$[8,12]=REF[8,6]/(EXC_{max}=2)$	6 223	...
N_2		
CAS[10,8]	396	...
$[10,16]=\{REF[10,5]/\Omega_{max}=0:1,16\}$	4 368	4 368
[10, 16] = Full-CI	2 388 528	...
HYBRID[10,16]: CAS + $1REF/\Omega_{max}=0:1,16$[a]	...	7 448
HYBRID[10, 16]: CAS + $1REF/\Omega_{max}=0:1,16$[a]	4 708	...
HYBRID[10, 16]: CAS + $1REF/\Omega_{max}=2:1,16$[a]	36 680	...
$[10,16]=REF[10,5]/(EXC_{max}=2)$	618	...
$[10,16]=REF[10,5]/(EXC_{max}=4)$	69 876	...
$[10,16]=REF[10,5]/(EXC_{max}=6)$	894 292	...

[a]Here, 1REF represents a single-reference wave function REF[10,5].

function, as well as the other CI expansions considered in this study, *the non-parallelity errors (NPEs) with respect to full CI* are listed in Table II along with the number of Slater determinants for convenience. The NPE for the $\{REF[8,4]/\Omega_{max}=0:1,12\}$ wave function is about 5 kcal/mol smaller than the one associated with FORS[8,6] or CAS[8,6] functions.

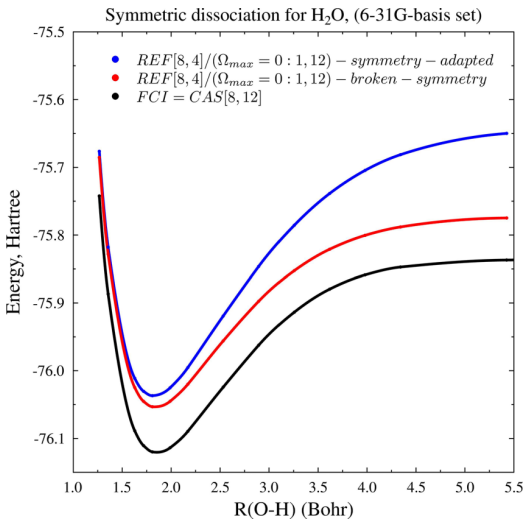

FIG. 1. Potential energy curves for symmetric dissociation of H_2O in the ground electronic state representing active space of 8 electrons in 12 orbitals (6-31G basis sets). Core 1s orbital for oxygen is kept doubly occupied at all times. All molecular orbitals are MCSCF-optimized.

TABLE II. Non-parallelity errors (NPE) [in kcal/mol] with respect to the FCI result for symmetric dissociation of H_2O, N_2, and C_2 (6-31G basis sets). Core 1s orbitals for oxygen and nitrogen are kept doubly occupied at all times. Symmetry-adapted molecular orbitals labeled by irreps of D_{2h} point-group are used unless explicitly noted otherwise. All the molecular orbitals are energy-optimized using MCSCF procedure.

Wave function	NPE	Number of determinants
H_2O		
CAS[8,6]	12.8	65
$[8,12]=\{REF[8,4]/\Omega_{max}=0:1,12\}$[a]	7.9	495
$[8,12]=\{REF[8,4]/\Omega_{max}=2:1,12\}$[a]	10.9	16 335
$[8,12]=REF[8,4]/(EXC_{max}=2)$	67.2	409
$[8,12]=REF[8,4]/(EXC_{max}=4)$	0.8	13 751
$[8,12]=REF[8,6]/(EXC_{max}=2)$	0.3	6 223
N_2		
CAS[10,8]	8.9	396
$[10,16]=\{REF[10,5]/\Omega_{max}=0:1,16\}$[a]	44.7	4 368
HYBRID[10, 16]: CAS + $1REF/\Omega_{max}=0:1,16$[a,b]	15.7	7 448
HYBRID[10, 16]: CAS + $1REF/\Omega_{max}=0:1,16$[b]	12.3	4 708
HYBRID[10, 16]: CAS + $1REF/\Omega_{max}=2:1,16$[b]	20.3	36 680
$[10,16]=REF[10,5]/(EXC_{max}=2)$	323.3	618
$[10,16]=REF[10,5]/(EXC_{max}=4)$	38.2	69 876
$[10,16]=REF[10,5]/(EXC_{max}=6)$	2.1	894 292
C_2		
CAS[8,8]	5.5	660
$[8,16]=\{REF[8,4]/\Omega_{max}=0:1,16\}$[a]	47.5	1 820
$[8,16]=\{REF[8,4]/\Omega_{max}=0:1,16\}$	32.4	1 820
$[8,16]=\{REF[8,4]/\Omega_{max}=2:1,16\}$	30.8	13 468
$[8,16]=\{REF[8,4]/\Omega_{max}=4:1,16\}$	5.3	106 924
HYBRID[8, 16]: CAS + $1REF/\Omega_{max}=0:1,16$[b]	4.8	2 410
$[10,16]=REF[10,5]/(EXC_{max}=2)$	54.9	469
$[10,16]=REF[10,5]/(EXC_{max}=4)$	7.4	35 663

[a]Broken-symmetry MOs have been used.
[b]Here, 1REF represents a single-reference wave function.

The main purpose of the data presented in Tables I and II is to compare Ω-restricted (symmetry-broken MOs) and excitation-level-restricted (symmetry-adapted MOs) CI expansions generated by electron substitutions out of MOs in a single-reference (a closed-shell Slater determinant) into orbitals in virtual space. From Table II, one can see that although $\{REF[8,4]/\Omega_{max}=0:1,12\}$ and $\{REF[8,4]/(EXC_{max}=2)\}$ contain approximately the same number of determinants, i.e., 495 and 409, respectively, the NPE for each of them differs greatly, clearly favoring the Ω-based approach. For clarity, these data are also displayed graphically in Figure 2 where $\{REF[8,4]/\Omega_{max}=0:1,12\}$ is represented in blue color while $\{REF[8,4]/(EXC_{max}=2)\}$ is represented in brown. For comparison purposes, in Figure 2, we also display potential energy curves representing the full-valence-space FORS[8,6] or CAS[8,6] (in purple) as well as the curve representing FCI (in black). In order to provide a better comparison, the graphical representations of energy deviations for each method (relative

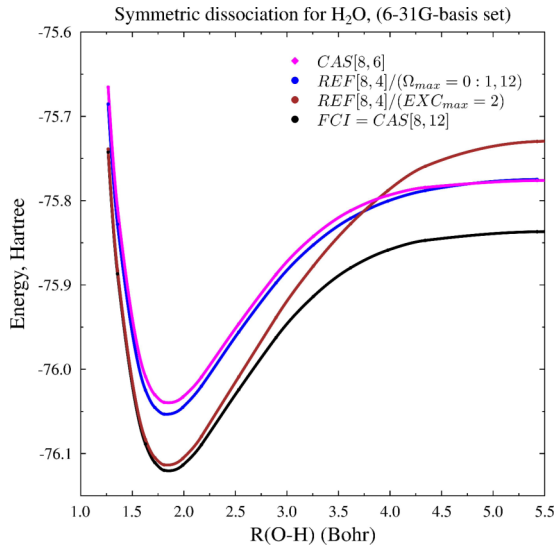

FIG. 2. Potential energy curves for symmetric dissociation of H_2O in the ground electronic state representing active space of 8 electrons in 12 orbitals (6-31G basis sets). Core 1s orbital for oxygen is kept doubly occupied at all times. All molecular orbitals are MCSCF-optimized. Note that only Ω-based selection scheme uses symmetry-broken molecular orbitals.

to FCI values) are plotted along the dissociation coordinate in Figure 3. However, as we explore the rate of convergence towards FCI by allowing higher values of Ω_{max} or EXC_{max} to be considered, it is clear that {REF[8,4]/Ω_{max} = 2: 1,12} exhibits significantly larger deviations from FCI (Figure 4) along the dissociation coordinate compared to {REF[8,4]/(EXC_{max} = 4)} (Figure 5) although the CI expansion sizes associated with these wave functions are about the same, i.e., 16 335 and 13 751, respectively. When we compare only Ω-based approaches, i.e., {REF[8,4]/Ω_{max} = 0: 1,12} and {REF[8,4]/Ω_{max} = 2: 1,12}, one can observe that although the absolute errors (relative to FCI) for the latter are reduced significantly

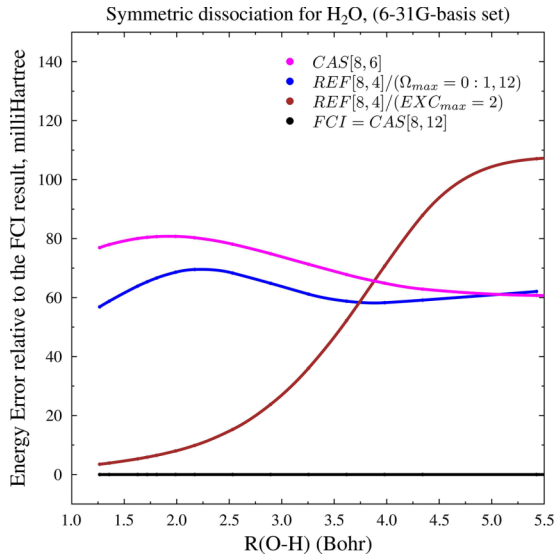

FIG. 3. Energy errors relative to the FCI result in symmetric dissociation of H_2O. The ground electronic state is described by an active space of 8 electrons in 12 orbitals (6-31G basis sets). Core 1s orbital for oxygen is kept doubly occupied at all times. All molecular orbitals are MCSCF-optimized. Note that only Ω-based selection scheme uses symmetry-broken molecular orbitals. All truncated CI methods use a single determinant as a reference function.

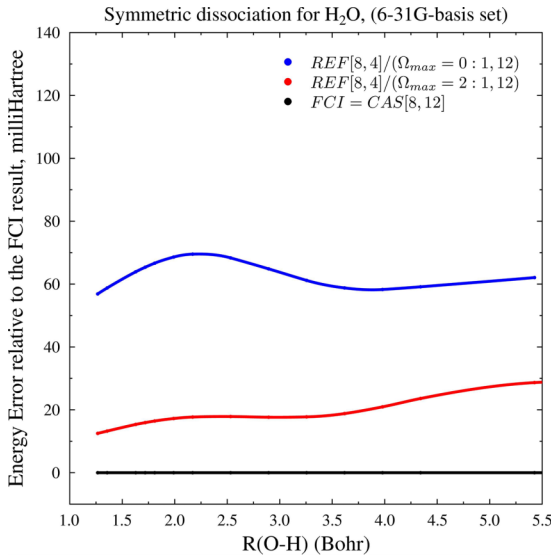

FIG. 4. Energy errors relative to the FCI result in symmetric dissociation of H_2O for the Ω-based CI selection scheme using broken-symmetry orbitals. The ground electronic state is described by an active space of 8 electrons in 12 orbitals (6-31G basis sets). Core 1s orbital for oxygen is kept doubly occupied at all times. All molecular orbitals are MCSCF-optimized. Note that truncated CI methods use a single determinant as a reference function.

(see Figure 4), the *relative errors* represented by NPEs are *actually larger* for {REF[8,4]/Ω_{max} = 2: 1,12} compared to {REF[8,4]/Ω_{max} = 0: 1,12}, as seen in Table II. On the other hand, when comparing only excitation-based approaches, i.e., {REF[8,4]/(EXC_{max} = 2)} and {REF[8,4]/(EXC_{max} = 4)}, one notices (Figure 5) that absolute and relative errors in terms of NPE are both significantly reduced going from EXC_{max} = 2 to EXC_{max} = 4.

It is of some interest to comment on the role of a reference function by going from a single-reference to a multireference

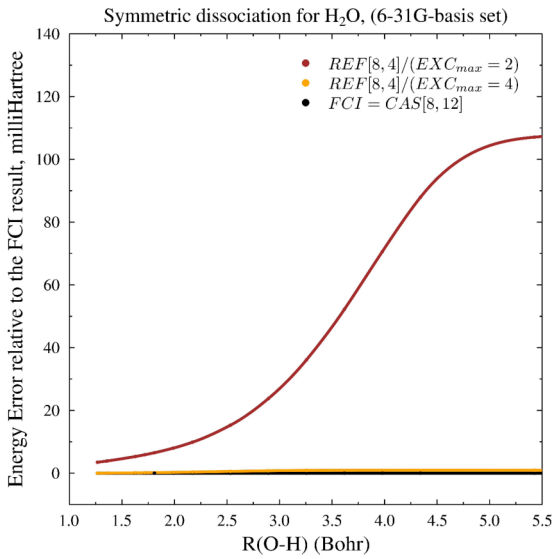

FIG. 5. Energy errors relative to the FCI result in symmetric dissociation of H_2O for the excitation-based CI selection scheme using symmetry-adapted orbitals. The ground electronic state is described by an active space of 8 electrons in 12 orbitals (6-31G basis sets). Core 1s orbital for oxygen is kept doubly occupied at all times. All molecular orbitals are MCSCF-optimized. All truncated CI methods use a single Hartree-Fock-like determinant as a reference function.

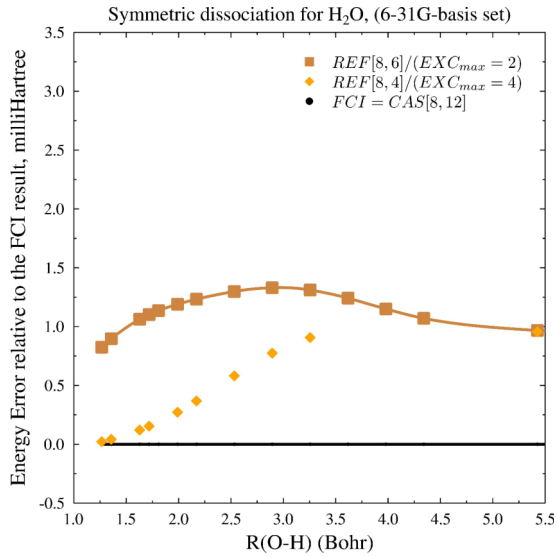

FIG. 6. Comparison of single-reference and multireference excitation-based methods for the symmetric dissociation of the H_2O molecule in the ground electronic state (6-31G basis set). The energy errors are relative to the FCI result. All molecular orbitals are MCSCF-optimized.

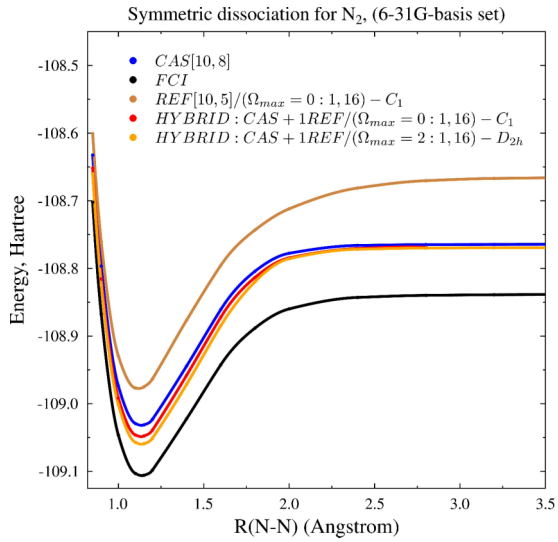

FIG. 7. Comparison of CI selection schemes for the ground-state potential energy curve of N_2 (6-31G basis set). All orbitals are MCSCF-optimized. Core 1s orbitals are kept doubly occupied at all times.

wave function. To this end, we have replaced the single-reference determinant by a FORS[8,6] (or CAS[8,6]) expansion keeping $EXC_{max} = 2$ restriction for the substitution of the reference MOs by virtual MOs in the CI wave function. This produces the $\{REF[8,6]/(EXC_{max} = 2)\}$ wave function. The performance of a single-reference-based method, i.e., $\{REF[8,4]/(EXC_{max} = 4)\}$ and multi-reference based method, $\{REF[8,6]/(EXC_{max} = 2)\}$ in describing the symmetric dissociation of water is displayed in Figure 6. One can see that although both methods yield small absolute deviations from FCI energy, it is clear that $\{REF[8,6]/(EXC_{max} = 2)\}$ produces smaller NPE values. In addition, the length of the CI expansion for $\{REF[8,6]/(EXC_{max} = 2)\}$ is about two times shorter than the one representing $\{REF[8,4]/(EXC_{max} = 4)\}$. Thus, it is evident that the performance of $\{REF[8,6]/(EXC_{max} = 2)\}$ is superior to the single-reference-based $\{REF[8,4]/(EXC_{max} = 4)\}$.

B. N_2 symmetric dissociation

Dissociation of first-raw diatomic molecules often provides a nice starting point for testing new methods in quantum chemistry.[7–14,49,50,79–91] The potential energy curve for the symmetric dissociation of N_2 in the ground electronic state is another example of a system which is frequently used to test the performance of new methods.[6,32,33,42,80–82,87] The case of N_2, however, is expected to be somewhat more challenging than the case of H_2O, since it involves breaking of a triple bond, while the symmetric dissociation for H_2O only deals with breaking two single bonds. The basis sets used are 6-31G. The 1s core orbitals are kept doubly occupied at all times; thus, n = 10 (10 valence-electrons) and M = 16. All molecular orbitals are fully optimized using a standard MCSCF procedure. The FCI wave function contains 2 388 528 Slater determinants (singlet-A_{1g} state) when symmetry-adapted (D_{2h}-point group) MOs are used. Table I displays the relevant data concerning the length of CI expansions in terms of Slater determinants for all wave functions used in this study. Table II and Figures 7–9 display the relevant data regarding the performance of various wave functions. As one can observe in Figure 7, the wave function $\{REF[10,5]/\Omega_{max} = 0: 1,16\}$ using broken-symmetry MOs (split-localized MOs) provides a *qualitatively* correct description of the PEC for N_2 (see the supplementary material[92] for graphical display of symmetry-adapted and symmetry-broken orbitals at various internuclear distances). However, the non-parallelity error compared to the FCI result is quite large, namely, 71.2 mHartree (44.7 kcal/mol) as seen in Table II. On the other hand, the CASSCF[10,8] result yields a lower-energy potential that dissociates into ROHF solutions for two nitrogen atoms (see Table III) with rather small non-parallelity errors of 14.2 mHartree (8.9 kcal/mol). The hybrid methods (CAS + 1REF/$\Omega_{max} = 0$: 1,16) using

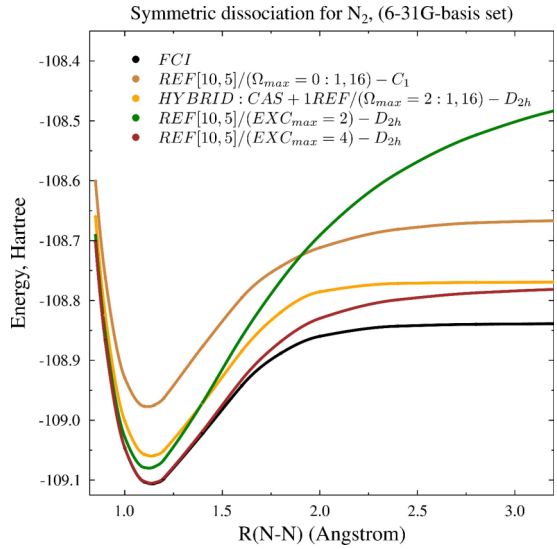

FIG. 8. Comparison of CI selection schemes for the ground-state potential energy curve of N_2 (6-31G basis set). All orbitals are MCSCF-optimized. Core 1s orbitals are kept doubly occupied at all times.

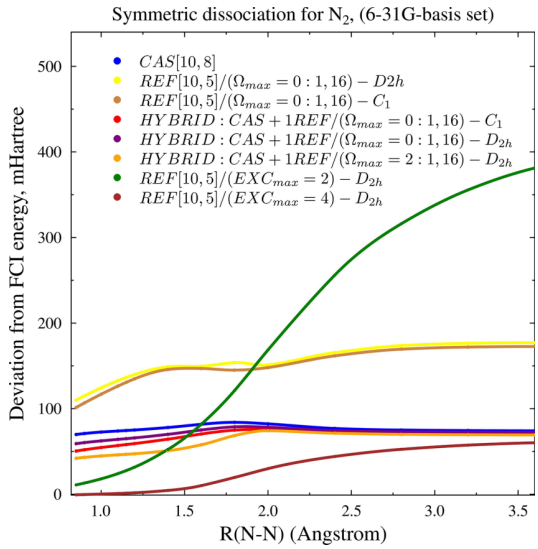

FIG. 9. Comparison of CI selection schemes for the ground-state potential energy curve of N_2 (6-31G basis set) in terms of deviations from the full CI result. All orbitals are MCSCF-optimized. Core 1s orbitals are kept doubly occupied at all times.

symmetry-broken MOs (C_1-point group) and (CAS + 1REF/Ω_{max} = 2: 1,16) using symmetry-adapted MOs (D_{2h}-point group) yield lower-energy potential energy curves compared to CASSCF solution; however, their NPEs are somewhat higher compared to the CASSCF result. When compared to their excitation-scheme counterparts, namely, {REF[10,5]/EXC_{max} = 2: 1,16} and {REF[10,5]/EXC_{max} = 4: 1,16}, the hybrid methods seem to perform quite well keeping in mind that they contain smaller number of determinants than the {REF[10,5]/EXC_{max} = 4: 1,16} wave function. Finally, Figure 9 exhibits plots of deviations from the FCI result along the dissociation coordinate in N_2 for all the methods considered in this study. It can be seen that the (CAS + 1REF/Ω_{max} = 2: 1,16) method using symmetry-adapted MOs (D_{2h}-point group) shows smallest deviations among the hybrid methods, while the excitation-based method {REF[10,5]/EXC_{max} = 2: 1,16} exhibits the largest errors for stretched internuclear distances. The method {REF[10,5]/EXC_{max} = 4: 1,16} improves dramatically over the latter, however at the expense of an increase in length of CI expansion (69 876 Slater determinants) while still exhibiting large NPE of 60.8 mHartree (38.2 kcal/mol). Thus, the hybrid methods considered in this study compare quite well with the excitation-based CI schemes in generating compact CI expansions that provide good quality description along the potential energy curve in N_2 addressing both static and dynamic electron correlations, simultaneously.

TABLE III. Size-consistency aspect of selected wave functions for H_2O and N_2 dissociation (6-31G basis set). Energy units are in Hartree.

Wave function	2 H	O	2H + O/H_2O
ROHF	2×(−0.498 233)	−74.778 234	−75.774 700
CAS[8,6]			−75.774 787 (R = 7.235 3 bohr)
[8, 12] = {REF[8, 4]/Ω_{max} = 0: 1, 12}			
C_1-MOs (symmetry broken orbitals)	...	...	−75.774 769 (R = 5.426 5 bohr)
D_{2h}-MOs (symmetry adapted orbitals)	...	...	−75.649 876 (R = 5.426 5 bohr)
Wave function	**2 N**		**N + N/N_2**
ROHF	2× (−54.382 051)		−108.764 102
CAS[10,8]	...		−108.764 313 (R = 3.6 Å)
[10, 16] = {REF[10, 5]/Ω_{max} = 0: 1, 16}			
C_1-MOs (symmetry broken orbitals)	...		−108.665 909 (R = 3.6 Å)
D_{2h}-MOs (symmetry adapted orbitals)	...		−108.661 418 (R = 3.6 Å)

[a]For molecular calculations, the internuclear distance R is listed in parentheses to indicate near-dissociation limit.

C. C_2 symmetric dissociation

Just like the N_2 system, the C_2 molecule[14,79,87–91] is a challenging case for calculating accurate potential energy curves for many *ab initio* methods. In contrast to the N_2 case however, the C_2 molecule requires multi-reference wave function for a qualitatively correct description even at near-equilibrium geometries.[87] The C_2 molecule is characterized by multiple bonds at its equilibrium geometry (one σ and two π bonds are involved) although the high occupancy of the $2\sigma^*$ orbital reduces the strength of the 2σ bond. Furthermore, a recent study by Shaik and co-workers[89] argues that the C_2 molecule can effectively be described as having a quadrupole bond at its equilibrium geometry. In addition, the C_2 system has a number of low-lying excited states[14,88,90,91] that affect the shape of the ground state potential energy curve due to avoided crossings.[90,91] Thus, the C_2 molecule represents an interesting case for our study. The basis sets used here are 6-31G. The 1s core orbitals are kept doubly occupied at all times; thus, n = 8 (8 valence-electrons) and M = 16. As before, all molecular orbitals are fully optimized using a standard MCSCF procedure. The FCI wave function contains 414 864 Slater determinants (singlet-A_{1g} state) when symmetry-adapted (D_{2h}-point group) MOs are used. Table II and Figures 10 and 11 display the relevant data regarding the performance of various wave functions. First of all, from Figure 10, one can observe that potential energy curves for FCI, CASSCF[8,8], {CAS + 1REF/Ω_{max}: 1,M}, and {REF[8,4]/(EXC_{max} = 4)} methods notably display a region of avoided crossing (around 1.7 Å) between states belonging to the same irreducible representation (see, e.g., Ref. 91). The Ω-based wave functions {REF[8,4]/Ω_{max} = 0: 1,16}, {REF[8,4]/Ω_{max} = 2: 1,16}, and {REF[8,4]/Ω_{max} = 4: 1,16} using symmetry-adapted MOs show a systematic convergence towards the FCI curve. The data show that Ω = 4 configurations are critical for achieving high-quality potential energy curves. Among the

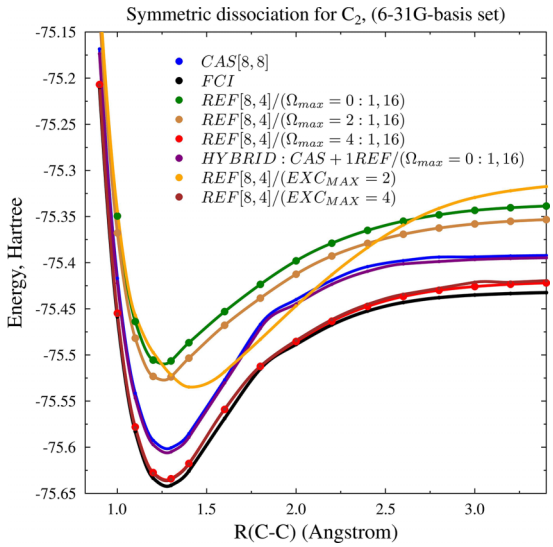

FIG. 10. Comparison of CI selection schemes for the ground-state potential energy curve of C_2 (6-31G basis set). All orbitals are MCSCF-optimized. Core 1s orbitals are kept doubly occupied at all times. All methods use symmetry-adapted molecular orbitals that transform as irreducible representations of the D_{2h} point group. The change in curvature in the region around 1.7 Å observed for some of the curves indicates an avoided-crossing between states belonging to the same irreducible representation.

excitation-based methods, one can see that {REF[8,4]/(EXC_{max} = 2)} generates a potential energy curve that has an elongated equilibrium bond-length compared to FCI and all the other methods that yield the equilibrium bond-length of about 1.3 Å. The closest agreement with the benchmark FCI curve among the wave functions considered for C_2 is observed for {REF[8,4]/Ω_{max} = 4: 1,16} and {REF[8,4]/(EXC_{max} = 4)}. It is noteworthy to mention that the hybrid method (CAS + 1REF/Ω_{max} = 0: 1,16) using symmetry-adapted MOs (D_{2h}-group) performs quite well yielding the smallest non-

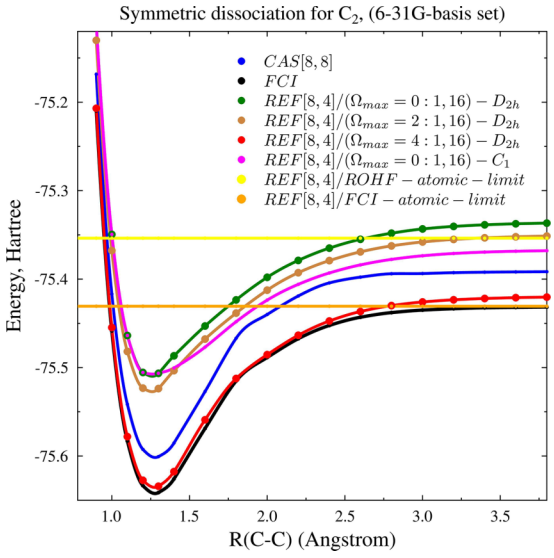

FIG. 11. Comparison of CI selection schemes for the ground-state potential energy curve of C_2 (6-31G basis set). All orbitals are MCSCF-optimized. Core 1s orbitals are kept doubly occupied at all times. Atomic dissociation limits are shown for ROHF and FCI levels of theory. The change in curvature in the region around 1.7 Å observed for some of the curves indicates an avoided-crossing between states belonging to the same irreducible representation.

parallelity errors of 4.8 kcal/mol among truncated CI wave functions. The use of symmetry-broken MOs in the case of {REF[8,4]/Ω_{max} = 0: 1,16} method (see magenta curve in Figure 11) yields a lower potential energy curve at stretched bond-lengths; however, at shorter distances (up to 1.4 Å), the lower energy solution is obtained by using symmetry adapted-molecular orbitals (see the supplementary material[92] for a graphical display of these orbitals). Thus, the {REF[8,4]/Ω_{max} = 0: 1,16} potential energy curve data points (magenta color) displayed in Figure 11 correspond to the lowest energy solutions at a given geometry where the optimal orbitals change from a symmetry-adapted set to a broken-symmetry set at R = 1.5 Å going towards the dissociation limit. It is interesting to note that CASSCF[8,8] method produces the dissociation limit of −75.3916 hartree which is significantly lower compared to atomic ROHF limit of two carbon atoms of −75.3537 hartree. One can argue that this is due to the fact that CASSCF[8,8] contains *more orbitals* than the minimum number needed to break the CC bond (in order to address the static correlation) and thus incorporates a significant amount of dynamic electron correlation at the dissociation-limit. This situation for C_2 differs from cases of H_2O and N_2 where CASSCF full valence spaces contain only the minimum number of MOs needed to break their chemical bonds as will be discussed in Sec. IV D.

D. Size-consistency and molecular dissociation in N_2 and H_2O

It is of some interest to address the question of *size-consistency* (or correct-separability) and *size-extensivity* (correct scaling of electron correlation energy)[93–95] of the methods that use seniority-number constraints to limit the size of CI expansions. For any method to be size consistent, the underlying reference must break all symmetries. The relevant data for H_2O and N_2 are presented in Table III. From these data, it can be seen that CASSCF solutions will converge to their atomic ROHF energies at their corresponding dissociation limits. As mentioned above, this can be argued to be due to a fact that these CASSCF spaces contain only the minimal number of MOs that are needed to break bonds. On the other hand, as can be observed from Figure 1, due to a fact that CI spaces using symmetry-adapted and symmetry-broken (split-localized) orbital sets are *not invariant*, the MCSCF procedure based on *symmetry-broken* MOs in the {REF[8,4]/Ω_{max} = 0: 1,12} wave function yields lower-energy solution (specially at the dissociation limit) compared to the symmetry-adapted MO set. See Figures 12–15 for a comparison between symmetry-adapted and broken-symmetry MOs produced by {REF[8,4]/Ω_{max} = 0: 1,12} wave functions at various geometries in H_2O. Since the energy of the {REF[8,4]/Ω_{max} = 0: 1,12} wave function at R(OH) = 3 × $R(OH)_{eq}$ (three times the equilibrium bond-length) is −75.7748 hartree which almost exactly matches the CASSCF dissociation limit (which is equal to the total sum of atomic ROHF energies), it *may appear* that the {REF[8,4]/Ω_{max} = 0: 1,12} method with symmetry-broken orbitals may produce a size-consistent result for this case. However, one must keep in mind that the {REF[8,4]/Ω_{max} = 0: 1,12} method uses 12 molecular orbitals and,

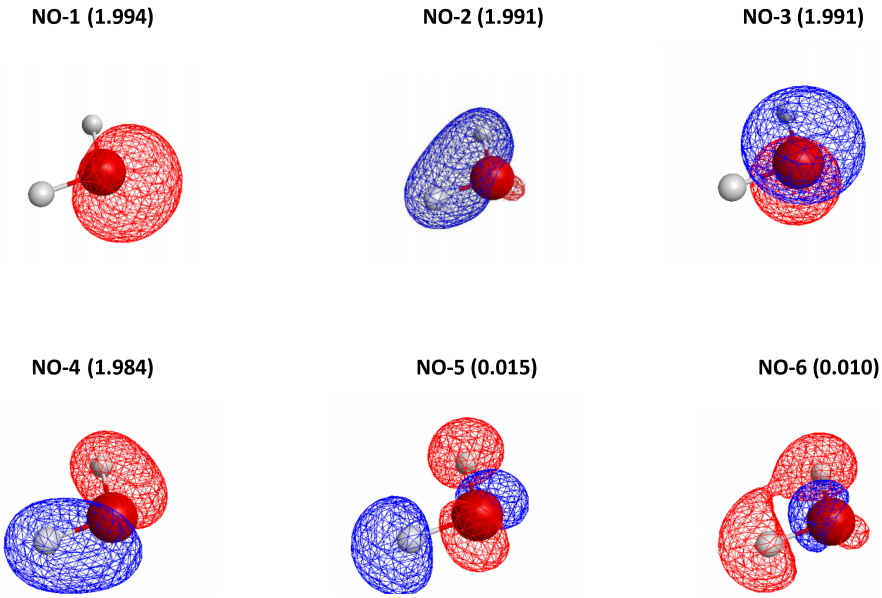

FIG. 12. The first 6 natural orbitals (occupation numbers shown in parentheses) based on the optimized MCSCF orbitals obtained using $\{\text{REF}[8,4]/\Omega_{\max} = 0: 1, 12\}$ method for the H_2O molecule (6-31G basis set) using C_{2v}-point group symmetry constraint. Core 1s orbital for oxygen is kept doubly occupied at all times and it is not shown. All the orbitals are symmetry-adapted and transform as irreducible representations of the C_{2v} point group and correspond to equilibrium geometry of $R_e(OH) = 1.808\ 85$ bohr.

therefore, includes *dynamical* electron correlation while the CASSCF[8,6] method using only the full valence MO space of 6 orbitals *does not*. In order to answer the question of *size-consistency conclusively*, one needs to identify only those configurations (Slater determinants) in the $\{\text{REF}[8,4]/\Omega_{\max} = 0: 1, 12\}$ wave function that account *only for static* electron correlation (without any contribution from dynamic correlation). Such configurations must be expressed in terms of only the *first six natural orbitals* (full valence space only). Then, the doubly occupied ($\Omega = 0$) configurations in $\{\text{REF}[8,4]/\Omega_{\max} = 0: 1, 6\}$ will yield energy that is only due to static correlation. The relevant data are presented in Table IV. Two wave functions can be considered: one contains only 4 determinants (wave function-I in this table) which are dominant, another contains the complete number, i.e., 15 of Slater determinants (wave function-II = $\{\text{REF}[8,4]/\Omega_{\max} = 0: 1, 6\}$ in this table). If we use the six dominant natural orbitals obtained from the MCSCF-optimized $\{\text{REF}[8,4]/\Omega_{\max} = 0: 1, 12\}$ wave function (at $3 \times R(OH)_{eq}$) and simply calculate the energy of wave function-I *without any further optimization* (see Table IV), the energy value is −75.7565 hartree (the value remains the same even if we try to optimize the orbitals). The inclusion of extra doubly occupied determinants in wave function-II leaves the energy practically unchanged. Since the energy of wave function-II of −75.7566 hartree is higher compared to the sum of atomic ROHF energies for this system of −75.7747 hartree by about 20 mHartree, it means that the $\{\text{REF}[8,4]/\Omega_{\max} = 0: 1, 12\}$ wave function *is not rigorously size-consistent*. The reason that the $\{\text{REF}[8,4]/\Omega_{\max} = 0: 1, 12\}$ expansion seems to approach the same dissociation limit as the CASSCF[8,6] method is due to a fact that the $\{\text{REF}[8,4]/\Omega_{\max} = 0: 1, 12\}$ wave function incorporates some dynamic correlation which, in a fortuitous way, appears to compensate for the lack of static correlation. The addition of $\Omega = 4$ configurations (6 determinants) to the CI expansion of wave function-I produces a 10-determinant wave function (wave function-III in Table IV) which matches the CASSCF[8,6] energy at the dissociation. Thus, the inclusion of these $\Omega = 4$ configurations in a wave function is critical for size-extensivity and proper description of simultaneous breaking of two single bonds in H_2O.

In the case of N_2, it is clear that the CASSCF[10,8] wave function is able to dissociate N_2 into N-atom fragments of

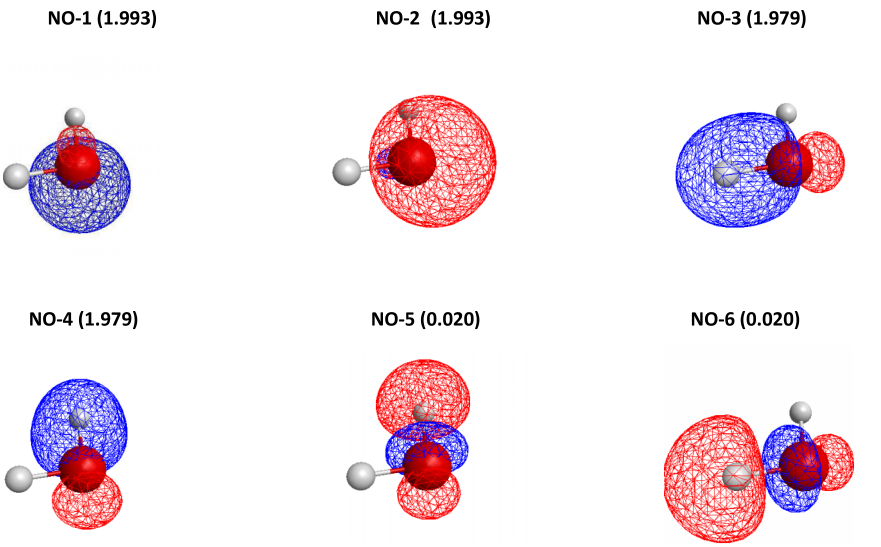

FIG. 13. The first 6 natural orbitals (occupation numbers shown in parentheses) based on the optimized MCSCF orbitals obtained using $\{\text{REF}[8,4]/\Omega_{\max} = 0: 1, 12\}$ method for the H_2O molecule (6-31G basis set) using no symmetry constraints (C_1-point group). Core 1s orbital for oxygen is kept doubly occupied at all times and it is not shown. All the orbitals are broken-symmetry (do not transform as irreducible representations of the C_{2v} point group of the molecular structure) and have a split-localized MO character. The results correspond to equilibrium geometry of $R_e(OH) = 1.808\ 85$ bohr.

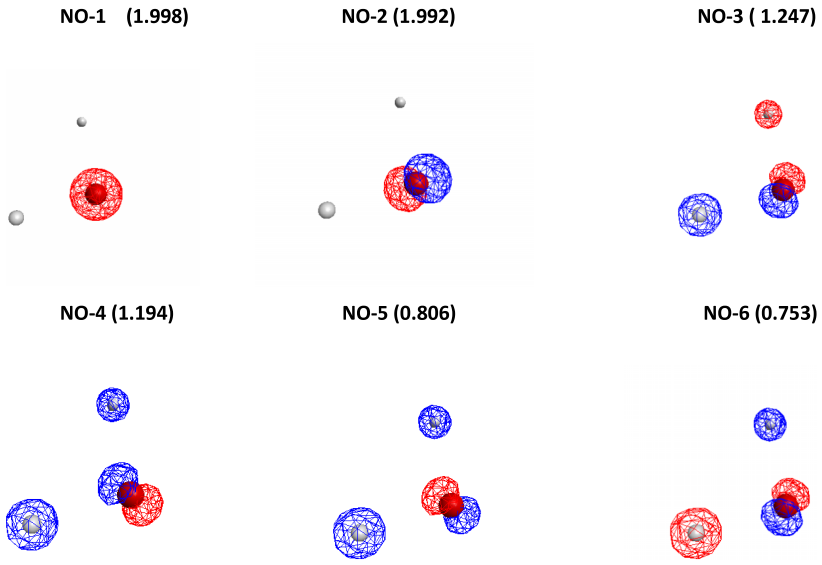

FIG. 14. The first 6 natural orbitals (occupation numbers shown in parentheses) based on the optimized MCSCF orbitals obtained using $\{\text{REF}[8,4]/\Omega_{\max}=0:1,12\}$ method for the H_2O molecule (6-31G basis set) using C_{2v}-point group symmetry constraint. Core 1s orbital for oxygen is kept doubly occupied at all times and it is not shown. All the orbitals are symmetry-adapted and transform as irreducible representations of the C_{2v} point group and are calculated at near-dissociation geometry of $3R_e$ (OH).

ROHF-type without any problems. On the other hand, the $\{\text{REF}[10,5]/\Omega_{\max}=0:1,16\}$ expansion yields a much higher energy at the dissociation-limit even with the use of symmetry-broken MOs (see the supplementary material[92] for the dominant natural orbitals after the optimization). The behavior of the $\{\text{REF}[10,5]/\Omega_{\max}=0:1,16\}$ potential curve is qualitatively similar to the potential energy curve obtained by using the GVB-PP wave function as reported by Faglioni and Goddard in Ref. 81.

To conclude, one can say that $\{\text{REF}[n,m]/\Omega_{\max}=0:m_{\min},M)\}$ wave functions do not yield size-consistent results for potential energy curves in general. Thus, higher $\Omega_{\max}$ values are often needed to ensure size-consistency for the Ω-based truncated CI expansions in terms of the sum of atomic-fragment ROHF (for open-shell systems) energy contributions. However, the use of Ω is helpful in identifying a small subset of configurations with $\Omega > 0$ that can produce a size-consistent result as shown above. Special cases where $\Omega = 0$ wave functions do produce size-consistent results for potential energy curves for full valence MO spaces are H_2 dissociation,[61] as well as H_8 linear chain symmetric dissociation.[39]

Of course, in order to obtain high-quality potential energy curves, one must go beyond static correlation (and beyond ROHF-atomic limit) and consider wave functions that recover dynamic correlation as well. In such cases, the consistent treatment using "super-atom" approach where atomic-limit energies are replaced by highly stretched molecular geometries may yield sufficiently small non-parallelity errors (see Refs. 7, 8, 10, and 12). As such, this approach can be useful for representing potential energy curves since the main focus is on energy differences rather than total energies.

Following the definition of Bartlett[93] referring to *size-extensivity* as the correct (linear) scaling of a method with the number of electrons, one can argue that $\Omega = 0$ wave functions are size-extensive with broken-symmetry molecular orbitals.

E. Be_2 symmetric dissociation

Now we shall focus on the Be_2 dissociation. This case has been the subject of numerous *ab initio* investigations (e.g., see Refs. 49 and 83–85) due to the fact that it is a weakly bound system (experimental binding energy at 2.4536 Å is only $-929.7\ \text{cm}^{-1}$) where both static and dynamic electron correlations are significant even for a qualitatively correct description. Recently, an in-depth analysis of the Be_2 dissociation curve has been reported by Schmidt, Ivanic, and Ruedenberg,[84] who

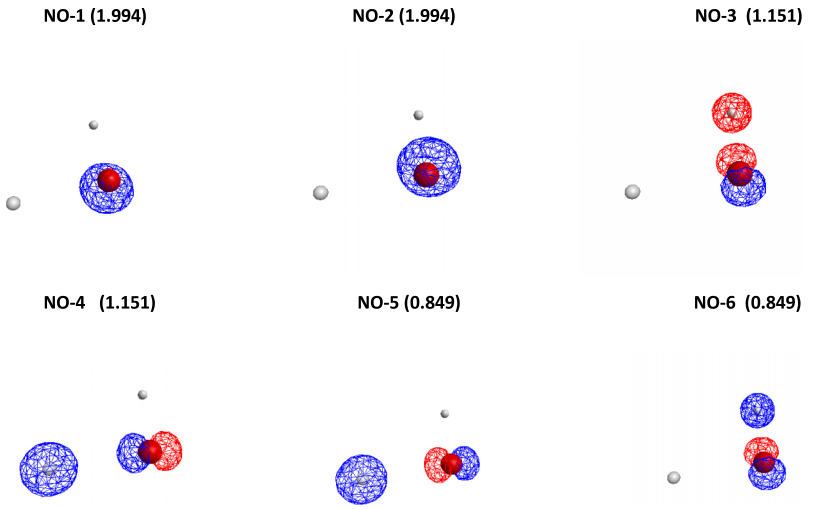

FIG. 15. The first 6 natural orbitals (occupation numbers shown in parentheses) based on the optimized MCSCF orbitals obtained using $\{\text{REF}[8,4]/\Omega_{\max}=0:1,12\}$ method for the H_2O molecule (6-31G basis set) using no symmetry constraints (C_1-point group). Core 1s orbital for oxygen is kept doubly occupied at all times and it is not shown. All the orbitals are broken-symmetry (do not transform as irreducible representations of the C_{2v} point group of the geometric structure) and have a split-localized MO character. The results correspond to near-dissociation geometry of $3R_e$(OH).

TABLE IV. Decomposition of FCI wave function to address the static correlation in terms of Slater determinants at near-dissociation limit ($3R(OH)_{eq}$) using strongly occupied natural orbitals from $\{REF[8,4]/\Omega_{max} = 0: 1, 12\}$ MCSCF calculation (energy = −75.774 77 hartree) for the H_2O molecule (6-31G basis set). The total sum of atomic ROHF energies is −75.774 70 hartree.

Strongly occupied natural orbital labels[a]						
1	2	3	4	5	6	CI energy
A. Wave function-I (orbital occupations listed below: 4-determinants, Ω = 0)						−75.756 52
2	2	2	2	0	0	
2	2	2	0	2	0	
2	2	0	2	0	2	
2	2	0	0	2	2	
B. Wave function-II (orbital occupations listed below: 15-determinants, Ω = 0)						−75.756 60
2	2	2	2	0	0	
2	2	2	0	2	0	
2	2	2	0	0	2	
2	2	0	2	2	0	
2	2	0	2	0	2	
2	2	0	0	2	2	
0	2	2	2	2	0	
0	2	2	2	0	2	
0	2	2	0	2	2	
0	2	0	2	2	2	
0	0	2	2	2	2	
2	0	2	2	2	0	
2	0	2	2	0	2	
2	0	2	0	2	2	
2	0	0	2	2	2	
C. Wave function-III (orbital occ. listed below: 10-determinants, Ω = 0, 4)						−75.775 88
2	2	2	2	0	0	
2	2	2	0	2	0	
2	2	0	2	0	2	
2	2	0	0	2	2	
2	2	1	1	1	1[b]	

[a]See Figure 15 for graphical display of these orbitals (full valence space).
[b]There are 6 determinants for this configuration involving all possible spin couplings between two α and two β electrons.

emphasized the significance of dynamic electron correlation in yielding the correct shape of the potential energy curve. These authors[84] performed a thorough analysis of the wave function along the dissociation coordinate. Thus, it is fair to say that a very high quality potential energy curve for Be_2 can be obtained by *ab initio* methods as can be confirmed by the agreement[85] between calculated and measured vibrational spectra for this system.

In the present study, our goal *is not* to generate a potential energy curve for Be_2 of spectroscopic quality, but rather to test the novel Ω-based CI selection scheme and compare it with a conventional, i.e., excitation-level based strategy. Furthermore, we shall explore the usefulness of the Ω-based CI selection scheme by applying it only to *the virtual* MO space (weakly occupied MOs), without any restrictions on the strongly occupied MOs in the full-valence-space. This means that *no restriction* (using Ω_{max} or EXC_{max}) will be imposed for the strongly correlating MOs within the full valence space of 4 electrons in 8 orbitals. The reference function is represented by a multi-reference wave function REF[4,8], i.e., 4 electrons in 8 orbitals, and the 1s core orbitals are being kept doubly occupied at all times. The dynamic electron correlation is addressed by allowing electronic substitutions into virtual MO space from this reference function resulting in a [4,58] wave function. The seniority-based approach is represented by $\{REF[4,8]/\Omega_{max} = 0: 9,58\}$ and $\{REF[4,8]/\Omega_{max} = 2: 9,58\}$, while the excitation-based method is represented by $\{REF[4,8]/(EXC_{max} = 2)\}$. For clarity, Table V compares the orbital occupations in the virtual MO space for Ω-based and excitation-based methods. While the use of the symmetry-broken MOs is expected to be more affective in producing lower-energy solutions as we have seen in cases of H_2O and N_2, in Be_2 we shall use only *symmetry-adapted* MOs simply to reduce the length of the CI expansion which can be easily handled by our pilot code. As before, full orbital (MCSCF) optimizations are performed for all CI expansions. The relevant data are displayed in Table VI and Figure 16. It is clear that the CASSCF[4,8] wave function (green curve in Figure 16) is not able to yield a bound system for Be_2 which is consistent with earlier studies[84] on this system (see Table SI of the supplementary material[92] for the list of CASSCF[4,8] dominant configurations at 4.0 Å geometry). At short internuclear distances, $\{REF[4,8]/\Omega_{max} = 0: 9,58\}$ yields a curve that is a bit lower but its quality is still rather poor and, in addition, the orbital-optimization step suffers from convergence problems beyond 2.5 Å. However, $\{REF[4,8]/\Omega_{max} = 2: 9,58\}$ is capable of yielding near-FCI energy values (~0.1 mhartree errors) at five selected geometries (see Figure 16) along the dissociation curve. The same high-quality performance is also observed for the excitation-based, i.e., MR-CISD type-method, $\{REF[4,8]/(EXC_{max} = 2)\}$. The latter wave function contains slightly fewer Slater determinants than the

TABLE V. The ground-state dissociation curve of Be_2 (cc-pVTZ basis set). Comparison of the Ω-based and excitation-based configuration interaction-selection schemes in terms of orbital occupations in the virtual MO space containing 50 molecular orbitals. The CI wave functions span the [4,58] active space.

Number of electrons (virtual MO space-only)[a]	MO-occupation Type[b]	Ω-based $\{REF[4,8]/\Omega_{max} = 2: 9,58\}$[c]	Excitation-based $REF[4,8]/(EXC_{max} = 2)$[d]
Are configurations with this MO-occupation-type allowed? (Yes/No)			
4	2, 2	Yes (Ω = 0)	No
	2, 1, 1	Yes (Ω = 2)	No
	1, 1, 1, 1	No (Ω = 4)	No
3	2, 1	Yes (Ω = 1)	No
	1, 1, 1	No (Ω = 3)	No
2	2	Yes (Ω = 0)	Yes (EXC = 2)
	1, 1	Yes (Ω = 2)	Yes (EXC = 2)
1	1	Yes (Ω = 1)	Yes (EXC = 1)

[a]The total number of electrons in the virtual MO space. The case of 0 electrons in the virtual space corresponds to the reference wave function.
[b]The number "2" indicates that the virtual MO is doubly occupied; the number "1" indicates that the virtual MO is singly occupied.
[c][4, 58] = $\{REF[4,8]/\Omega_{max} = 2: 9,58\}$ wave function contains 55 125 determinants.
[d][4, 58] = $\{REF[4,8]/(EXC_{max} = 2)$ wave function contains 32 652 determinants.

TABLE VI. The ground-state dissociation curve of Be_2 (cc-pVTZ basis set). All CI wave functions span the [4,58] active space. 1s-core electrons are kept doubly occupied at all times. All molecular orbitals are energy-optimized and are symmetry-adapted (D_{2h}-point symmetry group with ^1A_g wave function for the ground state).

R (Å)	FCI[a]	{REF[4, 8]/Ω_{max} = 0: 9, 58}[b]	{REF[4, 8]/Ω_{max} = 2: 9, 58}[c]	REF[4, 8]/(EXC_{max} = 2)[d]	CASSCF[e]
2.00	−29.227 713	−29.210 485			−29.205 154
2.20	−29.236 986	−29.221 236	−29.236 905		−29.216 320
2.30	−29.238 869	−29.223 971			29.219 408
2.40	−29.239 721	−29.225 737	−29.239 608		29.221 563
2.60	−29.239 825	−29.227 775	−29.239 785[f]	−29.239 761	−29.224 376
2.70	−29.239 544				−29.225 413
2.80	−29.239 216				−29.226 331
3.00	−29.238 636		−29.238 588[f]	−29.238 585	−29.227 939
3.60	−29.237 889				29.231 186
4.00	−29.237 708		−29.237 697[f]	29.237 696	−29.232 196
5.00	−29.237 379				−29.232 872
6.00	−29.237 205				−29.232 904
10.0	−29.237 093				−29.232 878
30.0	−29.237 088				−29.232 877

[a]FCI wave function contains 346 485 determinants.
[b][4, 58] = {REF[4, 8]/Ω_{max} = 0: 9, 58} wave function contains 1937 determinants.
[c][4, 58] = {REF[4, 8]/Ω_{max} = 2: 9, 58} wave function contains 55 125 determinants.
[d][4, 58] = {REF[4, 8]/(EXC_{max} = 2) wave function contains 32 652 determinants.
[e]Full-valence-space: FORS[4, 8] = CAS[4, 8] wave function contains 112 determinants.
[f]Energies are not completely converged (the MCSCF optimization was stopped due to an excessive number of iteration).

{REF[4, 8]/Ω_{max} = 2 : 9, 58} function, i.e., 15.9% versus 9.4% of the FCI expansion (spatial symmetry corresponds to a A_{1g}-irrep of D_{2h}-point group for the singlet state). Due to the fact that the {REF[4, 8]/Ω_{max} = 2: 9, 58} expansion contains additional configurations compared to the {REF[4, 8]/(EXC_{max} = 2)} function (see Table V), it yields slightly lower MCSCF-converged energies (see Table VI). Furthermore, when RHF molecular orbitals are used as the starting MOs, at the first MCSCF iteration, the {REF[4, 8]/Ω_{max} = 2: 9, 58} method yields a lower energy than {REF[4, 8]/(EXC_{max} = 2)} by about 0.6 and 0.5 mhartree for geometries R = 3.0 Å and R = 4.0 Å, respectively (see Table SII of the supplementary material[92] for more information). However, the difference in energies between these two methods quickly reduces as the number of MCSCF-iterations increases (see Table SII of the supplementary material[92]).

Thus, one can see that the Ω-based CI selection scheme exhibits a similar performance as the conventional, excitation-based method for describing the dynamic electron correlation of the beryllium dimer.

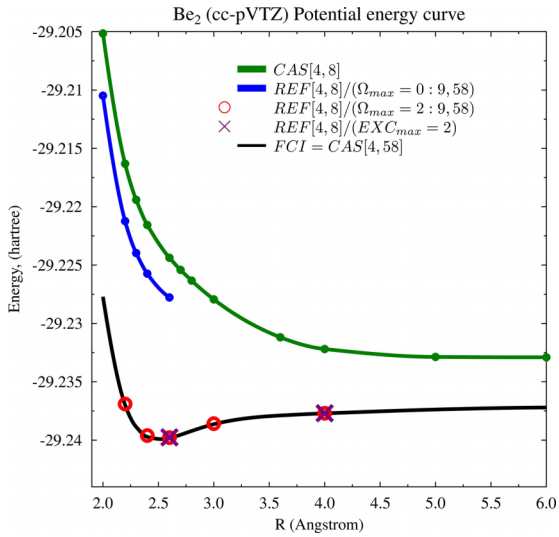

FIG. 16. Comparison of CI selection schemes for the ground-state potential energy curve of Be_2 (cc-pVTZ basis set). All orbitals are MCSCF-optimized. Core 1s orbitals are kept doubly occupied at all times. All truncated CI methods are multi-reference, i.e., the CAS[4,8] active space represents a reference function.

V. SUMMARY AND CONCLUSIONS

In this pilot study, we broadened the scope of studies focusing on the role of Ω in generating compact CI expansions by exploring its role in addressing dynamic correlation problem, especially in the presence of static (strong) electron correlation in situations when chemical bonds are broken or formed. Our earlier studies[39] used an Ω-concept focused primarily on the reliable description of *strong electron* correlation along reaction paths. By considering wave functions that go *beyond* full valence active spaces or minimal basis sets, in this work, we addressed *dynamic electron correlation* as well in addition to the static electron correlation. In the case of the symmetric dissociation of H_2O (6-31G basis), we find that {REF[8, 4]/Ω_{max} = 0: 1, 12} containing only 495 Slater determinants is capable of providing a robust description of the potential shape with NPE of only ∼8 kcal/mol while the CASSCF[8,6] wave function (65-determinants) yields NPE of ∼13 kcal/mol, as compared to the FCI result. This observation suggests that seniority number is also useful for addressing both strong and weak (dynamic) electron correlations. It is noteworthy that in the H_2O case it is essential to use symmetry-broken (split-localized, not atomic-like as shown in Figures 13

and 15) MOs in order to take the full advantage of Ω-based approach. We find that while a single-reference Ω-based approach, i.e., $\{\mathrm{REF}[8,4]/\Omega_{\max} = 0\text{:}\,1,12\}$ performs well for capturing the profile of the potential energy curve for the symmetric dissociation in H_2O, the increase in level of theory by increasing $\Omega_{\max}$ from 0 to 2 reduces absolute errors but does not lead to a reduction of NPEs, which is somewhat counterintuitive. In contrast to the Ω-based approach, excitation-based systematic CI truncations, also based on a single-reference of Hartree-Fock type, offer a much faster convergence towards FCI results although the $\{\mathrm{REF}[8,4]/(\mathrm{EXC}_{\max} = 2)\}$ wave function is associated with *very large* errors.

The use of seniority-number for the dissociation problem in N_2 (6-31G basis set) has demonstrated that the overall shape of the seniority-zero restricted CI expansion using symmetry-broken molecular orbitals is dominated by static electron correlation as it exhibits similarities to the curves in our earlier study[39] when only static correlation was considered. However, the hybrid methods combining the CI configurations representing CI determinants in the full-valence-space and a set of complementary configurations restricted to some maximum value of seniority-number (for the complete MO set {1,M}) offer a significant improvement in an overall shape of the potential curve. While the total energy deviations from the FCI curve for the hybrid ($\mathrm{CAS} + 1\mathrm{REF}/\Omega_{\max}$: 1,16) methods are comparable to those of the excitation-based approach $\{\mathrm{REF}[10,5]/\mathrm{EXC}_{\max} = 4\text{:}\,1,16\}$ which yields smaller deviations (especially at the near-equilibrium geometries), the NPE is considerably smaller for the former which is encouraging. While the NPEs for the hybrid methods are larger than the ones exhibited by the CASSCF[10,8] result for 6-31G basis set, it is expected that the performance of the seniority-restricted hybrid methods will be improved for larger basis sets when the contribution of dynamic correlation gets larger. In fact, for the case of the C_2 molecule, the hybrid approach using the $(\mathrm{CAS} + 1\mathrm{REF}/\Omega_{\max} = 0\text{:}\,1,16)$ wave function with symmetry-adapted MOs (D_{2h}-group) yields the smallest non-parallelity error among the truncated CI expansions considered in this study. Thus, using the hybrid approach seems to be very promising in generating accurate potential energy curves at a reduced-cost.

A further support for the usefulness of Ω-based approach is observed in the case of Be_2 dissociation (cc-pVTZ) where only symmetry-adapted MOs have been used. Both multi-reference methods, i.e., $\{\mathrm{REF}[4,8]/\Omega_{\max} = 2\text{:}\,9,58\}$ and $\{\mathrm{REF}[4,8] + (\mathrm{EXC}_{\max} = 2)\}$, reproduce near-FCI-quality potential with errors of ~0.1 mhartree. Overall, one can conclude that both multi-reference approaches, Ω-based and excitation-based, seem to address the dynamic electron correlation very well. Also, for both test cases, H_2O and Be_2, we find that multireference methods perform better in reducing non-parallelity errors along reaction paths.

It is important to note that in the present investigation, molecular orbitals have been fully MCSCF-optimized for all CI wave functions (not just in cases representing full-valence-spaces which is routinely done[60] for MCSCF wave functions). While for FCI wave functions, the orbital-optimization step is unnecessary; for truncated CI expansions, however, it is critical. A number of orbital-optimized electron correlation approaches have recently appeared in a literature representing only seniority-zero sectors of wave functions[43–47] that have been successfully applied to bond breaking processes. An interesting and challenging future case to further test the Ω-based CI selection methods would be the description of the dissociation process in the chromium dimer[27,96] where the sextuple bond is being broken. This case as well as other challenging systems (see, e.g., Ref. 97) will be a subject of our future investigations.

Overall, the present study provides an insight into the potential usefulness of the seniority number in describing *strong and weak (dynamical) electron correlations* simultaneously along reaction paths. The evidence presented in this study shows that seniority-number-based approaches can be an effective tool in describing potential energy curves.

ACKNOWLEDGMENTS

L.B. thanks Dr. Michael W. Schmidt for his continued assistance with the GAMESS code and feedback regarding the MacMolPlt software for plotting molecular orbitals. This work was supported by the National Science Foundation (No. CHE-1462434). G.E.S. is a Welch Foundation chair (No. C-0036). The work was also in part supported (for K.R.) by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences & Biosciences through the Ames Laboratory at Iowa State University under Contract DE-AC02-07CH11358. L. Bytautas gratefully acknowledges the computational resources of the theoretical chemistry group in the Department of Chemistry at Iowa State University.

[1]C. D. Sherrill, J. Chem. Phys. **132**, 110902 (2010).
[2]P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, Chem. Rev. **112**, 108 (2012).
[3]K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 18 (2000).
[4]A. I. Krylov, Chem. Phys. Lett. **350**, 522 (2001).
[5]G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).
[6]R. Shephard and M. Minkoff, Int. J. Quantum Chem. **106**, 3190 (2006).
[7]L. Bytautas, T. Nagata, M. S. Gordon, and K. Ruedenberg, J. Chem. Phys. **127**, 164317 (2007).
[8]L. Bytautas, N. Matsunaga, T. Nagata, M. S. Gordon, and K. Ruedenberg, J. Chem. Phys. **127**, 204313 (2007).
[9]L. Bytautas and K. Ruedenberg, J. Chem. Phys. **130**, 204101 (2009).
[10]L. Bytautas and K. Ruedenberg, J. Chem. Phys. **132**, 074109 (2010).
[11]L. Bytautas, N. Matsunaga, and K. Ruedenberg, J. Chem. Phys. **132**, 074307 (2010).
[12]L. Bytautas, N. Matsunaga, G. E. Scuseria, and K. Ruedenberg, J. Phys. Chem. A **116**, 1717 (2012).
[13]B. Csontos, B. Nagy, J. Csontos, and M. Kállay, J. Phys. Chem. A **117**, 5518 (2013).
[14]J. S. Boschen, D. Theis, K. Ruedenberg, and T. L. Windus, Theor. Chem. Acc. **133**, 1425 (2014).
[15]K. H. Marti and M. Reiher, Phys. Chem. Chem. Phys. **13**, 6750 (2011).
[16]D. W. Small and M. Head-Gordon, Phys. Chem. Chem. Phys. **13**, 19285 (2011).
[17]C. F. Bunge, J. Chem. Phys. **125**, 014107 (2006).
[18]M. L. Abrams and C. D. Sherrill, Chem. Phys. Lett. **412**, 121 (2005).
[19]M. L. Abrams and C. D. Sherrill, Mol. Phys. **103**, 3315 (2005).
[20]J. Ivanic and K. Ruedenberg, Theor. Chem. Acc. **106**, 339 (2001).
[21]L. Bytautas and K. Ruedenberg, Chem. Phys. **356**, 64 (2009).
[22]F. A. Evangelista, J. Chem. Phys. **140**, 124114 (2014).
[23]J. Ivanic, J. Chem. Phys. **119**, 9364 (2003).
[24]F. A. Evangelista, E. Prochnow, J. Gauss, and H. F. Schaefer III, J. Chem. Phys. **132**, 074107 (2010).
[25]D. Zgid and M. Nooijen, J. Chem. Phys. **128**, 144116 (2008).

[26]J. J. Phillips and D. Zgid, J. Chem. Phys. **140**, 241101 (2014).
[27]G. Li Manni, D. Ma, F. Aquilante, J. Olsen, and L. Gagliardi, J. Chem. Theory Comput. **9**, 3375 (2013).
[28]C. South, G. Schoendorff, and A. K. Wilson, Comput. Theor. Chem. **1040-1041**, 72 (2014).
[29]G. H. Booth, D. Cleland, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **135**, 084104 (2011).
[30]G. H. Booth, S. D. Smart, and A. Alavi, Mol. Phys. **112**, 1855 (2014).
[31]L. Bytautas, C. A. Jiménez-Hoyos, R. Rodríguez-Guzmán, and G. E. Scuseria, Mol. Phys. **112**, 1938 (2014).
[32]P. J. Knowles and B. Cooper, J. Chem. Phys. **133**, 224106 (2010).
[33]E. Neuscamman, J. Chem. Phys. **139**, 181101 (2013).
[34]A. Zen, E. Coccia, Y. Luo, S. Sorella, and L. Guidoni, J. Chem. Theory Comput. **10**, 1048 (2014).
[35]S. A. Varganov and T. J. Martínez, J. Chem. Phys. **132**, 054103 (2010).
[36]G. E. Scuseria, C. A. Jiménez-Hoyos, T. M. Henderson, K. Samanta, and J. K. Ellis, J. Chem. Phys. **135**, 124108 (2011).
[37]C. A. Jiménez-Hoyos, T. M. Henderson, T. Tsuchimochi, and G. E. Scuseria, J. Chem. Phys. **136**, 164109 (2012).
[38]C. A. Jiménez-Hoyos, R. Rodríguez-Guzmán, and G. E. Scuseria, J. Chem. Phys. **139**, 204102 (2013).
[39]L. Bytautas, T. M. Henderson, C. A. Jiménez-Hoyos, J. K. Ellis, and G. E. Scuseria, J. Chem. Phys. **135**, 044119 (2011).
[40]D. R. Alcoba, A. Torre, L. Lain, G. E. Massaccesi, and O. B. Oña, J. Chem. Phys. **139**, 084103 (2013).
[41]D. R. Alcoba, A. Torre, L. Lain, G. E. Massaccesi, and O. B. Oña, J. Chem. Phys. **140**, 234103 (2014).
[42]D. Alcoba, A. Torre, L. Lain, O. B. Oña, P. Capuzzi, M. Van Raemdonck, P. Bultinck, and D. Van Neck, J. Chem. Phys. **141**, 244118 (2014).
[43]T. Stein, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. **140**, 214113 (2014).
[44]T. M. Henderson, I. W. Bulik, T. Stein, and G. E. Scuseria, J. Chem. Phys. **141**, 244104 (2014).
[45]P. A. Johnson, P. W. Ayers, P. A. Limacher, S. De Baerdemacker, D. Van Neck, and P. Bultinck, Comput. Theor. Chem. **1003**, 101 (2013).
[46]K. Boguslawski, P. Tecmer, and P. W. Ayers, Phys. Rev. B **89**, 201106(R) (2014).
[47]K. Boguslawski, P. Tecmer, P. A. Limacher, P. A. Johnson, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck, J. Chem. Phys. **140**, 214114 (2014).
[48]P. A. Limacher, T. D. Kim, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, Mol. Phys. **112**, 853 (2014).
[49]S. Sharma, T. Yanai, G. H. Booth, C. J. Umrigar, and G. K.-L. Chan, J. Chem. Phys. **140**, 104112 (2014).
[50]H. Nakatsuji, Acc. Chem. Res. **45**, 1480 (2012).
[51]G. Gidofalvi and D. A. Mazziotti, J. Phys. Chem. A **118**, 495 (2014).
[52]G. Knizia and G. K.-L. Chan, J. Chem. Theory Comput. **9**, 1428 (2013).
[53]K. Boguslawski, P. Tecmer, O. Legeza, and M. Reiher, J. Phys. Chem. Lett. **3**, 3129 (2012).
[54]R. J. Bartlett and J. F. Stanton, Rev. Comput. Chem. **5**, 65 (1994).
[55]R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
[56]K. Boguslawski, P. Tecmer, G. Barcza, O. Legeza, and M. Reiher, J. Chem. Theory Comput. **9**, 2959 (2013).
[57]L. M. Mentel, R. van Meer, O. V. Gritsenko, and E. J. Baerends, J. Chem. Phys. **140**, 214105 (2014).
[58]K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert, Chem. Phys. **71**, 41 (1982).
[59]B. O. Roos, Adv. Chem. Phys. **69**, 399 (1987).
[60]M. W. Schmidt and M. S. Gordon, Annu. Rev. Phys. Chem. **49**, 233 (1998).
[61]A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (McGraw-Hill, 1989).
[62]L. Bytautas, J. M. Bowman, X. Huang, and A. J. C. Varandas, Adv. Phys. Chem. **2012**, 679869 (2012).
[63]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. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).
[64]B. M. Bode and M. S. Gordon, J. Mol. Graphics Modell. **16**, 133 (1999).
[65]D. S. Koltun and J. M. Eisenberg, *Quantum Mechanics of Many Degrees of Freedom* (John Wiley & Sons, New York, 1988).
[66]M. Couty and M. B. Hall, J. Phys. Chem. A **101**, 6936 (1997).
[67]A. C. Hurley, J. Lennard-Jones, and J. A. Pople, Proc. R. Soc. London, Ser. A **220**, 446 (1953).
[68]W. Kutzelnigg, J. Chem. Phys. **40**, 3640 (1964).
[69]A. J. Coleman, Rev. Mod. Phys. **35**, 668 (1963).
[70]M. Casula and S. Sorella, J. Chem. Phys. **119**, 6500 (2003).
[71]W. A. Goddard III, T. H. Dunning, Jr., W. J. Hunt, and P. J. Hay, Acc. Chem. Res. **6**, 368 (1973).
[72]P. Jeszenszki, P. R. Nagy, T. Zoboki, Á. Szabados, and P. R. Surján, Int. J. Quantum Chem. **114**, 1048 (2014).
[73]K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).
[74]K. G. Wilson, Rev. Mod. Phys. **55**, 583 (1983).
[75]W. Klopper, J. Chem. Phys. **102**, 6168 (1995).
[76]L. Bytautas and K. Ruedenberg, J. Chem. Phys. **124**, 174304 (2006).
[77]L. Bytautas and K. Ruedenberg, Mol. Phys. **100**, 757 (2002).
[78]L. Bytautas, J. Ivanic, and K. Ruedenberg, J. Chem. Phys. **119**, 8217 (2003).
[79]H. Nakatsuji and H. Nakashima, J. Chem. Phys. **142**, 194101 (2015).
[80]G. K.-L. Chan, M. Kállay, and J. Gauss, J. Chem. Phys. **121**, 6110 (2004).
[81]F. Faglioni and W. A. Goddard III, Int. J. Quantum Chem. **73**, 1 (1999).
[82]D. W. Small and M. Head-Gordon, J. Chem. Phys. **130**, 084103 (2009).
[83]R. J. Bartlett, J. Phys. Chem. **93**, 1697 (1989).
[84]M. W. Schmidt, J. Ivanic, and K. Ruedenberg, J. Phys. Chem. A **114**, 8687 (2010).
[85]J. M. Merritt, V. E. Bondybey, and M. C. Heaven, Science **324**, 1548 (2009).
[86]M. W. Schmidt, J. Ivanic, and K. Ruedenberg, J. Chem. Phys. **140**, 204104 (2014).
[87]L. Bytautas and K. Ruedenberg, J. Chem. Phys. **122**, 154110 (2005).
[88]C. D. Sherrill and P. Piecuch, J. Chem. Phys. **122**, 124104 (2005).
[89]S. Shaik, D. Danovich, W. Wu, P. Su, H. S. Rzepa, and P. C. Hiberty, Nat. Chem. **4**, 195 (2012).
[90]A. J. C. Varandas, J. Chem. Phys. **129**, 234103 (2008).
[91]S. Sharma, J. Chem. Phys. **142**, 024107 (2015).
[92]See supplementary material at http://dx.doi.org/10.1063/1.4929904 for the numerical data describing wave functions and convergence patterns in the case of beryllium dimer, as well as molecular orbitals for various wave functions for N_2 and C_2 molecules.
[93]R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).
[94]P. R. Taylor, "Coupled-cluster methods in quantum chemistry," in *Lecture Notes in Quantum Chemistry, European Summer School*, edited by B. O. Roos (Springer-Verlag, Berlin, 1994).
[95]M. Nooijen, K. R. Shamasundar, and D. Mukherjee, Mol. Phys. **103**, 2277 (2005).
[96]S. M. Casey and D. G. Leopold, J. Phys. Chem. **97**, 816 (1993).
[97]L. Bytautas, Croat. Chem. Acta **86**, 453 (2013).