UC Berkeley

UC Berkeley Previously Published Works

Title

Non-orthogonal configuration interaction for the calculation of multielectron excited states.

Permalink

<https://escholarship.org/uc/item/2ks868dj>

Journal

The Journal of chemical physics, 140(11)

ISSN

0021-9606

Authors

Sundstrom, Eric J
Head-Gordon, Martin

Publication Date

2014-03-01

DOI

10.1063/1.4868120

Peer reviewed

Non-Orthogonal Configuration Interaction for the Calculation of Multielectron Excited States

Eric J. Sundstrom^{1, a)} and Martin Head-Gordon¹

*Department of Chemistry University of California, Berkeley Berkeley, CA 94720 and
Chemical Sciences Division Lawrence Berkeley National Laboratory Berkeley, CA,
94720 Phone: 510-642-5957 Fax: 510-643-1255*

(Dated: 9 February 2014)

We apply Non-orthogonal Configuration Interaction (NOCI) to molecular systems where multielectron excitations, in this case double excitations, play a substantial role: the linear polyenes and β -carotene. We demonstrate that NOCI when applied to systems with extended conjugation, provides a qualitatively correct wavefunction at a fraction of the cost of many other multireference treatments. We also present a new extension to this method allowing for purification of higher-order spin states by utilizing Generalized Hartree-Fock (GHF) Slater determinants and the details for computing $\langle S^2 \rangle$ for the ground and excited states.

I. INTRODUCTION

With vital applications in the photophysics of vision, photosynthesis, and solar energy conversion, proper modeling of molecular excited states is an important challenge for quantum chemistry. Multielectron excitations are particularly interesting, as they play a significant role in molecules with extended conjugation, such as those in biology¹ and those used in organic photovoltaics², and are a necessary component for understanding multiexciton formation^{3,4}, which could potentially lead to the enhancement of solar cell efficiencies. Modeling of multielectron excitations is difficult because the most commonly used and computationally efficient ground-state methods, when applied to excited states, fail to capture excitations of this character.

The most commonly used electronic structure method to compute excited states is the time dependent variety of density functional theory (TDDFT)^{5,6}, but because the most computationally tractable solution to the TDDFT equations utilize a linear response formalism they are unable to treat multielectron excitations.^{7,8} Although there have been a number of developments toward a DFT-based theory to treat double excitations⁹⁻¹¹ these methods are not in common use and still suffer from many of the same issues which plague ground-state DFT: choice of functional, approximate functionals may yield non-variational energies, etc.

The coupled-cluster formalism is a successful ground state theory which when expanded to treat excited states is able to capture double excitations. These methods include equation-of-motion coupled-cluster¹² and linear response theories based upon approximations of this theory, CIS(D)¹³, CC2, and CC3¹⁴. While coupled cluster is formally exact, in practice, the excitation level must be truncated at a given number of excited electrons, typically single and double excitations (EOM-CCSD). In order to provide a satisfactory level of correlation for an ex-

cited state with an M electron excitation, an $M + 1$ level of truncation is required¹⁵. Using a high-level truncation is problematic because the scaling of these methods are N^5 (CIS(D) and CC2), N^6 (EOM-CCSD), N^7 (CC3 and perturbative triples varieties of EOM-CC), and N^8 (EOM-CCSDT) where N is the size of the atomic orbital (AO) basis. Thus even a double excitation requires a 7th or 8th order scaling method.

Spin-flip (SF) methods, which have been utilized in configuration interaction (CI)¹⁶, DFT¹⁷, and CC¹⁸ frameworks are another set of tools for treating multielectron excitations which do not suffer from the high scaling of the CC formalism. These methods begin with a high spin reference and then employ linear combinations of spin-flipping excitations ($\alpha \rightarrow \beta$) to access states with the desired M_s value, such as the ground state. The first of these methods was SF-CIS (spin-flip configuration interaction singles); this method did not provide spin eigenstates and thus SC-SF-CIS¹⁹ (spin-complete SF-CIS) was developed to ameliorate this problem. SF-XCIS²⁰, which we shall apply to systems in this paper, is an extension to SC-SF-CIS which includes single excitations from the doubly-occupied space to the virtual space for all occupations of two electrons in the singly-occupied space. Restricted active space spin-flip configuration interaction variants (RAS-SF)²¹⁻²³ are the natural extension of SF-XCIS, which include larger numbers of spin-flips and use RAS techniques to include larger active spaces. Despite their success, the RAS-SF family of methods still suffers from some deficiencies: the excitation energies can be sensitive to the quality of the high-spin reference, these methods lack a full treatment of dynamic correlation, and these methods have exponential scaling with respect to active space size.

In order to properly treat multielectron excitations, quantum chemists typically resort to using a multiconfigurational wavefunction; of these, the most popular is the complete active space self-consistent-field (CASSCF) method²⁴, and its second-order perturbation-corrected variant (CASPT2)²⁵. CASSCF involves optimizing the orbitals over all the configurations within an active space and thus scales exponentially with the active space size.

^{a)}Electronic mail: eric.jon.sundstrom@berkeley.edu

This poor scaling limits this method to small basis sets and since appropriate active spaces generally grow with molecular size, to fairly small molecules. Substantial effort has been dedicated to reducing the exponential scaling of CASSCF; one particularly promising direction is the Density Matrix Renormalization Group (DMRG)²⁶. DMRG when paired with the CASSCF problem, effectively replaces the CI solver of CASSCF with the DMRG algorithm which is then able to solve the CASSCF problem exactly in an active space with non-exponential cost for pseudo-linear systems. For 3-dimensional systems DMRG-CASSCF scales exponentially but still favorably with respect to traditional CASSCF solvers.

In this work we present an alternative multiconfigurational wavefunction for treating excited states, non-orthogonal configuration interaction (NOCI)^{27,28}. NOCI scales as $\mathcal{O}(M^2 \max(n_e^3, N^2))$, where M is the number of determinants included, n_e is the number of electrons, and N is the size of atomic orbital basis. As we will show, the minimal size of M required to obtain qualitative results can be remarkably small for interesting molecules and thus this method is tractable for even large systems. In traditional CI approaches, one uses excitations out of the reference as the basis for interaction, typically truncating at some number of excitations; singles, doubles, triples, etc. This basis is orthogonal because each state is formed from a combination of the canonical occupied and virtual orbitals which are mutually orthogonal spaces. Orthogonality is not a necessary requirement to perform CI, but chosen because of the ease with which one can compute the Hamiltonian matrix and thus diagonalize it to compute the energies and wavefunctions of the states.

When using different Hartree-Fock solutions as the basis for CI, as is done throughout this paper, the configurations are non-orthogonal. Non-orthogonality makes computation of the Hamiltonian in this basis slightly more complicated, but also potentially allows the basis to span more of the relevant configuration space than a truncated orthogonal one of the same size would. The NOCI wavefunction is a linear combination of multiple Slater determinants and thus applicable to both ground-state and excited-state multireference problems. NOCI may also be used to treat ground-state strong correlation problems, but that will not be our focus here. In order to provide more spin-pure excited states, we will be utilizing generalized Hartree Fock (GHF)^{29,30} solutions. In GHF, the single particle functions are a linear combination of both up (α) and down (β) spin $|\chi_i(\mathbf{r}, \xi)\rangle = |\psi_i^\alpha(\mathbf{r})\rangle |\alpha(\xi)\rangle + |\psi_i^\beta(\mathbf{r})\rangle |\beta(\xi)\rangle$ ²⁹. Despite considerable discourse on symmetry breaking in HF^{31–35}, many of the broken-symmetry methods (other than simple unrestricted) never gained popularity within the quantum chemistry community, likely due to complications with interpretation (i.e. bad quantum numbers) and computational concerns.

Specifically we would like to point to the recent work of Scuseria and collaborators on the projected Hartree-Fock method (PHF)^{36–39} (which is related to both NOCI and

the older EHF)⁴⁰. After being relatively ignored by the electronic structure community, GHF wavefunctions have been used by Scuseria and collaborators in their work on the projected Hartree-Fock method (PHF)^{36–39}. In PHF, one applies projection operators to a broken-symmetry determinant to form an effective Hamiltonian which is then diagonalized until self-consistency. This scheme is a variation after projection (VAP) scheme which allows one to obtain the variationally minimum energy with respect to these broken-symmetry determinants. However, the variationally minimized wavefunction obtained only describes the ground state and thus PHF in the current form is not used to compute excited states for molecular systems (although research in this area is ongoing)⁴¹. Despite the recent success of this method it suffers from two flaws: it is neither size-consistent nor size-extensive³⁶. By size-consistency we mean that the energy evaluated on two non-interacting fragments, A and B , is equal to the sum of their individually evaluated energies: $E_{AB} = E_A + E_B$. Size-extensivity implies that when treating an infinite system, a method calculates a correlation energy per particle which does not vanish. A final important property of an excited state method is size-intensivity, which denotes the ability of a method to produce excited state energies which do not change with the addition of non-interacting fragments.

It is well known that orthogonal varieties of CI truncated at a particular level of excitation beyond singles, e.g. configuration interaction single and doubles excitations (CISD), are variational but not size-consistent nor size-intensive. On the other hand NOCI or an orthogonal version of the same method may be made size-consistent and the excitation energies made size-intensive. This is because a non-interacting fragment which by its very nature must not participate in the excitation will not affect the calculation of the excitation energy i.e. the Hamiltonian will be block diagonal in the fragments. Of course this has the caveat that the number of excited determinants required to obtain the same accuracy for the two fragment calculation, assuming the fragments are identical, would be the square of that required for the independent systems. We should also state that for an interacting system the number of determinants necessary would also increase, likely in a non-linear way. NOCI does not provide size-extensive correlation energies for infinite completely interacting systems, but this limitation does not prevent us from studying the bulk of chemistry.

The remainder of this article is organized as follows: 1) we shall recapitulate the basic theory of NOCI, 2) describe our new spin-purification scheme and discuss our choice of CI basis, 3) we shall explore NOCI's effectiveness in calculating excited states by applying the method to the first four all-trans linear polyenes and the β -carotene molecule, 4) finally we shall draw some conclusions about how this theory fits into the current quantum chemistry toolbelt and discuss some future extensions for the theory.

II. THEORY

The notation utilized in this work is as follows: NOCI wavefunctions: Φ ; Slater determinants: Ψ with indices w, x to distinguish between different solutions; Spatial Molecular Orbitals (MO): ψ ; Spatial Atomic Orbitals (AO): ω ; AO Indices: lower case Greek $\mu, \nu, \dots$; Virtual MO indices: a, b ; Occupied MO indices: i, j ; Generic MO indices: l, k . Bold case is employed for matrices. We employ the Einstein summation convention and the tensor notation with both covariant (subscript) and contravariant (superscript) indices of Head-Gordon et al.⁴²

A. Basics

The general idea of NOCI is to build and diagonalize the Hamiltonian in a basis of Slater determinants, henceforth basis states, which are in general non-orthogonal. The non-orthogonality of the basis states introduces an overlap metric, $\mathbf{S}$, into the standard CI eigenvalue problem:

$$\mathbf{H}\mathbf{D} = \mathbf{S}\mathbf{D}\mathbf{E} \quad (1)$$

where $\mathbf{H}$ is the Hamiltonian matrix, $\mathbf{D}$ are the expansion coefficients and $\mathbf{E}$ is a diagonal matrix of the energies. Thus NOCI requires one to compute the overlap and Hamiltonian between two generally non-orthogonal basis states; our method for this is outlined below for more details see Reference 28. Let us consider two solutions (Slater determinants):

$$|^w\Psi\rangle \quad \text{and} \quad |^x\Psi\rangle \quad (2)$$

each formed from n orthogonal occupied molecular orbitals:

$$|^w\phi_i\rangle = \sum_{\mu} |\omega_{\mu}\rangle {}^wC_{\cdot i}^{\mu} \quad (3)$$

$$|^x\phi_i\rangle = \sum_{\mu} |\omega_{\mu}\rangle {}^xC_{\cdot i}^{\mu} \quad (4)$$

These orbitals are orthogonal within a given set but non-orthogonal between the sets. We would like to form two sets of orbitals such that the occupied overlap matrix between the determinants is diagonal:

$$\langle {}^w\phi_i | {}^x\phi_j \rangle = s_i \delta_{ij} \quad \text{and} \quad 0 \leq |s_i| \leq 1 \quad (5)$$

To obtain these sets we use the corresponding orbitals of Amos and Hall^{43,44}, also known as Löwdin-Paired orbitals⁴⁵. The classic equations (see the appendix of Amos and Hall⁴³) suggest diagonalizing the product ${}^{wx}\mathbf{S}^\dagger {}^{wx}\mathbf{S}$, where ${}^{wx}\mathbf{S} = {}^wC_{\cdot i}^{*\mu} S_{\mu\nu} {}^xC_{\cdot j}^{\nu}$ and $S_{\mu\nu}$ is the atomic orbital overlap matrix. We found it more numerically stable to use the singular value decomposition

(SVD):

$${}^{wx}\mathbf{S} = {}^wC_{\cdot i}^{*\mu} S_{\mu\nu} {}^xC_{\cdot j}^{\nu} \quad (6)$$

$${}^{wx}\mathbf{S} = \mathbf{U}\sigma\mathbf{V}^t \quad (\text{SVD}) \quad (7)$$

$${}^x\tilde{\mathbf{C}} = {}^x\mathbf{C}\mathbf{V}^t \quad (8)$$

$${}^w\tilde{\mathbf{C}} = {}^w\mathbf{C}\mathbf{U}^t \quad (9)$$

where σ is a diagonal matrix of the singular values. It must be noted that in the formulation presented above there is a loss of phase information in the eigenvalues of ${}^{wx}\mathbf{S}$. Early proponents of these orbitals computed the determinants of both transformation matrices in order to preserve the phase. In our implementation, we compute the eigenvalues of the non-Hermitian matrix, ${}^{wx}\mathbf{S}$, and compare the product of its diagonal to that of the transformed set, changing the sign as necessary. After each pair of basis states are made biorthogonal, a set of generalized Slater-Condon rules^{28,43} may be used to compute the Hamiltonian matrix elements and then the general eigenvalue problem can be solved.

B. Extended Spin-Purification Scheme

Here we present a new spin-purification scheme for NOCI utilizing GHF-type determinants. In prior implementations of NOCI, in an attempt to purify spin-contaminated UHF solutions, two degenerate configurations, per one original UHF determinant, were included in the expansion: the excess α configuration and the excess β configuration²⁸. When applied to only one UHF determinant this concept is known as the Half-Projected Hartree-Fock Method, an approximation to PHF in which one applies a spin projection operator that selects for either the even or odd quantum number $\mathcal{S}$. Consider a given UHF solution which is exactly a linear combination of two spin-pure components such as a singlet and a triplet, by interacting the excess α and excess β determinants one is able to obtain the two spin-pure configurations by adding and subtracting the determinants appropriately. Within NOCI, the diagonalization of the Hamiltonian provides the proper signs to accomplish this.

In order to extend this spin-purification to determinants which are linear combinations of more than two spin components, we interact non-collinear (GHF) determinants. Applying the same corresponding orbital treatment as above, one obtains a similar set of Slater-Condon rules for matrix elements of the Hamiltonian operator in a basis of GHF solutions. In order to use these wavefunctions for spin-purification, we apply an AO based unitary rotation parameterized by an angle θ to the coefficient matrix which changes the direction of the S_z axis without changing the energy. Choosing a rotation angle of $\frac{\pi}{2}$ transforms the α component of each orbital into the β component of each orbital and is exactly equivalent to the UHF spin-purification scheme discussed above. Al-

though these basis states are degenerate with their unrotated counterparts, they provide a new set of overlap and Hamiltonian matrix elements with the other basis states in the expansion and provide more variational flexibility in the CI. In this work we shall use a quadrature of n_q points equidistant between 0 and $\frac{\pi}{2}$ to allow spin-contaminated states to purify. That is if one originally began with 5 UHF determinants and $n_q = 5$, the NOCI Hamiltonian and overlap would be square matrices of dimension 25. This idea is very similar to the discretization

of the gauge integration used for spin-projection in the PHF method³⁶ and thus we suspect that a small number of quadrature points will remove much of the spin-contamination.

In order to compute $\langle S^2 \rangle$ for a given state we evaluate the $\langle S^2 \rangle$ matrix elements for each pair of GHF determinants and then use the eigenvectors of the CI diagonalization to calculate the value of a given state. The matrix element for a pair of determinants ($|^w\Psi\rangle$ and $|^x\Psi\rangle$) is:

$$\begin{aligned}\langle ^w\Psi | S^2 | ^x\Psi \rangle &= \langle ^w\Psi | S_x^2 | ^x\Psi \rangle + \langle ^w\Psi | S_y^2 | ^x\Psi \rangle \\ &\quad + \langle ^w\Psi | S_z^2 | ^x\Psi \rangle \\ \langle ^w\Psi | S_\sigma^2 | ^x\Psi \rangle &= \langle ^w\Psi | S_\sigma | ^w\Psi \rangle \langle ^w\Psi | ^x\Psi \rangle \langle ^x\Psi | S_\sigma | ^x\Psi \rangle \\ &\quad + \sum_{ia} \langle ^w\Psi | S_\sigma | ^w\Psi_i^a \rangle \langle ^w\Psi_i^a | ^x\Psi \rangle \langle ^x\Psi | S_\sigma | ^x\Psi \rangle \\ &\quad + \sum_{jb} \langle ^w\Psi | S_\sigma | ^w\Psi \rangle \langle ^w\Psi | ^x\Psi_j^b \rangle \langle ^x\Psi_j^b | S_\sigma | ^x\Psi \rangle \\ &\quad + \sum_{ia,jb} \langle ^w\Psi | S_\sigma | ^w\Psi_i^a \rangle \langle ^w\Psi_i^a | ^x\Psi_j^b \rangle \langle ^x\Psi_j^b | S_\sigma | ^x\Psi \rangle\end{aligned}$$

where $\sigma = x, y, z$. Here we inserted the resolution of the identity for each determinant ($|^w\Psi\rangle$) noting that the occupied and virtual spaces of each determinant are orthogonal, the occupied spaces between the determinants have been made biorthogonal (corresponding orbitals), and the virtual spaces remain non-orthogonal between the two determinants. In order to simplify these resolutions, we have used the property that the spin component operators are one-electron operators and therefore only the ground-state and single-excitations couple through them. In Appendix A, we present a derivation for each matrix element in the above equation, resulting in a computation of $\langle S^2 \rangle$ for each determinantal pair which scales as N^3 where N is the size of the atomic orbital basis.

C. Basis States

Below, we present the algorithm used to obtain the set of non-orthogonal Slater determinants which form the basis for NOCI. We begin by converging a symmetry-preserving RHF solution. From this “ground-state”, we form some subset of singly- and doubly-excited orthogonal determinants which are of chemical relevance. These explicitly unrestricted determinants are then allowed to optimize using the Maximum Overlap Method (MOM)⁴⁷ combined with the Direct Inversion of the Iterative Subspace (DIIS) extrapolation method^{48,49} until the DIIS error vector is less than 1.0×10^{-8} . It has been discussed at length^{28,47,50–53} that these states provide approxima-

tions to excited states despite not being orthogonal to the ground state.

The orbital optimization allows for both valence and core relaxation in the presence of the constrained occupancy of what would be virtual orbitals. This relaxation is an important component of the non-orthogonal expansion of the wavefunction and results in a more compact representation than its orthogonal counterpart. As we shall show, an orthogonal set of the same determinants performs poorly in the case of wavefunctions with multireference character. Admittedly, this relaxation is only occurring at the mean field level and thus cannot fully account for these correlations. The relaxation in NOCI is not like that of a CASSCF type wavefunction where all the orbitals are optimized in the presence of all the configurations included in the active space, but rather is a basis state specific optimization.

A development version of QCHEM 3.2⁵⁴ utilizing the Armadillo linear algebra library⁵⁵ and 6-31G*⁵⁶ atomic orbital basis was used for all calculations. All geometries were obtained from optimizations performed with the same basis set using the B3LYP functional⁵⁷.

III. RESULTS

A. Polyenes

The linear polyenes have long been used for testing new excited state methods in electronic structure theory. This

is primarily due to the fact that two of the low singlet excited states are of significantly different character. The $1^1B_u^+$ state is experimentally “bright” (dipole allowed) and primarily of single excitation character, while the $2^1A_g^-$ state is “dark” (dipole forbidden) and has considerable double excitation character. These two excitations will be the main focus of this study because they are low in the energy spectrum for the whole range of systems investigated here. For the longer polyenes, it has been shown that the $1^1B_u^-$ state becomes lower in energy than the $1^1B_u^+$ state for $n = 7$, and by $n = 11$ (where n is defined by $C_{2n}H_{2n+2}$), the $3^1A_g^-$ state is also below the $1^1B_u^+$ state⁵⁸.

Due to the chemical similarities of the systems studied in this work (all-*trans* polyenes and β -carotene), the basis states chosen to represent the final wavefunction were the same. We limited the number of basis states to 10 for two reasons: this allowed us to use a higher number of quadrature points for the new spin-purification scheme and showed that even in this greatly reduced space, NOCI performs quite admirably. The set of basis states consisted of:

1. The symmetric RHF determinant.
2. $b_g \rightarrow a_u^*$ (HOMO $\rightarrow$ LUMO).
3. $a_u \rightarrow a_u^*$ (HOMO-1 $\rightarrow$ LUMO).
4. $b_g \rightarrow b_g^*$ (HOMO $\rightarrow$ LUMO+1).
5. $a_u \rightarrow b_g^*$ (HOMO-1 $\rightarrow$ LUMO+1).
6. $(b_g)^2 \rightarrow (a_u^*)^2$ (HOMO² $\rightarrow$ LUMO²).
7. $(b_g)^2 \rightarrow (b_g^*)^2$ (HOMO² $\rightarrow$ LUMO+1²).
8. $(a_u)^2 \rightarrow (a_u^*)^2$ (HOMO-1² $\rightarrow$ LUMO²).
9. $b_g a_u \rightarrow (a_u^*)^2$ (HOMO, HOMO-1 $\rightarrow$ LUMO²).
10. $(b_g)^2 \rightarrow a_u^* b_g^*$ (HOMO² $\rightarrow$ LUMO, LUMO+1).

Specifically determinant 6 was included in order to target the known double-excitation character of the $2^1A_g^-$ state in the polyenes. For all systems, the excited determinants were optimized within the UHF framework but despite this some of the double excitation 6-8 were found to be solutions of the RHF equations; the determinants had equal α and β coefficients.

1. Butadiene

For theoreticians, even the simplest of the linear polyenes, butadiene, has long been a challenge. Many theoretical techniques have been applied to butadiene, including configuration interaction singles (CIS), TDDFT⁵⁹, second-order algebraic diagrammatic construction (ADC(2))⁶⁰, equation of motion coupled-cluster (EOM-CC)^{61,62}, CC2 and CC3⁶³, symmetry adapted cluster configuration interaction (SAC-CI)^{64,65}, CASSCF^{66,67}, CASPT2^{66,68}, restricted active space SCF (RASSCF)⁶⁹, RASCI-SF²³, and multireference Møller-Plesset (MRMP)⁶⁷. The relative ordering of the $1^1B_u^+$ and $2^1A_g^-$ states in butadiene has a long record of controversy in theoretical work, due to the sheer number of methods applied. Recently, this controversy has been

laid to rest by the work of Watson and Chan⁶². They applied high-level EOM-CC theory, including extrapolation to the complete basis set limit and estimates for core excitations. They found that the vertical excitation energy of the $1^1B_u^+$ state (6.21 ± 0.02 eV) lies below the $2^1A_g^-$ state (6.41 ± 0.07 eV) in contrast to some previous results. In their work, they conclude that the most cited experimental value for the $1^1B_u^+$ excitation, 5.92 eV from electron impact energy loss spectroscopy⁷⁰, is likely not a truly vertical excitation and that the experimental results for the notoriously difficult dipole forbidden $2^1A_g^-$ state are not precise enough for comparison.

In order to assess the value of utilizing the non-collinear (GHF) quadrature for spin-purification, we applied the method with various values of n_q to butadiene. In Table I, we present both the vertical excitation energies and the $\langle S^2 \rangle$ values. The character of all of the NOCI states were assigned by inspecting the components of the CI eigenvector for large weights on determinants of the approximately the correct symmetry and character. When only applying the two state spin-purification scheme ($n_q = 2$), despite having a reasonable excitation energy, the $2^1A_g^-$ state is fairly spin contaminated; this is likely due to the fact that this state contains a linear combination of two open shell singlets i.e. $a_u \rightarrow a_u^*$ (HOMO-1 $\rightarrow$ LUMO) and $b_g \rightarrow b_g^*$ (HOMO $\rightarrow$ LUMO+1), with $\langle S^2 \rangle$ values of 1.07 and 1.14 respectively. It looks as though increasing the non-collinear basis can only reduce spin contamination to a finite value without inclusion of more states. Small eigenvalues of the overlap matrix in the NOCI generalized eigenvalue problem are representative of linear dependencies in our basis states. If one investigates these values as n_q is increased from 10 to 30, the number of these below a threshold increases by almost the same number as the number of new determinants. This indicates that for this system, the Hilbert space has been saturated and increasing the quadrature beyond as $n_q = 10$ will not change the answer greatly. The fact that this saturation occurs before $\langle S^2 \rangle = 0$ is interesting and will be investigated in the future.

For butadiene it should be noted that the ordering of the two singlet excited states is incorrect with respect to the results of Watson and Chan⁶². This can be explained by the fact that we are primarily targeting static correlation in our basis states, the primary type of correlation for the $2^1A_g^-$ state; a proper treatment of the $1^1B_u^+$ state requires more dynamic correlation.

2. $C_{2n}H_{2n+2}$ with $2 \leq n \leq 11$

We shall now test NOCI on the longer polyenes comparing to both reference methods and methods with similar scaling. In Table II, the vertical excitation energies of the linear polyenes with NOCI $n_q = 2$ and $n_q = 10$ are shown in comparison with other obtained theoretical values. We have included CASPT2 because of this method’s popularity in treating multireference excited states and

TABLE I. Absolute ground state energies (a.u.) and vertical excitation energies (eV) and $\langle S^2 \rangle$ for butadiene with various levels of GHF quadrature (n_q).

State	$1^1A_g^-$		$1^1B_u^+$		$2^1A_g^-$	
	a.u.	$\langle S^2 \rangle$	eV	$\langle S^2 \rangle$	eV	$\langle S^2 \rangle$
$n_q = 2$	-154.93643	0.001	7.63	0.007	6.72	0.353
$n_q = 5$	-154.93647	0.000	7.61	0.002	6.63	0.096
$n_q = 10$	-154.93647	0.000	7.61	0.002	6.63	0.095
$n_q = 30$	-154.93647	0.000	7.61	0.002	6.63	0.094

FIG. 1. $\langle S^2 \rangle$ values for the linear polyenes for NOCI with a non-collinear (GHF) quadrature (n_q) of 2 and 10. Increasing the quadrature decreases the spin-contamination by a factor of 3 for all states. The ground states are not shown here as they are very close to zero and remain so with increasing quadrature.

the more recent CASCI-MRMP results⁵⁸ in order to provide a benchmark. It should be noted that when comparing these results, the excitation energies will be dependent on many computational parameters such as geometry and basis set and thus the CASPT2 and CASCI-MRMP results cannot be directly compared with those calculated in this work. In order to provide a comparison with methods which have similar scaling and expected accuracy, we have also included results from CI Singles (CIS) and SF-XCIS²⁰, calculated using the same geometry and basis set. We have chosen not to include TDDFT results despite previous work⁵⁹ because, as stated early, TDDFT would not accurately reproduce double excitations and will suffer from the same failings as CIS. Table II also includes data for “OCI”, by which we mean an orthogonal version of CI including the same 10 states as NOCI albeit their unrelaxed counterparts.

First, we will consider the NOCI spin-purification scheme, comparing the results of $n_q = 2$ with that of $n_q = 10$ shows that the change in excitation energies increases as n increases, reaching a final value of ~ 0.3 eV for the $2^1A_g^-$ state of 11-ene. These changes in energy are directly correlated with the amount of original spin-contamination in these states at the $n_q = 2$

FIG. 2. Polyene singlet vertical excitation energies (eV), comparing the NOCI ($n_q = 10$) and OCI results. OCI is a multireference CI where the only included configurations are the same configurations as those included in the NOCI, albeit their orthogonal unrelaxed counterparts.

level. Figure 1 shows the change in $\langle S^2 \rangle$ for both excited states as a function of chain length for both levels of quadrature. Again the quadrature is able to reduce spin-contamination by about a factor of three for all states. It seems it may be beneficial to increase n_q as the amount of spin-contamination increases in the basis states but in order to maintain a consistent comparison we have remained at $n_q = 10$.

Let us begin now compare the energies of OCI to those of NOCI; Table II contains a column for both OCI and energy change with respect to NOCI ($n_q = 10$) ($\Delta(\text{OCI-NOCI})$) and Figure 2 contains the OCI and NOCI ($n_q = 10$) results plotted with respect to chain length. Clearly, OCI does not provide an accurate description of the doubly excited $2^1A_g^-$ state and the relaxation of the orbitals by NOCI decreases the energy of this state by greater than 1 eV for the whole set of molecules. NOCI is able to account for substantial additional correlation energy when compared to an orthogonal CI with the same number of determinants; i.e. NOCI contains a more compact representation of the true wavefunction and encompasses more of the relevant configuration space. The difference between OCI and NOCI is not as pronounced for the $1^1B_u^+$ state, where the state is primarily a single configuration and thus treated comparably by both OCI and NOCI. As the chain length increases (approximately at $n = 8$), the single configuration picture of the $1^1B_u^+$ breaks down and OCI and NOCI begin to diverge. This is an indication that the relaxation in the determinants which make up the $1^1B_u^+$ state in NOCI are beginning to take on character of more than one orthogonal state.

A subset of the results in Table II are included in Figure 3 in order to highlight the trends of these methods as the chain length increases. Moving on to the results for CIS when compared to the reference CASCI-MRMP, we see that CIS cannot qualitatively reproduce the excita-

TABLE II. Absolute ground state energies (a.u.), vertical excitation energies (eV), and $\langle S^2 \rangle$ shown in () for the all-*trans* polyenes $C_{2n}H_{2n+2}$ with $2 \leq n \leq 11$ from different theories. Here, OCI is a multireference CI where the only included configurations are the same configurations as those included in the NOCI, albeit their orthogonal unrelaxed counterparts.

n	State	CIS	SF-XCIS	NOCI ($n_q = 2$)	NOCI ($n_q = 10$)	OCI	$\Delta(\text{OCI-NOCI})$	CASCI- MRMP	CASPT2
2	$1^1A_g^-$		-154.98077	-154.93643 (0.00)	-154.93647 (0.00)	-154.93537	0.00110		
	$1^1B_u^+$	6.79	7.23	7.63 (0.01)	7.61 (0.00)	7.85	0.24	6.21 ^a	6.27 ^c
	$2^1A_g^-$	9.24	7.23	6.72 (0.35)	6.63 (0.10)	8.61	1.98	6.31	6.23
3	$1^1A_g^-$		-231.86320	-231.81963 (0.00)	-231.81965 (0.00)	-231.81768	0.00196		
	$1^1B_u^+$	5.69	6.22	6.57 (0.04)	6.51 (0.01)	6.38	-0.13	5.25 ^b	5.19
	$2^1A_g^-$	8.12	6.53	6.32 (0.29)	6.23 (0.08)	7.37	1.14	5.10	5.01
4	$1^1A_g^-$		-308.75426	-308.70398 (0.00)	-308.70398 (0.00)	-308.70264	0.00134		
	$1^1B_u^+$	4.97	5.65	5.59 (0.07)	5.52 (0.02)	5.67	0.16	4.57	4.38 ^d
	$2^1A_g^-$	7.23	5.36	5.22 (0.48)	5.09 (0.14)	6.69	1.60	4.26	4.42
5	$1^1A_g^-$		-385.63800	-385.58980 (0.00)	-385.58982 (0.00)	-385.58825	0.00157		
	$1^1B_u^+$	4.47	5.11	4.94 (0.09)	4.86 (0.03)	5.10	0.24	4.17	
	$2^1A_g^-$	6.54	4.90	4.73 (0.53)	4.59 (0.16)	6.27	1.68	3.68	
6	$1^1A_g^-$		-462.52693	-462.47573 (0.00)	-462.47575 (0.00)	-462.47427	0.00149		
	$1^1B_u^+$	4.11	4.83	4.45 (0.12)	4.35 (0.03)	4.70	0.34	3.87	
	$2^1A_g^-$	6.00	4.58	4.25 (0.67)	4.10 (0.21)	6.00	1.90	3.19	
7	$1^1A_g^-$		-539.41218	-539.36198 (0.00)	-539.36201 (0.00)	-539.36047	0.00154		
	$1^1B_u^+$	3.83	4.52	4.09 (0.14)	3.98 (0.04)	4.25	0.28	3.60	
	$2^1A_g^-$	5.55	4.25	3.93 (0.78)	3.77 (0.25)	5.37	1.60	2.80	
8	$1^1A_g^-$		-616.29917	-616.24832 (0.00)	-616.24837 (0.00)	-616.24692	0.00145		
	$1^1B_u^+$	3.61	4.34	3.80 (0.17)	3.67 (0.05)	4.05	0.37	3.38	
	$2^1A_g^-$	5.19	4.07	3.63 (0.92)	3.46 (0.33)	5.16	1.69	2.50	
9	$1^1A_g^-$		-693.18618	-693.13482 (0.00)	-693.13488 (0.00)	-693.13348	0.00140		
	$1^1B_u^+$	3.43	4.17	3.57 (0.21)	3.43 (0.06)	4.02	0.59	3.18	
	$2^1A_g^-$	4.88	3.91	3.39 (1.07)	3.19 (0.37)	5.63	2.44	2.25	
10	$1^1A_g^-$		-770.07316	-770.02148 (0.00)	-770.02157 (0.00)	-770.02012	0.00145		
	$1^1B_u^+$	3.29	4.06	3.39 (0.24)	3.24 (0.07)	3.89	0.66	3.00	
	$2^1A_g^-$	4.62	3.80	3.17 (1.25)	2.94 (0.48)	5.52	2.58	2.04	
11	$1^1A_g^-$		-846.96019	-846.90829 (0.01)	-846.90844 (0.00)	-846.90683	0.00160		
	$1^1B_u^+$	3.17	3.96	3.23 (0.29)	3.07 (0.09)	3.80	0.74	2.84	
	$2^1A_g^-$	4.40	3.71	3.11 (1.23)	2.83 (0.47)	5.31	2.49	1.86	

^a Reference 67. Basis set: Dunning “QZ3p” (Carbon:5s4p3d, Hydrogen:3s2p), Active Space: (4,8), Geometry: Experimental

^b Reference 58. Basis set: cc-pvDZ, Active Space: (10,10) Geometry: Møller-Plesset 2nd Order Perturbation Theory

^c Reference 71. Basis set: ANO (Carbon:4s3p2d, Hydrogen:3s2p)+diffuse functions, Active Space: Full π -space, Geometry: Experimental

^d Reference 66. Basis set: ANO (Carbon:4s3p1d, Hydrogen:2s1p)+diffuse functions, Active Space: Full π -space, Geometry: CASSCF

tion energy of the $2^1A_g^-$ state. It places the $2^1A_g^-$ state well above the $1^1B_u^+$ state for all systems studied. This is not just an error in the excitation energy; CIS cannot correctly describe the character of the $2^1A_g^-$ state, due to a lack of double excitations. This deficiency also appears in the wavefunction where the two largest CIS amplitudes for the $2^1A_g^-$ state are HOMO-1 $\rightarrow$ LUMO and HOMO $\rightarrow$ LUMO+1; this same type of behavior would occur in TDDFT as well. On the other hand CIS treats the $1^1B_u^+$ state quite well, paralleling the CASCI-MRMP results.

Unlike CIS, SF-XCIS, by the inclusion of a (2,2) active space treats the doubly excited character of the $2^1A_g^-$

excited state appropriately. For all systems, the largest component of the SF-XCIS wavefunction of the $2^1A_g^-$ state is the HOMO $\rightarrow$ LUMO doubly excited state, with smaller components on other states of the correct symmetry. Although SF-XCIS shows the two states reversing order twice between $2 \leq n \leq 4$ (when the correct result is one crossing around $n = 3$ (hexatriene)), the method does recover the correct ordering of these two states for the longer chains. Comparing NOCI with SF-XCIS, it is amazing that NOCI using only 10 determinants can compete with SF-XCIS, a method which includes more doubles excitations than NOCI, some triple

FIG. 3. Polyene singlet vertical excitation energies (eV), comparing the NOCI ($n_q = 10$) results with CIS and SF-XCIS. Solid lines: $1^1B_u^+$. Dashed lines: $2^1A_g^-$. Using CIS is insufficient to treat the $2^1A_g^-$ excited state. NOCI and SF-XCIS do comparably well for the shorter chains but NOCI outperforms SF-XCIS in the long chain limit.

excitations, and an increasing configuration space with molecular size. This again shows the power of using relaxed non-orthogonal configurations when compared to their orthogonal counter parts. For longer polyenes, the relative energy of these two states becomes constant in both NOCI and SF-XCIS, which is in contrast to the behavior of the reference method CASCI-MRMP, which for longer chains shows an increase in the separation of these two states. For SF-XCIS, this degradation indicates that one may need to increase the active space and potentially the number of spin-flips. For NOCI, one should use more than 10 determinants to treat a polyene chain of length $n = 11$ but for the purposes of illustrating the applicability of this method and for a fair comparison across these systems we thought this number of basis states to be sufficient.

B. β -Carotene

Carotenoids are an important class of biological molecules because they, along with chlorophylls, comprise a substantial fraction of the chromophores involved in light-harvesting in plants and cyanobacteria. Carotenoids provide two functions in photosynthetic organisms: they directly absorb light energy for use in photosynthesis and also participate in a photoprotection mechanism. For a thorough discussion of experimental and theoretical results for these systems, including the difficulty in peak assignment, we recommend the review of Polívka and Sundström.⁷² Having been featured in two prior electronic structure studies, we chose β -carotene (a carotenoid with 11 double bonds) as a sample carotenoid to examine the performance of NOCI on these biologically-relevant molecules. This is an interesting test case for NOCI, because the lower symmetry of this molecule lends itself naturally to the breaking of spatial symmetry within the excited state determinants. Historically the excited states of the carotenoids

TABLE III. Vertical excitation energies (eV) and $\langle S^2 \rangle$ shown in () for of the β -carotene with a variety of methods.

State	$2^1A_g^-$	$1^1B_u^+$	$\Delta E (1^1B_u^+ - 2^1A_g^-)$
$n_q = 10$	2.94 (0.53)	3.20 (0.08)	0.26
SF-XCIS	3.87	3.97	0.10
DMRG-CASSCF	2.99	> 4.31	> 1.32
DFT/MRCI	1.96	2.42	0.42

^a Reference 73.

^b Reference 75.

have been labeled S1, S2, etc; here we shall opt to reuse the polyene excited state nomenclature despite β -carotene having a lower point group symmetry. Table III shows the vertical excitation energies and $\langle S^2 \rangle$ values for β -carotene with NOCI, SF-XCIS and two other methods: Density Matrix Renormalization Group CASSCF (DMRG-CASSCF)⁷³, which was discussed in the Introduction, and DFT/MRCI^{74,75}, which we shall describe briefly now.

DFT/MRCI is a semi-empirical procedure combining DFT for dynamic correlation and MRCI for strong correlations and excited states. This method involves a number of features including an efficient parallelization scheme, an extensive configuration selection, and use of the resolution of the identity (RI) approximation for four-index integrals. Despite these advantages, DFT/MRCI still possesses drawbacks such as choice of functional, empirical parameters for forming the effective Hamiltonian and use of an empirical recipe to avoiding double counting of correlation effects. Even with the considerable configuration selection, the method requires millions of configuration state functions for a molecule of the size of β -carotene. Because this method does include dynamic correlation, we will consider it a possible benchmark, keeping in mind that it is difficult to estimate the errors for a method which is non-variational.

To our knowledge, these are the only two studies of β -carotene applying modern electronic structure methods which can qualitatively treat double excitations. TDDFT has been applied to study the $S_0 \rightarrow S_2$ transition⁷⁶, but as stated previously, standard linear-response TDDFT cannot accurately treat double excitations^{7,8} and thus should not be trusted for any transitions involving the $2^1A_g^-$ state, including investigations of potential energy surfaces.

In the DMRG-CASSCF work, the authors averaged over the four lowest eigenstates and thus within their calculation were unable to obtain the $1B_u^+$ state. They found a state ordering of $1^1A_g^- < 2^1A_g^- < 1^1B_u^- < 3^1A_g^-$ with excitation to the $3^1A_g^-$ at 4.31 eV, which gives a lower bound to the $1B_u^+$ state. This ordering is correct within the active space but CASSCF does not include any dynamic correlation outside of this space (e.g. σ - π), which is known to be important for the $1B_u^+$ state. In fact, DFT-MRCI, a method which includes these σ - π cor-

relations, orders the states as $1^1A_g^- < 2^1A_g^- < 1^1B_u^+ < 1^1B_u^- < 3^1A_g^-$. If we consider DFT/MRCI to include all correlation effects and DMRG-CASSCF to include all static correlation effects, then the difference between these two methods gives us an estimate for the total amount of dynamic correlation in each state: the $2^1A_g^-$ and $1^1B_u^+$ states have dynamic correlation of ~ 1 eV and > 2 eV, respectively. From this estimate, we can see that neither NOCI nor SF-XCIS (methods designed primarily to treat static correlation) are truly able to determine the state ordering as neither value is separated by the > 1 eV difference that would be added by a more complete treatment of dynamic correlation. In comparison to SF-XCIS, NOCI does provide a better estimate of the gap between these two states. Considering the difference in the complexity of the wavefunctions of these methods, it is striking that NOCI is able to achieve the results that it does. Comparing the results of both NOCI and SF-XCIS on β -carotene to that of the polyenes, we see that these methods treat this systems as a polyene of approximately length $n = 10$, which is consistent with the fact that the conjugation of β -carotene is broken by the slight twist in the polyene backbone of the molecule. Again the differences between NOCI and the reference could likely be reduced by the inclusion of more states; specifically one would like to include more double excitations, which are of known importance to molecules with this length of conjugation. We want to stress that in this current form of “hand”-picking excited determinants, NOCI cannot be expected to stand up against DFT/MRCI.

IV. CONCLUSIONS

We have shown that NOCI may be used to calculate excitation energies for double excitations in conjugated systems. Although the excitation energies presented in this work were not quantitative, we believe this method shows promise for future use in this direction. NOCI scales favorably compared to many other multireference methods and could be applied to systems even larger than those presented here. We applied a new spin-purification scheme resulting in better $\langle S^2 \rangle$ values and excitation energies for both the ground and excited states. NOCI also shows promise where orbital relaxation may be important such as singlet-triplet gaps in transition metal systems. We also feel this method could be applied to excited states with charge transfer character as the relaxation of the orbitals would allow for charge redistribution after “excitation”.

Further developments and generalizations of NOCI are possible in many directions. Currently, the most unsatisfactory portion of NOCI, as it was used in this study, is the selection of basis states by human input and can lead to biasing of the results. This is only exacerbated by the reliance of this procedure on the Maximum Overlap Method (MOM) which, because it involves a non-linear optimization can be difficult to converge to the correct

determinants. A black-box method for selection of basis states and a reduced dependence on MOM is our next target of research and progress is already underway in this direction. In this work, we used GHF determinants for spin-purification, but one could also search for non-collinear GHF determinants to provide more basis states. As a wavefunction based CI theory, NOCI would also benefit from many of the known routes to target an improved treatment of dynamic correlation, such as that used in CASPT2. Finally, we plan on implementing nuclear gradients in order to follow the potential energy surfaces of the excited states calculated above as well as orbital gradients, which would allow for an orbital optimized NOCI, very related to the resonating Hartree-

Fock approach of Fukutome⁷⁷.

ACKNOWLEDGMENTS

E.J.S. would like to thank the Head-Gordon group for fruitful discussions. This work was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Appendix A: Calculation of $\langle S^2 \rangle$

As discussed above the equation for computing $\langle S^2 \rangle$ between to non-orthogonal determinants may be written as:

$$\begin{aligned} \langle {}^w\Psi | S^2 | {}^x\Psi \rangle &= \langle {}^w\Psi | S_x^2 | {}^x\Psi \rangle + \langle {}^w\Psi | S_y^2 | {}^x\Psi \rangle \\ &\quad + \langle {}^w\Psi | S_z^2 | {}^x\Psi \rangle \\ \langle {}^w\Psi | S_\sigma^2 | {}^x\Psi \rangle &= \langle {}^w\Psi | S_\sigma | {}^w\Psi \rangle \langle {}^w\Psi | {}^x\Psi \rangle \langle {}^x\Psi | S_\sigma | {}^x\Psi \rangle \\ &\quad + \sum_{ia} \langle {}^w\Psi | S_\sigma | {}^w\Psi_i^a \rangle \langle {}^w\Psi_i^a | {}^x\Psi \rangle \langle {}^x\Psi | S_\sigma | {}^x\Psi \rangle \\ &\quad + \sum_{jb} \langle {}^w\Psi | S_\sigma | {}^w\Psi \rangle \langle {}^w\Psi | {}^x\Psi_j^b \rangle \langle {}^x\Psi_j^b | S_\sigma | {}^x\Psi \rangle \\ &\quad + \sum_{ia,jb} \langle {}^w\Psi | S_\sigma | {}^w\Psi_i^a \rangle \langle {}^w\Psi_i^a | {}^x\Psi_j^b \rangle \langle {}^x\Psi_j^b | S_\sigma | {}^x\Psi \rangle \end{aligned}$$

where $\sigma = x, y, z$. We shall now show how to compute each matrix element in the above equation. The equations for the spin components for a given ground state determinant with itself are given by:

$$\begin{aligned} \langle \Psi | S_z | \Psi \rangle &= \frac{1}{2} (\text{tr}(\mathbf{S}_{AO} \mathbf{P}_{\alpha\alpha}) - \text{tr}(\mathbf{S}_{AO} \mathbf{P}_{\beta\beta})) \\ \langle \Psi | S_x | \Psi \rangle &= \frac{1}{2} (\text{tr}(\mathbf{S}_{AO} \mathbf{P}_{\alpha\beta}) + \text{tr}(\mathbf{S}_{AO} \mathbf{P}_{\beta\alpha})) \\ \langle \Psi | S_y | \Psi \rangle &= \frac{1}{2} (\text{tr}(\mathbf{S}_{AO} \mathbf{P}_{\alpha\beta}) - \text{tr}(\mathbf{S}_{AO} \mathbf{P}_{\beta\alpha})) \end{aligned}$$

where $\mathbf{P}_{\alpha\beta}$ etc. are the blocks of the GHF density matrix and $\mathbf{S}_{AO}$ is the atomic orbital overlap. These are ob-

tained by applying the the Pauli spin matrices to our non-collinear (GHF) wavefunction.^{29,78} To compute $\langle {}^w\Psi | {}^x\Psi \rangle$ we use the fact that in the corresponding orbital basis the inter-determinantal overlap matrix ${}^w\mathbf{S}$ is diagonal, which gives us:

$$\langle {}^w\Psi | {}^x\Psi \rangle = \det({}^w\mathbf{S}) = \prod_i s_i$$

where s_i are the elements of that diagonal. To compute $\langle {}^w\Psi | S_\sigma | {}^w\Psi_i^a \rangle$ we recall that each individual Slater determinant maintains orthogonality between their respective occupied and virtual subspaces, therefore:

$$\begin{aligned} \langle \Psi | S_z | \Psi_i^a \rangle &= \sum_{\mu\nu} {}^\alpha C_{i\cdot}^{\cdot\mu} \langle \omega_\mu | \omega_\nu \rangle {}^\alpha C_{\cdot a}^{\nu\cdot} - {}^\beta C_{i\cdot}^{\cdot\mu} \langle \omega_\mu | \omega_\nu \rangle {}^\beta C_{\cdot a}^{\nu\cdot} \\ \langle \Psi | S_x | \Psi_i^a \rangle &= \sum_{\mu\nu} {}^\beta C_{i\cdot}^{\cdot\mu} \langle \omega_\mu | \omega_\nu \rangle {}^\alpha C_{\cdot a}^{\nu\cdot} + {}^\alpha C_{i\cdot}^{\cdot\mu} \langle \omega_\mu | \omega_\nu \rangle {}^\beta C_{\cdot a}^{\nu\cdot} \\ \langle \Psi | S_y | \Psi_i^a \rangle &= \sum_{\mu\nu} {}^\beta C_{i\cdot}^{\cdot\mu} \langle \omega_\mu | \omega_\nu \rangle {}^\alpha C_{\cdot a}^{\nu\cdot} - {}^\alpha C_{i\cdot}^{\cdot\mu} \langle \omega_\mu | \omega_\nu \rangle {}^\beta C_{\cdot a}^{\nu\cdot} \end{aligned} \tag{A1}$$

To compute $\langle {}^w\Psi_i^a | {}^x\Psi \rangle$ and $\langle {}^w\Psi | {}^x\Psi_j^b \rangle$ we recognize that in the corresponding orbital basis a single substitution in the coefficient matrix replaces either a row or column of ${}^{wx}\mathbf{S}$, changing the determinant of that matrix by one value:

$$\begin{aligned}\langle {}^w\Psi_i^a | {}^x\Psi \rangle &= \frac{\det({}^{wx}\mathbf{S}) \langle {}^w\phi_a | {}^x\phi_i \rangle}{{}^{wx}S_{ii}} \\ \langle {}^w\Psi | {}^x\Psi_j^b \rangle &= \frac{\det({}^{wx}\mathbf{S}) \langle {}^w\phi_j | {}^x\phi_b \rangle}{{}^{wx}S_{jj}}\end{aligned}$$

In order to compute the final element, $\langle {}^w\Psi_i^a | {}^x\Psi_j^b \rangle$, we will utilize the matrix determinant lemma which states:

$$\det(\mathbf{A} + \mathbf{U}\mathbf{V}^\dagger) = \det(\mathbf{I} + \mathbf{V}^\dagger\mathbf{A}^{-1}\mathbf{U}) \det(\mathbf{A}) \quad (\text{A2})$$

where $\mathbf{A}$ is an n -by- n invertible matrix and $\mathbf{U}$ and $\mathbf{V}$ are n -by- m matrices and are not related to the $\mathbf{U}$ and $\mathbf{V}$ of the SVD. The lemma becomes useful when we notice

that, akin to the element $\langle {}^w\Psi_i^a | {}^x\Psi \rangle$ computed above, the element $\langle {}^w\Psi_i^a | {}^x\Psi_j^b \rangle$ involves replacing a row and a column of ${}^{wx}\mathbf{S}$. In our case the matrix $\mathbf{A}$ will be the corresponding orbitals overlap matrix ${}^{wx}\mathbf{S}$, the determinant of which is $\langle {}^w\Psi | {}^x\Psi \rangle$ which we have already computed above. Recalling this matrix is diagonal the inverse may also be computed easily. The $\mathbf{U}$ and $\mathbf{V}$ must be constructed so that they replace the elements of the ${}^{wx}\mathbf{S}$ appropriately; in our case because we are replacing a row and a column of this matrix $\mathbf{U}$ and $\mathbf{V}$ will be n -by-2 matrices making the so far unknown determinant on the right hand side the determinant of 2-by-2 matrix. The construction of $\mathbf{U}$ and $\mathbf{V}$ involves recognizing that the substitutions in $|{}^w\Psi\rangle$ make for rows consisting of elements $\langle {}^w\phi_a | {}^x\phi_i \rangle$, substitutions in $|{}^x\Psi\rangle$ make for columns consisting of elements $\langle {}^w\phi_j | {}^x\phi_b \rangle$ and both substitutions make a point of intersection $\langle {}^w\phi_a | {}^x\phi_b \rangle$. Utilizing Kronecker deltas to take into account when to add and subtract the values, we can write down $\mathbf{U}$ and $\mathbf{V}$:

$$\begin{aligned}U_{1,k} &= \delta_{ki} \\ U_{2,k} &= (1 - \delta_{ki})(\langle {}^w\phi_k | {}^x\phi_b \rangle - \delta_{kj}s_j) \\ V_{l,1} &= (1 - \delta_{lj})(\langle {}^w\phi_a | {}^x\phi_l \rangle - \delta_{li}s_i) + \delta_{lj}(\langle {}^w\phi_a | {}^x\phi_b \rangle - \delta_{li}s_i) \\ V_{l,2} &= \delta_{lj}\end{aligned}$$

where i represents the occupied orbital substituted by a , the virtual orbital, of $|{}^w\Psi\rangle$ and j and b likewise for $|{}^x\Psi\rangle$,

and the indices k and l represent dummy occupied orbital indices (the n rows of these skinny matrices). After some algebra we arrive at:

$$\begin{aligned}
\det(\mathbf{I} + \mathbf{V}^{\dagger wx} \mathbf{S}^{-1} \mathbf{U}) &= \\
&= \det \begin{pmatrix} \clubsuit & \heartsuit \\ \diamond & \spadesuit \end{pmatrix} \\
\clubsuit &= \frac{\langle x\phi_i | w\phi_a \rangle}{s_i} + \delta_{ji} \left(\frac{\langle x\phi_b | w\phi_a \rangle}{s_i} - \frac{\langle x\phi_i | w\phi_a \rangle}{s_i} \right) \\
\heartsuit &= \sum_l \frac{\langle w\phi_a | x\phi_l \rangle \langle w\phi_l | x\phi_b \rangle}{s_l} - \frac{\langle w\phi_a | x\phi_i \rangle \langle w\phi_i | x\phi_b \rangle}{s_i} \\
&\quad + \frac{\langle w\phi_a | x\phi_b \rangle \langle w\phi_j | x\phi_b \rangle}{s_j} - \frac{\langle w\phi_a | x\phi_j \rangle \langle w\phi_j | x\phi_b \rangle}{s_j} - \langle w\phi_a | x\phi_b \rangle \\
&\quad + \delta_{ij} \left(\frac{\langle w\phi_a | x\phi_i \rangle \langle w\phi_i | x\phi_b \rangle}{s_i} - \frac{\langle w\phi_a | x\phi_b \rangle \langle w\phi_j | x\phi_b \rangle}{s_j} + \langle w\phi_a | x\phi_b \rangle \right) \\
\diamond &= \delta_{ij} s_i^{-1} \\
\spadesuit &= \frac{\langle w\phi_j | x\phi_b \rangle}{s_j} - \delta_{ji} \left(\frac{\langle w\phi_i | x\phi_b \rangle}{s_i} + 1 \right) \\
\det(\mathbf{I} + \mathbf{V}^{\dagger wx} \mathbf{S}^{-1} \mathbf{U}) &= \frac{\langle w\phi_a | x\phi_i \rangle \langle w\phi_j | x\phi_b \rangle}{s_i s_j} \\
&\quad + \delta_{ji} \left(\frac{\langle x\phi_b | w\phi_a \rangle}{s_i} - s_i^{-1} \sum_l \frac{\langle w\phi_a | x\phi_l \rangle \langle w\phi_l | x\phi_b \rangle}{s_l} \right)
\end{aligned}$$

This expression can then be split up into the on and off diagonal terms, combined with the equations of A1, and contracted appropriately to form an evaluation of $\langle S^2 \rangle$ which scales as N^3 per determinantal pair where N is the size of atomic orbital basis.

- ¹K. Palczewski, *Annu. Rev. Biochem.* **75**, 743 (2006).
- ²J. L. Delgado, P.-A. Bouit, S. Filippone, M. A. Herranz, and N. Martín, *Chem. Commun.* **46**, 4853 (2010).
- ³P. M. Zimmerman, F. Bell, D. Casanova, and M. Head-Gordon, *JACS* **133**, 19944 (2011).
- ⁴P. M. Zimmerman, C. B. Musgrave, and M. Head-Gordon, *Acc. Chem. Res.*, 1339 (2013).
- ⁵K. Burke, J. Werschnik, and E. K. U. Gross, *J. Chem. Phys.* **123**, 62206 (2005).
- ⁶E. Runge and E. K. U. Gross, *Phys. Rev. Lett.* **52**, 997 (1984).
- ⁷N. T. Maitra, F. Zhang, R. J. Cave, and K. Burke, *J. Chem. Phys.* **120**, 5932 (2004).
- ⁸P. Elliott, S. Goldson, C. Canahui, and N. T. Maitra, *Chem. Phys.* **391**, 110 (2011).
- ⁹R. J. Cave, F. Zhang, N. T. Maitra, and K. Burke, *Chem. Phys. Lett.* **389**, 39 (2004).
- ¹⁰G. Mazur and R. Wodarczyk, *J. Comput. Chem.* **30**, 811 (2009).
- ¹¹J. Cullen, M. Krykunov, and T. Ziegler, *Chem. Phys.* **391**, 11 (2011).
- ¹²A. I. Krylov, *Annu. Rev. Phys. Chem.* **59**, 433 (2008).
- ¹³M. Head-Gordon, R. R. J. Rico, M. Oumi, and T. T. J. Lee, *Chem. Phys. Lett.* **219**, 21 (1994).
- ¹⁴O. Christiansen, H. Koch, and P. Jørgensen, *Chem. Phys. Lett.* **243**, 409 (1995).
- ¹⁵S. Hirata, M. Nooijen, and R. J. Bartlett, *Chem. Phys. Lett.* **326**, 255 (2000).
- ¹⁶A. I. Krylov, *Chem. Phys. Lett.* **350**, 522 (2001).
- ¹⁷Y. Shao, M. Head-Gordon, and A. I. Krylov, *J. Chem. Phys.* **118**, 4807 (2003).

- ¹⁸S. V. Levchenko and A. I. Krylov, *J. Chem. Phys.* **120**, 175 (2004).
- ¹⁹J. S. Sears, C. D. Sherrill, and A. I. Krylov, *J. Chem. Phys.* **118**, 9084 (2003).
- ²⁰D. Casanova and M. Head-Gordon, *J. Chem. Phys.* **129**, 064104 (2008).
- ²¹D. Casanova and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **11**, 9779 (2009).
- ²²P. M. Zimmerman, F. Bell, M. Goldey, A. T. Bell, and M. Head-Gordon, *J. Chem. Phys.* **137**, 164110 (2012).
- ²³D. Casanova, *J. Chem. Phys.* **137**, 084105 (2012).
- ²⁴B. O. Roos, in *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry Part 2, Volume 69*, *Advances in Chemical Physics*, Vol. 69, edited by K. P. Lawley (John Wiley & Sons, Inc., Hoboken, NJ, USA, 1987) pp. 399–442.
- ²⁵K. Andersson, P.-A. Malmqvist, and B. O. Roos, *J. Chem. Phys.* **96**, 1218 (1992).
- ²⁶G. K.-L. Chan and S. Sharma, *Annu. Rev. Phys. Chem.* **62**, 465 (2011).
- ²⁷P. A. Malmqvist, *Int. J. Quantum Chem.* **30**, 479 (1986).
- ²⁸A. J. W. Thom and M. Head-Gordon, *J. Chem. Phys.* **131**, 124113 (2009).
- ²⁹P. Löwdin and I. Mayer, *Adv. Quantum Chem.* **24**, 79 (1992).
- ³⁰C. A. Jimenez-Hoyos, T. M. Henderson, and G. E. Scuseria, *J. Chem. Theory Comput.* **7**, 2667 (2011).
- ³¹A. Overhauser, *Phys. Rev.* **128**, 1437 (1962).
- ³²J. Cizek and J. Paldus, *J. Chem. Phys.* **407**, 3976 (1967).
- ³³J. Paldus and J. Cizek, *J. Chem. Phys.* **52**, 2919 (1970).
- ³⁴H. Fukutome, *Int. J. Quantum Chem.* **20**, 955 (1981).
- ³⁵J. Calais, *Adv. Quantum Chem.* **17**, 225 (1985).
- ³⁶C. A. Jiménez-Hoyos, T. M. Henderson, T. Tsuchimochi, and G. E. Scuseria, *J. Chem. Phys.* **136**, 164109 (2012).
- ³⁷G. E. Scuseria, C. A. Jimenez-Hoyos, T. M. Henderson, K. Samanta, and J. K. Ellis, *J. Chem. Phys.* **135**, 124108 (2011).
- ³⁸J. Ellis, R. Martin, and G. Scuseria, *J. Chem. Theory Comput.* **9**, 2857 (2013).

- ³⁹K. Samanta, C. A. Jiménez-Hoyos, and G. E. Scuseria, *J. Chem. Theory Comput.* **8**, 4944 (2012).
- ⁴⁰I. Mayer, *Adv. Quantum Chem.* **12**, 189 (1980).
- ⁴¹R. Rodríguez-Guzmán, K. W. Schmid, C. A. Jiménez-Hoyos, and G. E. Scuseria, *Phys. Rev. B* **85**, 245130 (2012).
- ⁴²M. Head-Gordon, P. E. Maslen, and C. A. White, *J. Chem. Phys.* **108**, 616 (1998).
- ⁴³A. T. Amos and G. G. Hall, *Proc. R. Soc. A* **263**, 483 (1961).
- ⁴⁴H. F. King, R. E. Stanton, H. Kim, R. E. Wyatt, and R. G. Parr, *J. Chem. Phys.* **47**, 1936 (1967).
- ⁴⁵P.-O. Lowdin, *J. Appl. Phys.* **33**, 251 (1962).
- ⁴⁶Y. G. Smeyers and G. Delgado-Barrio, *Int. J. Quantum Chem.* **8**, 733 (1974).
- ⁴⁷A. T. B. Gilbert, N. a. Besley, and P. M. W. Gill, *J. Phys. Chem. A* **112**, 13164 (2008).
- ⁴⁸P. Pulay, *Chem. Phys. Lett.* **73** (1980).
- ⁴⁹P. Pulay, *J. Comput. Chem.* **3**, 556 (1982).
- ⁵⁰A. Thom and M. Head-Gordon, *Phys. Rev. Lett.* **101**, 1 (2008).
- ⁵¹N. A. Besley, A. T. B. Gilbert, and P. M. W. Gill, *J. Chem. Phys.* **130**, 124308 (2009).
- ⁵²B. Peng, B. E. Van Kuiken, F. Ding, and X. Li, *J. Chem. Theory Comput.*, 130829161949007 (2013).
- ⁵³S. R. Yost, T. Kowalczyk, and T. Van Voorhis, *J. Chem. Phys.* **139**, 174104 (2013).
- ⁵⁴Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. a. DiStasio, R. C. Lochan, T. Wang, G. J. O. Beran, N. a. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. a. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. a. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. I. Schaefer, J. Kong, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **8**, 3172 (2006).
- ⁵⁵C. Sanderson, Technical Report **NICTA**, 1 (2010).
- ⁵⁶P. C. Hariharan and J. Pople, *Theor. Chim. Acta* **28**, 213 (1973).
- ⁵⁷A. D. Becke, *J. Chem. Phys.* **98**, 6 (1993).
- ⁵⁸Y. Kurashige, H. Nakano, Y. Nakao, and K. Hirao, *Chem. Phys. Lett.* **400**, 425 (2004).
- ⁵⁹C.-P. Hsu, S. Hirata, and M. Head-Gordon, *J. Phys. Chem. A* **105**, 451 (2001).
- ⁶⁰J. H. Starcke, M. Wormit, J. Schirmer, and A. Dreuw, *Chem. Phys.* **329**, 39 (2006).
- ⁶¹J. D. Watts, S. R. Gwaltney, and R. J. Bartlett, *J. Chem. Phys.* **105**, 6979 (1996).
- ⁶²M. A. Watson and G. K.-L. Chan, *J. Chem. Theory Comput.* **8**, 4013 (2012).
- ⁶³O. Lehtonen, D. Sundholm, R. Send, and M. P. Johansson, *J. Chem. Phys.* **131**, 024301 (2009).
- ⁶⁴O. Kitao and H. Nakatsuji, *Chem. Phys. Lett.* **143**, 528 (1988).
- ⁶⁵B. Saha, M. Ehara, and H. Nakatsuji, *J. Chem. Phys.* **125**, 014316 (2006).
- ⁶⁶L. Serrano-Andres, R. Lindh, B. O. Roos, and M. Merchán, *J. Phys. Chem.* **97**, 9360 (1993).
- ⁶⁷K. Nakayama, H. Nakano, and K. Hirao, *Int. J. Quantum Chem.* **66**, 157 (1998).
- ⁶⁸B. Ostojić and W. Domcke, *Chem. Phys.* **269**, 1 (2001).
- ⁶⁹M. Boggio-Pasqua, M. J. Bearpark, M. Klene, and M. A. Robb, *J. Chem. Phys.* **120**, 7849 (2004).
- ⁷⁰J. P. Doering and R. McDiarmid, *J. Chem. Phys.* **73**, 3617 (1980).
- ⁷¹L. Serrano-Andres, M. Merchán, I. Nebot-Gil, R. Lindh, and B. O. Roos, *J. Chem. Phys.* **98**, 3151 (1993).
- ⁷²T. Polívka and V. Sundström, *Chem. Phys. Lett.* **477**, 1 (2009).
- ⁷³D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.* **128**, 144117 (2008).
- ⁷⁴S. Grimme and M. Waletzke, *J. Chem. Phys.* **111** (1999).
- ⁷⁵M. Kleinschmidt, C. M. Marian, M. Waletzke, and S. Grimme, *J. Chem. Phys.* **130**, 044708 (2009).
- ⁷⁶N. Tschirner, M. Schenderlein, K. Brose, E. Schlodder, M. A. Mrogiński, C. Thomsen, and P. Hildebrandt, *Phys. Chem. Chem. Phys.* **11**, 11471 (2009).
- ⁷⁷H. Fukutome, *Prog. Theor. Phys.* **80**, 417 (1988).
- ⁷⁸B. Sykja and J. L. Calais, *J. Phys. C: Solid State Phys.* **15**, 3079 (1982).