JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2019, 15, 4790−4803

pubs.acs.org/JCTC

Tracking Excited States in Wave Function Optimization Using Density Matrices and Variational Principles

Lan Nguyen Tran,*[,†,‡] Jacqueline A. R. Shea,[†] and Eric Neuscamman*[,†,§]

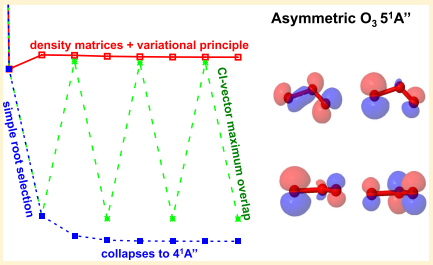
[†]Department of Chemistry, University of California, Berkeley, California 94720, United States

[‡]Ho Chi Minh City Institute of Physics, VAST, Ho Chi Minh City 700000, Vietnam

[§]Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

Supporting Information

ABSTRACT: We present a method for finding individual excited states' energy stationary points in complete active space self-consistent field theory that is compatible with standard optimization methods and highly effective at overcoming difficulties due to root flipping and near-degeneracies. Inspired by both the maximum overlap method and recent progress in excited-state variational principles, our approach combines these ideas in order to track individual excited states throughout the orbital optimization process. In a series of tests involving root flipping, near-degeneracies, charge transfers, and double excitations, we show that this approach is more effective for state-specific optimization than either the naive selection of roots on the basis of energy ordering or a more direct generalization of the maximum overlap method. We provide evidence that this state-specific approach improves the performance of complete active space perturbation theory for vertical excitation energies. Furthermore, we find that the state-specific optimization can help avoid state-averaging-induced discontinuities on potential energy surfaces. With a simple implementation, a low cost, and compatibility with large active space methods, the approach is designed to be useful in a wide range of excited-state investigations.



1. INTRODUCTION

While linear response theory has indisputably been the most widely successful paradigm for making predictions about electronic excitations in chemistry, there remain important situations in which its use is either difficult or entirely inadvisible. The most affordable methods in this category,[1] such as time-dependent density functional theory (TD-DFT) and configuration interaction singles (CIS), are well-known to face challenges when an excitation significantly alters a molecule's charge density, as occurs in charge transfer, Rydberg, and core excitations. Although TD-DFT has additional concerns,[2] one major issue in these cases comes from the fact that significant charge density changes induce orbital relaxation effects that are not captured in the linear response of a Slater determinant.[3] Even equation of motion coupled cluster theory with singles and doubles excitations (EOM-CCSD),[4] which responds around a much more sophisticated wave function, has difficulty in capturing these relaxation effects in double excitations[4−6] and some core excitations.[7] Orbital relaxation effects aside, these linear response methods—as well as many other excited-state methods including GW theory,[8] the Bethe−Salpeter equation,[8] and most variants of the algebraic diagrammatic construction[9]—rely on the assumption that the ground state is single-reference in character. When this assumption fails, as for example at conical intersections where at least two states mix strongly, one is typically forced to abandon the realm of weakly correlated methods altogether.

For decades now, complete active space self-consistent field theory (CASSCF)[10−13] and its generalizations[14−16] have been the go-to starting points for strongly correlated treatments of molecular ground and excited states alike. Just as Hartree−Fock (HF) theory[17] provides a reference state in which the molecular orbitals are relaxed in the presence of strong Pauli correlations, CASSCF relaxes the molecular orbitals in the presence of both Pauli correlations and any other strong correlations that exist within an active set of orbitals and electrons, typically chosen as a small set near the Fermi level. While expanding the size of the active space that can be treated remains a high priority and has been the focus of much recent work,[18−24] an additional important concern in excited-state modeling is the issue of root flipping.[25−29] In the same way that HF and ΔSCF methods[30−33] relax orbitals by finding stationary points on the single-Slater-determinant energy surface, one would ideally like to prepare a CASSCF reference state by finding the CASSCF energy stationary point corresponding to the excited state in question. However, the common approach of separating the configuration interaction (CI) coefficient equations and the orbital rotation equations can create problems when two CI stationary points (typically called roots) cross each other in energy when the orbitals are updated in pursuit of the (initially) higher-energy root's overall stationary point. Simple root selection (SRS)—in which one

Received: April 10, 2019
Published: August 8, 2019

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

naively seeks to make the *n*th root's energy stationary during each orbital rotation step—will fail in this scenario because the *n*th root and the root below it exchange ordering indefinitely during the "two-step" cycle of CI and orbital relaxation steps. While damping schemes aid convergence in some cases, they do not address the underlying issue of a lower state's root being much too high in energy when expressed in an active space built from a different state's orbitals.

Although early methods in the coupled nonlinear optimization of CI coefficients and orbitals showed promise in bypassing this issue,[34] the most common remedy in use today is the state-averaged (SA) approach,[35] in which one simply minimizes the average energy of multiple states. The SA approach has the benefit of being compatible with modern, highly efficient two-step optimization schemes and is largely immune to root flipping by virtue of being insensitive to the ordering of the states within the average. Although this resolution of the root-flipping problem is very important, it is not the only reason that SA-CASSCF is often preferred. It has been shown, for example, that the SA approach can help facilitate error cancellation within fixed-node diffusion Monte Carlo,[36] which is not a trivial advantage given the near-inscrutability of nodal surfaces. More generally, SA-CASSCF ensures that different states are orthogonal and represented in a common orbital basis, which greatly simplifies the evaluation of important state-interaction properties such as spin−orbit and nonadiabatic couplings.[37,38] The orthogonality of SA-CASSCF states also supports "perturb-then-diagonalize" schemes, such as (extended) multistate complete active space second-order perturbation theory [(X)MS-CASPT2], which offer substantial benefits in the vicinity of the conical intersections and other difficult regions of molecular potenital energy surfaces (PESs).[39−41]

While these advantages make SA-CASSCF a powerful tool for modeling molecular excited states, the requirement that all states share a common orbital basis prevents them from being energy stationary points, which any exact Hamiltonian eigenstate should be.

This lack of stationarity can be inconvenient when evaluating gradients,[42] but the more troubling issue is the potential loss of accuracy incurred by producing reference states whose orbitals are not fully relaxed. For instance, it has been shown that SA-CASSCF orbitals are often biased toward the ground state and thus not reliable in some charge-transfer situations.[43,44] Another issue that is exacerbated by SA is the unfortunate reality that, like those of essentially any other nonlinear wave function approximation, CASSCF's PESs are not always smooth. The more states that are included in a state average, the higher are the chances that one of them will find a nonanalytic point, and because all of the states are linked through the shared orbitals, all of the states in the average will develop a nonanalyic feature at that point. This issue is particularly common when using fixed-weight SA-CASSCF and poses major challenges for nonadiabatic dynamics simulations, where dynamic weighting schemes have provided some relief.[45−47] Whether dealing with a case in which different states require drastically different orbital relaxations due to large charge-transfer-induced dipole shifts[48] or a case in which state averaging introduces PES discontinuities, it is interesting to ask how helpful it would be to instead fully relax individual states to their associated energy stationary points and, crucially, whether this can be achieved in the face of root flipping and while retaining compatibility with efficient two-step optimization approaches.

Recently, fully state-specific orbital relaxations have been achieved for weakly correlated excited states[49] by exploiting simple approximations to excited-state variational principles.[50−53] This progress came in the form of a nonlinear optimization in which the approximated variational principle was minimized under the constraint that the energy end up stationary with respect to both CI coefficients and orbital rotations. While one can imagine a straightforward application of these ideas to the CASSCF wave function, this would lead to a coupled nonlinear optimization of CI and orbital variables that would not necessarily be cost-competitive with modern two-step optimization schemes. In this study, we instead explore whether affordable approximations to excited-state variational principles can aid CASSCF in successfully tracking individual excited states in the presence of root flipping. The general idea is similar in spirit to the maximum overlap method (MOM),[54,55] which works to prevent variational collapse during ΔSCF optimizations, and indeed we find that state tracking is most effective when we combine approximate variational principles with a MOM-inspired matching condition based on reduced density matrices. In particular, we demonstrate that this combination is more effective at finding CASSCF excited-state stationary points than either the naive SRS approach or an approach based on approximate wave function overlaps as estimated via the CI vectors. In doing so, we also provide evidence that although the improvement is sometimes modest, state-specific CASSCF (SS-CASSCF) is a better reference for post-CASSCF methods than SA-CASSCF. Furthermore, initial testing on MgO shows that SS-CASSCF is able to provide smooth PESs (which are interestingly quasi-diabatic in character) in a case where equal-weight SA-CASSCF cannot, which suggests that it could act as a complement to dynamic weighting schemes in such cases.[45] Finally, as the CASSCF step is rarely the bottleneck when post-CASSCF methods are in use, our approach does not significantly increase the overall computational cost, so we are able to recommend its use in general. In particular, the approach requires only one- and two-body reduced density matrices and thus should be immediately compatible with the rapidly expanding collection of large active space methods that have come onto the scene in recent years.[18−23]

2. THEORY

2.1. Excited-State Variational Principles. It has long been recognized[52,53,56] that the variational principle

$$W = \frac{\langle \Psi | (\omega - \hat{H})^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{1}$$

has the Hamiltonian eigenstate with energy closest to ω as its global minimum. Recently, this and similar forms involving $\hat{H}^2$ have found tractable applications in spite of the challenges posed by the squared Hamiltonian operator, both in single-reference excited-state approaches[49,57] and in quantum Monte Carlo,[48,58−64] where one can separate the two powers of $\hat{H}$ using a resolution of the identity $\sum_I |I\rangle\langle I|$ over which statistical sampling may be performed:

$$W = \frac{1}{\langle \Psi | \Psi \rangle} \sum_I \langle \Psi | (\omega - \hat{H}) | I \rangle \langle I | (\omega - \hat{H}) | \Psi \rangle \tag{2}$$

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

In the context of CASSCF, one can simplify this resolution of the identity significantly by exploiting the fact that the space of wave functions to which $\hat{H}$ can connect the CASSCF reference is exactly spanned by the active space itself and all internally contracted singles and doubles excitations out of it. For ease of discussion, we will consider the case in which there are no closed-shell orbitals, noting that the generalization to the full theory with a closed shell is straightforward and has been realized in our implementation. With active spin orbitals labeled by i and j and virtual spin orbitals labeled by a and b and the trivial assumption that the CASSCF wave function Φ is normalized, the resolution of the identity form of the variational principle can thus be organized as

$$W = W_0 + W_1 + W_2 \tag{3}$$

in which

$$W_0 = (\omega - E)^2 \tag{4}$$

$$W_1 = \sum_{ia} |\langle \Phi | \hat{H} \hat{a}_a^\dagger \hat{a}_i | \Phi \rangle|^2 \tag{5}$$

$$W_2 = \sum_{ijab} \left| \langle \Phi | \hat{H} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i | \Phi \rangle \right|^2 \tag{6}$$

Of particular note is the fact that by virtue of being limited to internally contracted single and double excitations, a full evaluation of W, should one wish to pursue it, should be similar in complexity to constructing the right-hand side of the first-order wave function equation in CASPT2.[65,66] In particular, if we denote the numbers of active and virtual orbitals by n_a and n_v, the formal scalings of W_1 and W_2 are $n_a^3 n_v$ and $n_a^4 n_v^2$, respectively.

Inspecting the three components of W, we find that they have simple interpretations if we take Φ to be a root of the complete active space CI problem. First, W_0 simply states that the eigenstate we are seeking should have an energy close to ω. Second, noting that the energies of CI roots are already stationary with respect to the CI coefficients, we see that W_1 is simply a measure of how close the wave function is to being an overall energy stationary point, as each of its terms is proportional to the energy derivative with respect to an active-to-virtual orbital rotation. Indeed, if Φ is a CI root and $W_1 = 0$, then the CASSCF energy is stationary. Finally, noting that W_1 and W_2 are unaffected if we make the replacement $\hat{H} \rightarrow \hat{H} - E$, we see that W_2 contains all of the terms that would need to be zero in addition to those in W_1 in order for the energy variance $\sigma^2 = \langle (\hat{H} - E)^2 \rangle$ to be zero. As an exact eigenstate of $\hat{H}$ will be both an energy stationary point and a zero-variance state, we see that together W_0, W_1, and W_2 are simply a least-squares way of saying that we want the CASSCF wave function that is closest to the exact energy eigenstate near ω.

In principle, we could use these expressions to follow the approach of excited-state mean field (ESMF) theory[49] and minimize the Lagrangian

$$L = W - \boldsymbol{\mu} \cdot \frac{\partial E}{\partial \boldsymbol{\nu}} \tag{7}$$

with respect to a variable set $\boldsymbol{\nu}$ containing both orbital rotations and the CI coefficients. This approach uses the variational principle W, or an approximation to it, to guide an optimization to the desired energy stationary point via constrained Lagrangian minimization. However, making such an approach cost-competitive with CASSCF two-step optimization methods would not be trivial, and in this study we seek to exploit approximations to W in a much simpler context.

Within the standard two-step approach of switching back and forth between Davidson CI diagonalizations and orbital optimization steps, it is typically the CI step that dominates the cost when the active space gets large. We will therefore leave the Davidson step unchanged for now and consider using approximations to W in the remainder of the optimization where, unlike in the ground-state case, we will need an objective function that is not simply the total energy in order to avoid collapses to lower states. In our initial testing, we have found that while a limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) minimization of $|\nabla L|^2$ with respect to orbital rotation variables is effective when we set $W \approx W_0 + W_1$, it is in most cases equally effective to relax the orbitals by a simple L-BFGS minimization of $|\nabla E|^2$. This observation suggests that simple generalizations of standard Newton–Raphson-style orbital optimizations in which the line search is set to minimize $|\nabla E|^2$ rather than E itself are likely to be adequate in many cases (and much faster than L-based quasi-Newton methods), even if they lack the strong resistance to ground-state collapse offered by L. However, orbital relaxation for a particular CI root is only part of the process of converging to a desired stationary point during a CASSCF two-step optimization. The method used to select the CI root for which the orbitals are to be relaxed is equally important and in general quite challenging. As we now discuss, it is in this area that we find excited-state variational ideas to be most helpful.

2.2. Maximum Overlap Analogues. First introduced by Gill and co-workers,[54] the maximum overlap method helps prevent variational collapse to the ground state in the attempt to locate excited-state solutions to the Roothan equations in Hartree–Fock or density functional theory. The idea is to choose the excited-state orbital occupation for the orbitals generated by a newly diagonalized Fock matrix by selecting the orbitals that give the largest combined overlap with the molecular orbitals from the previous iteration of the method or some target set of molecular orbitals believed to be similar to those of the desired excited state. In essence, MOM is an attempt to follow the trail of a particular excited state through the sequence of discrete (and sometimes large) orbital relaxations that occur over the course of the self-consistent field iterations. This goal is very similar to what we desire when faced with a root-flipping problem in CASSCF: we wish to track a particular excited state through the sequence of discrete (and sometimes large) changes to the Davidson CI roots that occur over the course of a two-step CASSCF optimization.

If one wished to pursue a strategy similar in spirit to that of the orbital-overlap-based MOM, a simple strategy would be to hope that changes to the CI vector $\mathbf{c}$ between iterations are never too large and simply define a tracking function Q_{MOM} based on the approximate wave function overlap between each of the current iteration's CI roots and some target CI vector $\mathbf{c}_t$, taken here to be the CI vector selected in the previous iteration:

$$Q_{\mathrm{MOM}}(\mathbf{c}) = \mathbf{c}_t \cdot \mathbf{c} \tag{8}$$

We should stress that to keep the evaluation inexpensive, this Q_{MOM} metric ignores changes in the orbital basis from one iteration to the next and thus is not an exact evaluation of the overlap between the current and previous wave functions. Of

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790–4803

course, like MOM itself, this strategy, and any tracking strategy based on measuring wave function similarities across iterations, will not necessarily be robust in cases where the iterative method makes large changes to the wave function in a single iteration. Ideally, we would therefore like to augment this strategy with a component that is less dependent on iteration-to-iteration similarity. The central finding of this study is that in combination with a measure of similarity that is more robust than CI vector dot products, the excited-state variational principle W can help in this way.

In hopes of creating a more robust state-tracking function while keeping the cost low, we define the following quality measure for a newly generated Davidson root with CI vector $\mathbf{c}$:

$$Q_{W\Gamma}(\mathbf{c}) = W_0(\mathbf{c}) + W_1(\mathbf{c}) + D(\mathbf{c}) \tag{9}$$

in which

$$D(\mathbf{c}) = \frac{\|\Gamma_t - \Gamma(\mathbf{c})\|}{n_{CAS}} \tag{10}$$

where Γ_t and $\Gamma(\mathbf{c})$ are the one-body reduced density matrices (1RDMs) for the target wave function and the current root in question, respectively, and n_{CAS} is the number of active orbitals. Γ_t can be either fixed at the beginning of the calculation or updated on the basis of the selected root at each iteration (in the present work, we keep it fixed). In either case, we rotate it into the current orbital basis before evaluating D in order to reduce sensitivity to orbital changes. One could achieve similar insensitivity in Q_{MOM} by evaluating the exact overlap between the previous and current iterations' roots, but it is not obvious whether this could be achieved at low cost. It should be noted that we divide the Frobenius norm of the 1RDM difference by the number of active orbitals in order to make the relative importance of W_0, W_1, and D less sensitive to the size of the chosen active space. This choice is built on the idea that the out-of-active-space parts of the density matrix difference are small, as the occupation vector is fixed for these orbitals and we do not expect qualitative changes in the closed-shell orbital shapes. While it may be that a different balance between W_0, W_1, and the density matrix difference is optimal, in practice we have found that this choice is effective in most cases. While one could of course also include W_2 in this quality measure, we have chosen to omit it because of its relatively high cost of evaluation and our observation that $Q_{W\Gamma}$ is quite effective even without it.

As for why we chose a density matrix difference for our measure of iteration-to-iteration similarity, the logic is that although we could measure the similarity via a combination of any number of wave function properties (e.g., dipole moment), a large number of properties are themselves determined via the 1RDM. Of course, one could also consider 2RDM differences, but for simplicity's sake we have for now limited our investigation to differences of 1RDMs. As our results will demonstrate, the quality measure $Q_{W\Gamma}$, although not perfect and certainly less sophisticated than it could be, is far more effective at dealing with root flipping in attempts to track a specific excited state through a CASSCF two-step optimization than either SRS or Q_{MOM}. Finally, as shown in section 3.4, we emphasize that despite the intuitive role of D, the combined $W\Gamma$ quality measure is essential in cases where two roots have similarly structured density matrices.

2.3. Optimization Procedure. To summarize, the overall optimization procedure that we test here involves the following steps:

1. An orbital basis is chosen as an initial guess, an initial CASCI calculation is performed, and the root that will be targeted for state-specific convergence is selected. ω is initialized as this root's energy and Γ_t as its one-body density matrix.
2. The orbital coefficient matrix $\mathbf{C} = \mathbf{C}_0 e^{\mathbf{X}}$ is relaxed via L-BFGS minimization of either $|\nabla_{\mathbf{X}} L|^2 + |\nabla_{\boldsymbol{\mu}} L|^2$ (with $W \approx W_0 + W_1$) or the even simpler objective function $|\nabla_{\mathbf{X}} E|^2$. While the latter is in principle more prone to variational collapse, we have found that in the cases studied here, these two orbital optimization approaches lead to the same results. It should be noted that for convenience we have implemented the L-based approach using TensorFlow's[67] automatic differentiation[49,68] framework. For the approach of minimizing $|\nabla_{\mathbf{X}} E|^2$ alone, we implemented a GMRES solver on top of pySCF's[69] existing CASSCF code in order to step along the Newton–Raphson update direction with a line search that is set to minimize $|\nabla_{\mathbf{X}} E|^2$. With the total number of orbitals $n_o = n_a + n_v$ (again, core orbitals can be included, but we suppress them for simplicity of presentation), the scaling of this step (dominated in both approaches by the four-index integral transformation) is $n_a n_o^4$.
3. A new CASCI calculation is performed (via pySCF) to obtain the low-lying roots of the appropriate space and spin symmetry in the new orbital basis. The scaling of this step is combinatorial in the size of the active space.
4. The root with the lowest value of $Q_{W\Gamma}$ or the largest value for Q_{MOM} is selected, depending on which quality measure is being employed. In the $Q_{W\Gamma}$ approach, ω is then reset to the selected root's energy. Γ_t is not updated, but it is rotated into the current orbital basis before the density matrix difference is evaluated. If instead one is following the SRS approach, then the root is simply selected on the basis of its position in the energy ordering. With a CASCI expansion of n_{det} determinants, the cost of this step scales as n_{det} for Q_{MOM} and as $n_a^3 n_v$ (dominated by the W_1 term) for $Q_{W\Gamma}$.
5. The procedure is repeated starting with step 2 until an overall CASSCF energy stationary point is found.

3. RESULTS AND DISCUSSION

We now turn to a collection of molecular examples with which we seek to gain insight into four key questions: First, how effective is the $W\Gamma$ approach at overcoming root flipping in comparison with MOM and SRS? Second, can the $W\Gamma$ approach succeed in cases where there are nearly degenerate states that cannot be distinguished by W alone? Third, do the SS-CASSCF solutions that this approach helps us find outperform their SA-CASSCF counterparts as reference functions for post-CASSCF methods like CASPT2 and Davidson-corrected multireference configuration interaction with singles and doubles (MRCI+Q)? Finally, can SS-CASSCF yield smooth potential energy surfaces when SA-CASSCF does not?

These questions are studied in the molecular systems summarized in Table 1. The cc-pVDZ basis set was used throughout, as were CASSCF energy, orbital gradient, and density matrix convergence thresholds of 10^{-7}, 10^{-4}, and 10^{-4},

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

Table 1. Summary of Frozen and Active Orbitals Used for All Systems

molecule	frozen orbitals	active orbitals
LiH	none	4e,4o: Li 1s2s2p_z H 1s
O_3	O 1s2s	12e,9o: O 2p
CH_2O	C 1s2s O 1s2s	8e,8o: C 2p O 2p H 1s
MgO	none	8e,8o: Mg 3s3p O 2s2p

respectively. In most cases, we began our optimizations in the HF orbital basis, but in MgO we began in the local density approximation (LDA) basis instead because both the MOM and $W\Gamma$ approaches converged more rapidly from an LDA starting point. All of the CASPT2 and MRCI+Q calculations were carried out using Molpro,[70] while EOM-CCSD calculations were performed using QChem.[71] All of the post-CASSCF methods employed the usual frozen-core approximation, but for CASSCF itself the core was frozen or not as described in Table 1. For MRCI+Q, we used Molpro's default convergence thresholds for the energy and density matrix in all systems except for MgO, where we found it necessary to set them both to 10^{-5} in order to avoid unstable oscillations in the energies.

3.1. LiH. Let us begin with the well-known example of root flipping that occurs in LiH. Near the equilibrium bond length (~1.6 Å), this molecule's ground state ($X^1\Sigma^+$) is basically ionic, while the first excited state ($A^1\Sigma^+$) is predominantly neutral as a result of charge transfer excitation. As the bond length is increased, however, the ground state becomes increasingly neutral and the excited state increasingly ionic. At intermediate distances, the system shows an avoided crossing between these states that causes a well-recognized example of root flipping for SRS, as shown in the left panel of Figure 1. As the orbitals are optimized for the excited state, the

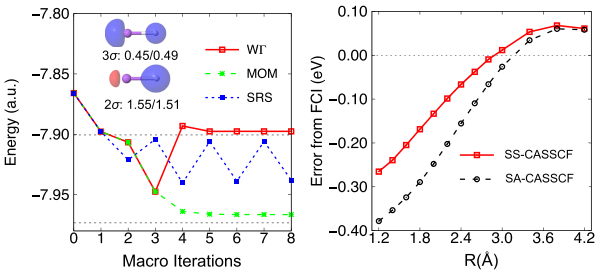


Figure 1. (left) Change in energy during CASSCF optimizations of the first excited state ($A^1\Sigma^+$) of LiH at a bond distance of 2.6 Å. The SRS, MOM, and $W\Gamma$ approaches all start with the Hartree−Fock orbitals as the initial guess. For reference, the horizontal lines show the full CI (FCI) energies for the ground and excited states. The inset shows the 2σ and 3σ natural orbitals and occupation numbers for the excited state as produced by the $W\Gamma$ approach. The occupation numbers from FCI (the second numbers) are given for comparison. Li and H atoms are shown in purple and white, respectively. (right) Excitation energy errors relative to FCI at different bond distances R.

energy of the ground-state CASCI root rises while that of the excited-state root falls. Soon the two roots flip in the energy ordering, at which point SRS is now effectively trying to optimize the orbitals for the ground state, which in turn causes the energy ordering to flip back. This process continues indefinitely, preventing the SRS approach from converging at all.

After the first orbital relaxation, both the MOM and $W\Gamma$ approaches recognize that the ordering of the roots has changed and select the lower root, thus diverging from the optimization path of SRS. Their two quality measures then select the same root for one more iteration, at which point orbital relaxations clearly work to push the state toward the ground state, as can be seen by the large energy lowering between macro iterations 2 and 3. It is here that the two methods diverge, with MOM's quality measure selecting the root that ultimately becomes the ground state. The $W\Gamma$ measure, on the other hand, successfully keeps track of the excited-state root and ultimately converges to a CASSCF stationary point that clearly corresponds to the desired excited state. As shown in the right panel of Figure 1 and in more detail in Table S1 in the Supporting Information, this success is repeated at all bond distances, and we find that the excitation energies based on SS-CASSCF energy differences tend to be a bit more accurate than those based on equally weighted SA-CASSCF energies.

3.2. Asymmetric O_3. We now turn to asymmetrically stretched ozone and focus our attention on two nearly degenerate excitations at the somewhat arbitrary asymmetric geometry $R_{O_1O_2}$ = 1.3 Å, $R_{O_2O_3}$ = 1.8 Å, $\angle O_1O_2O_3$ = 120°. These states correspond to the fourth and fifth SS-CASSCF excitation energies (relative to the $1^1A'$ ground state) in this C_s geometry's A″ representation, so we label them as $4^1A''$ and $5^1A''$ even though we will see that when out-of-active-space correlation is included it is no longer obvious which actually has the lower energy. The $4^1A''$ state's primary component is the configuration with four singly occupied orbitals resulting from a double excitation that moves one electron each from the two highest-occupied A′ orbitals into the lowest unoccupied A′ and A″ orbitals, whereas the $5^1A''$ state's primary component is the single excitation from the highest-occupied A″ orbital into the lowest unoccupied A′ orbital. That said, both of these states have nontrivial coefficients on the other state's primary component, meaning that neither state can be described as a simple single excitation. To add to the confusion, these states show up in the reverse order ($5^1A''$ before $4^1A''$) in the energy-ordered CASCI roots in the initial HF orbital basis, and so root flipping is essentially guaranteed during state-specific optimization.

Unlike the lowest three A″ states—all of which converge without trouble using SRS—this pair of states proves a nontrivial challenge for state-specific optimization. As shown in Figure 2, MOM finds the $4^1A''$ stationary point without difficulty when we start from the corresponding (fifth) A″ root of the HF-orbital CASCI but becomes trapped in a limit cycle when we target $5^1A''$ by starting from the corresponding (fourth) root. SRS succeeds in finding both stationary points, but it does so by making a root-flipping-induced qualitative swap in the state being tracked. This (or perhaps collapse to a lower state) is to be expected, as the energy ordering of these states is reversed at their individual stationary points compared with the initial HF-orbital CASCI. The practical results are twofold. First, the state the user gets is not the same excited state as was initially targeted. Second, the convergence of the $4^1A''$ state proves to be very slow, with the gradient still not quite converged even after 100 macro iterations. The $W\Gamma$ approach, in contrast, converges rapidly for both the $4^1A''$ and $5^1A''$ states when it is initiated with the corresponding root from the initial HF-orbital CASCI. Unlike SRS, these

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

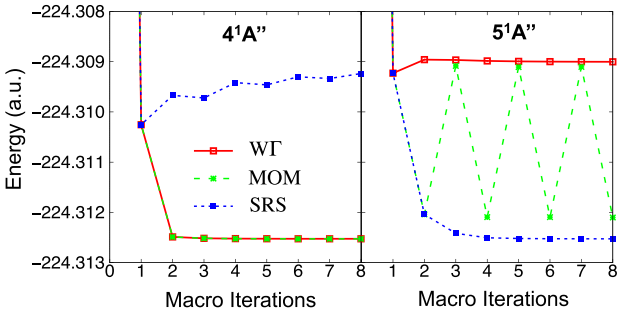


Figure 2. Energy of the selected root after each CASCI calculation during SS-CASSCF optimizations of asymmetric O_3. The left and right panels show results when the SRS, MOM, and *W*Γ approaches are initialized with the $4^1A''$ and $5^1A''$ CASCI roots, respectively, in the initial Hartree−Fock orbital basis.

optimizations do not involve any switching between qualitatively different states, despite the fact that the energy ordering does change. We explicitly verify that the desired state is tracked at every iteration of the optimization in Figure 3,

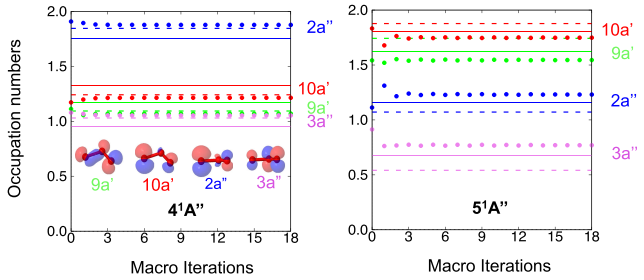


Figure 3. Natural orbital occupation numbers after each CASCI calculation during the *W*Γ-guided CASSCF optimizations of the (left) $4^1A''$ and (right) $5^1A''$ states in asymmetric O_3, with the zeroth iteration corresponding to the initial HF-orbital CASCI. For comparison, the solid and dashed horizontal lines give the corresponding final occupation numbers from three- and six-state SA-CASSCF calculations, respectively. The inset shows the final shapes of these four natural orbitals for the SS-CASSCF $4^1A''$ case. It should be noted that the shapes of the SS-CASSCF $5^1A''$ natural orbitals look very similar even though the natural occupation numbers largely differ.

which shows both how little the relevant active space natural orbital occupation numbers change and how clearly qualitatively different the $4^1A''$ and $5^1A''$ states are. At convergence, we see that the stationary points found in the *W*Γ approach correspond closely to both the initial CASCI wave functions and the results from equal-weighted three-state ($1^1A'$, $4^1A''$, $5^1A''$) and six-state ($1^1A'$, $1^1A''$, $2^1A''$, $3^1A''$, $4^1A''$, $5^1A''$) SA-CASSCF calculations.

Of course, in comparison to SA-CASSCF, the more pressing question is whether the ability to find the correct stationary points in SS-CASSCF offers any advantages in final accuracy. To investigate this question, we performed both CASPT2 and MRCI+Q calculations based on the different SS and SA wave functions, the results of which are displayed in Figure 4. We find that MRCI+Q predicts the states to be nearly degenerate, with excitation energies within 0.02 eV of each other, regardless of whether the orbitals from either SA-CASSCF calculation or from SS-CASSCF are used. It should be noted that the SS-MRCI+Q case amounts to three different MRCI

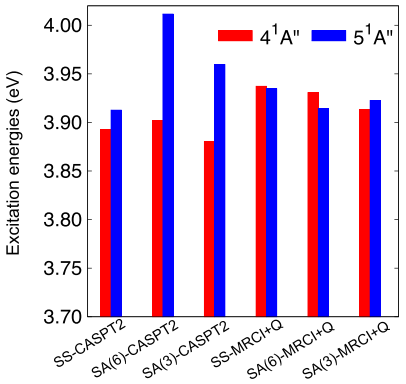


Figure 4. CASPT2 (with no level shift) and MRCI+Q excitation energies for the $4^1A''$ and $5^1A''$ states of asymmetric O_3 starting from SS-CASSCF or three- or six-state SA-CASSCF wave functions.

+Q calculations, one each in the ground, $4^1A''$, or $5^1A''$ state's orbital basis, with the appropriate MRCI+Q root's energy extracted in each case. On the other hand, CASPT2 predicts the degeneracy to be within 0.02 eV only when the calculation is based on the SS-CASSCF wave functions, showing gaps of more than 0.07 eV in both the three- and six-state SA cases (it should be noted that single-state zeroth-order Hamiltonians were used for CASPT2 starting from either SA-CASSCF or SS-CASSCF). One way to explain these findings is to note that the out-of-active-space single excitations that provide state-specific orbital relaxations are treated perturbatively in CASPT2, so the fact that SS-CASSCF has already provided state-specific orbitals puts us in a regime where the perturbative assumption of small singles (and doubles) coefficients is more likely to be satisfied in practice. Although the improvement is modest, SS-CASPT2 is in closer agreement with MRCI+Q than is SA-CASPT2, which is encouraging because it suggests that SS-CASSCF may help the lower-cost CASPT2 method agree better with the higher-cost and typically higher-accuracy MRCI+Q method. As we will see, this same pattern plays out again and again in the states studied below.

3.3. CH_2O. We now turn to formaldehyde—with C_{2v} geometry R_{CO} = 1.2 Å, R_{CH} = 1.1 Å, ∠HCH = 116.43°—to investigate a case of two excited states with very similar components. According to EOM-CCSD, there are two excited states in the $^1A_1$ symmetry sector whose principal components are superpositions of the $1b_1 \rightarrow 2b_1$ and $2b_2 \rightarrow 3b_2$ single-electron transitions (the orbitals are plotted in Figure 5). These two states, which in the HF-orbital CASCI are the 2^1A_1 and 4^1A_1 states, are dominated by the sum of or difference between these components and turn out to have very similar natural orbital occupation numbers (see Figure 5). One might worry that such similarity could confuse a density-matrix-based approach, whereas it looks at first glance like the MOM approach should be effective because the CI vectors are clearly quite different, at least according to EOM-CCSD. As we will see, however, the more difficult of these two states confounds both the MOM and *W*Γ approaches as we defined them in section 2, and we were able to succeed with SS-CASSCF only by *increasing* the weighting of the density matrix difference within our $Q_{W\Gamma}$ measure.

Starting all of the CASSCF optimizations with HF orbitals as the initial guess, we found that the SRS, MOM, and *W*Γ approaches all succeed at finding the energy stationary point associated with the 2^1A_1 state. However, even though we used

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

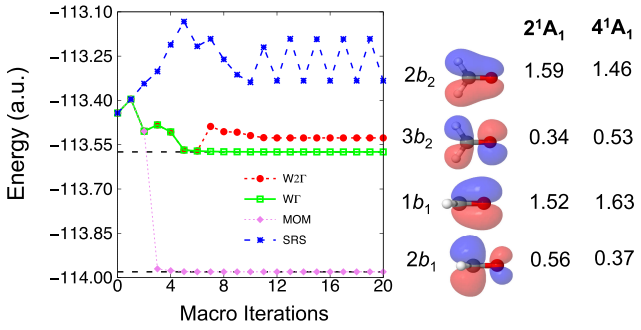


Figure 5. (left) Changes in the energy during SRS, MOM, $W\Gamma$, and reweighted $W2\Gamma$ CASSCF optimizations of the CH_2O 4^1A_1 state. The dashed horizontal lines indicate the ground and 2^1A_1 states. (right) Plots of the four most important orbitals contributing to these states (the SS-CASSCF natural orbitals for 4^1A_1 are shown; the corresponding 2^1A_1 orbitals are quite similar) and the corresponding SS-CASSCF natural orbital occupation numbers.

a nearly full-valence active space, none of these approaches succeeded for the 4^1A_1 state when applied as described in Theory. For example, Figure 5 shows how the MOM approach collapses to the ground state after just a few macro iterations. Interestingly, this 4^1A_1 state provides an example of a case in which the relative weighting of the W_0, W_1, and D components matters when employing the $W\Gamma$ quality measure defined in eq 9. When we doubled the weighting of the density matrix difference, i.e., $Q_{W\Gamma} \rightarrow Q_{W2\Gamma} = W_0 + W_1 + 2D$, our approach converged successfully to the 4^1A_1 state, as can be seen in Figure 5. Thus, even though this is a case where it was not obvious a priori that root flipping would occur and the states in question had easily distinguished CI vectors at the outset of the optimization, we found that the approach incorporating density matrix differences is more effective at converging to the relevant stationary points than either the SRS or MOM approach.

To ascertain how much the ability to find these states' stationary points matters at the end of the day, we compare CASPT2 and MRCI+Q results based on SS- and SA-CASSCF in Table 2. Here we employed two SA schemes, one giving equal weighting to the first four $^1A_1$ states (w1) and one using weightings of $(0.6, 0.1, 0.1, 0.2)$ and $(0.2, 0.1, 0.1, 0.6)$ when modeling the ground and 4^1A_1 states, respectively (w2). As CASPT2 was found to suffer from intruder states, we employed two different level shifts in each case. One difference that is noticed immediately is that compared with SA-CASPT2, the excitation energies for SS-CASPT2 are less sensitive to changes in the level shift, which is certainly a desirable property. For the 2^1A_1 state, we see that the CASPT2 and MRCI+Q excitation energies are not very sensitive to the choice between SA-CASSCF and SS-CASSCF. It should be noted that for this state there is some question about how accurate EOM-CCSD is expected to be, as the CASSCF CI vectors all show a significant weight on the $(2b_2)^2 \rightarrow (2b_1)^2$ double excitation and EOM-CCSD is known to overestimate the energies of double excitations because of its inability to relax their orbitals.[4−6] The 4^1A_1 state, on the other hand, does not have any significant doubly excited components, so the coupled cluster result should be more reliable for it. Indeed, we find that as we go in order of increasing state specificity, SA(w1) → SA(w2) → SS, the CASPT2 and MRCI+Q results for this state move toward the coupled cluster value, suggesting that there is a modest accuracy improvement to be had for this state by achieving SS-CASSCF. Recalling that the overall cost in large systems is dominated by post-CASSCF methods, we see that accuracy improvements of this type should be achievable at a negligible extra cost compared with an SA approach.

Table 2. Excitation Energies for the 2^1A_1 and 4^1A_1 States of CH_2O along with the IPEA Level Shifts Used for CASPT2

		excitation energy (eV)	
method	shift (au)	2^1A_1	4^1A_1
SA(w1)-CASPT2	0.1	9.71	10.94
SA(w1)-CASPT2	0.2	9.83	11.15
SA(w1)-MRCI+Q	N/A	9.86	11.10
SA(w2)-CASPT2	0.1	N/A	11.03
SA(w2)-CASPT2	0.2	N/A	11.19
SA(w2)-MRCI+Q	N/A	N/A	11.21
SS-CASPT2	0.1	9.74	11.24
SS-CASPT2	0.2	9.83	11.35
SS-MRCI+Q	N/A	9.83	11.30
EOM-CCSD	N/A	10.08	11.38

3.4. MgO: Vertical Excitation Energies. The next system we investigate is MgO at a bond distance of 1.8 Å. Like LiH, the ground state at this near-equilibrium geometry has ionic character, but in MgO the closed-shell determinant that dominates the ground-state wave function is somewhat doubly ionic, with both of Mg's 3s electrons moving into a bonding σ orbital with substantial O 2p character. We therefore expect to find a challenging assortment of low-lying excitations, including double charge transfers that return the system to more neutral states as well as excitations that retain the ground state's ionic nature. Indeed, Table 3 shows that the eight lowest-lying $^1A_1$ states in an initial CASCI calculation in the LDA orbital basis (LDA-CASCI) contain three excited states whose dipoles suggest that they largely retain the ground state's ionic nature as well as four charge-transfer states whose electron densities have shifted significantly toward the Mg

Table 3. Information on the Eight Lowest MgO $^1A_1$ States in the Initial LDA-CASCI Calculation, Including Their Dipole Moments (μ), Dominant Excitation Characters (DECs), and Primary Active Space Configurations (PASCs)[a]

state	label	μ (D)	DEC(s)	PASC(s)
1^1A_1	GS	−3.95	N/A	$5\sigma^2 2\pi^4 6\sigma^2$
2^1A_1	M1	−5.39	$6\sigma \rightarrow 7\sigma$	$5\sigma^2 2\pi^4 6\sigma^1 7\sigma^1$
			$2\pi \rightarrow 3\pi$	$5\sigma^2 2\pi^3 6\sigma^2 3\pi^1$
3^1A_1	V1	−4.88	$2\pi \rightarrow 3\pi$	$5\sigma^2 2\pi^3 6\sigma^2 3\pi^1$
4^1A_1	V2	−5.93	$6\sigma \rightarrow 8\sigma$	$5\sigma^2 2\pi^4 6\sigma^1 8\sigma^1$
5^1A_1	CT1	3.84	$2\pi^2 \rightarrow 7\sigma^2$	$5\sigma^2 2\pi^2 6\sigma^2 7\sigma^2$
6^1A_1	CT2	3.93	$2\pi^2 \rightarrow 7\sigma^2$	$5\sigma^2 2\pi^2 6\sigma^2 7\sigma^2$
			$6\sigma^2 \rightarrow 7\sigma^2$	$5\sigma^2 2\pi^4 7\sigma^2$
7^1A_1	CT4	2.33	$2\pi 6\sigma \rightarrow 3\pi 7\sigma$	$5\sigma^2 2\pi^3 6\sigma^1 3\pi^1 7\sigma^1$
8^1A_1	CT3	3.66	$2\pi 6\sigma \rightarrow 3\pi 7\sigma$	$5\sigma^2 2\pi^3 6\sigma^1 3\pi^1 7\sigma^1$
			$6\sigma^2 \rightarrow 7\sigma^2$	$5\sigma^2 2\pi^4 7\sigma^2$
			$2\pi^2 \rightarrow 7\sigma^2$	$5\sigma^2 2\pi^2 6\sigma^2 7\sigma^2$

[a]For states with two important configurations of nearly equal weight, we have listed both. The LDA orbital shapes are shown in Figure S1. The states' labels are ordered within a category (GS for ground state, M for missing, V for valence, CT for charge transfer) by their SS-CASSCF energies; the order of CT3 and CT4 inverts during SS optimization.

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

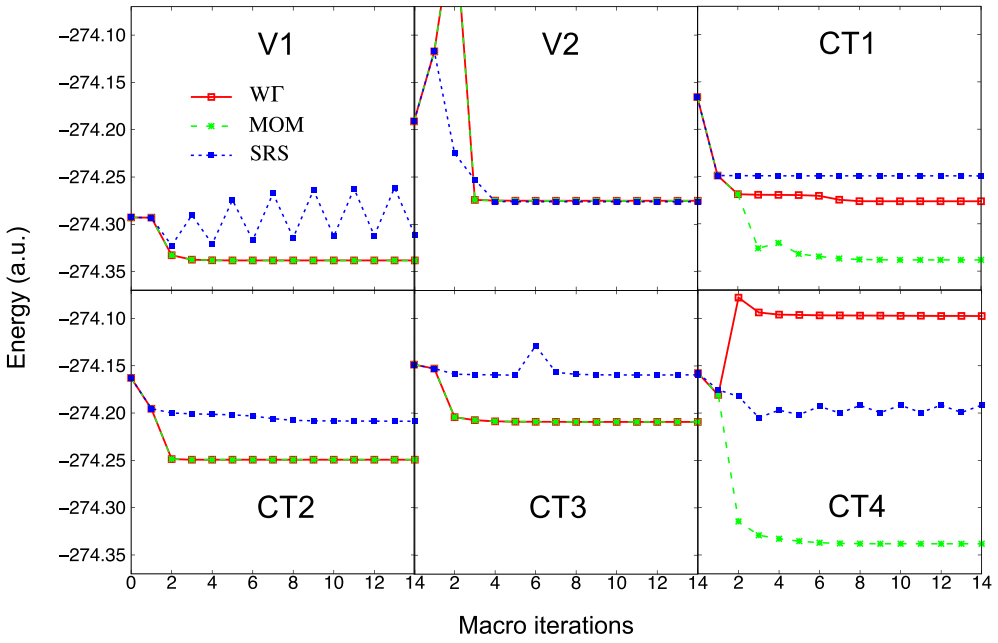


Figure 6. Energies of the selected CASCI roots at each macro iteration during SS-CASSCF optimizations of six MgO excited states under the SRS, MOM, and *W*Γ approaches to root selection.

atom. The prevalence of double excitations and even double charge transfers among these states makes MgO a clear example where a multireference treatment is necessary, but the mixture of neutral and ionic states makes it hard to know a priori how to construct an SA scheme that treats all states fairly. Ideally, this concern could be bypassed via state-specific optimization, in which the CASSCF energy stationary point corresponding to each of these states is located. However, as we now discuss, MgO proves to be especially difficult in this regard, with SRS failing to converge to the targeted excited state in every case and the MOM and *W*Γ approaches succeeding in only four and six of the seven cases, respectively.

To begin, we emphasize that the SRS approach fails to converge to the initially targeted state in every one of the seven MgO excited states studied here. That said, Figure 6 shows that although there are cases where SRS does not converge at all, it often finds a stationary point corresponding to a different state than the one originally targeted. This is true even for targeting of state V2, which in Figure 6 looks like a success. However, inspecting the converged wave function's properties, including natural orbital occupation numbers and dipole moments, and comparing them with those in Figure 7 and Table S2, we find that when targeting state V2, SRS converges to the nearly degenerate stationary point belonging to state CT1. Similarly, when targeting CT1, CT2, and CT3, SRS converges to the stationary points for CT2, CT3, and a state outside our set of seven, respectively.

Compared with SRS, both the MOM and *W*Γ approaches are substantially more effective. For the six states shown in Figure 6, they at least both converge to a stationary point in every case. Careful inspection of the final wave functions, including the natural orbital occupation and dipole moment analyses in Figure 7 and Table S2, shows that *W*Γ in fact converges to the intended stationary point in all six cases. In contrast, MOM succeeds in only four cases: when attempting to find the stationary point for the CT1 and CT4 states, it collapses to the V1 stationary point. We therefore see that, as

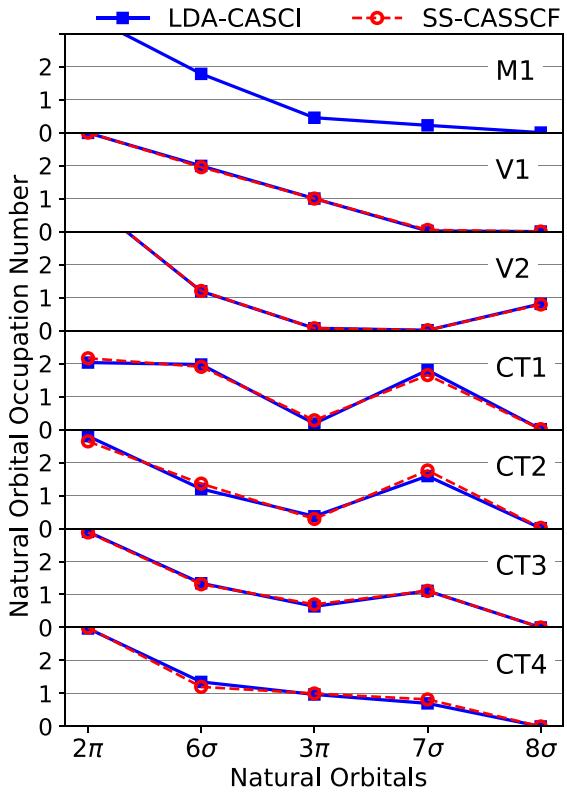


Figure 7. MgO's excited states may be distinguished using natural orbital occupation numbers.

in the other systems in this study, the *W*Γ approach is more effective for SS-CASSCF than CI-vector-based MOM, while both of these greatly outperform SRS. That said, we note that none of these methods succeeded in converging to the M1 stationary point. Indeed, we have tried different tracking schemes, even including the W_2 term and starting from SA-CASSCF orbitals obtained from a 10%/90% weighting, but we

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

have not been able to locate this point (see Figure S2). Thus, while the $W\Gamma$ approach is promising, it would clearly benefit from further improvements in future work.

In most cases, the success of $W\Gamma$ is primarily due to the Γ contribution, but the CT2 state was found to be an exception. As shown in Figure 8, when only the Γ criterion is used, there

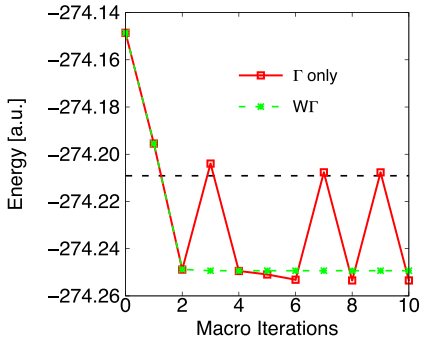


Figure 8. Changes in the total energy during SS-CASSCF optimizations for the CT2 state of MgO when selecting the CASCI root using Γ alone and when using the $W\Gamma$ composite approach. The horizontal dashed line indicates the energy of CT3's energy stationary point.

is an oscillation between CT2 and CT3. In order to gain insight into this situation, we provide a detailed analysis of the different contributions, W and Γ, to the composite $W\Gamma$ in Table S3. At some iterations, the Γ contributions for CT2 and CT3 are very similar. Therefore, it is difficult to distinguish these two states using Γ alone. Nevertheless, at all iterations the CT2 $W\Gamma$ values are smaller than the CT3 ones, implying that the composite $W\Gamma$ approach is essential in order to target this state. Curiously, despite the difficulties faced when using Γ alone, the approximate overlap in Q_{MOM} succeeds for this state.

As an aside, one could of course start from SA-CASSCF orbitals instead of LDA or HF orbitals. Whether this is helpful will of course be state-specific, but it is certainly not a panacea. For example, the CT1 state that was difficult for both SRS and MOM (see Figure 6) is still a problem for MOM starting from SA orbitals, as can be seen in Figure 9. Although SRS now succeeds (along with $W\Gamma$), MOM still collapses to the V1 stationary point as it did starting with LDA orbitals.

To investigate whether these SS-CASSCF solutions offer advantages when post-CASSCF methods are used, we first had to decide on SA schemes against which to compare them.

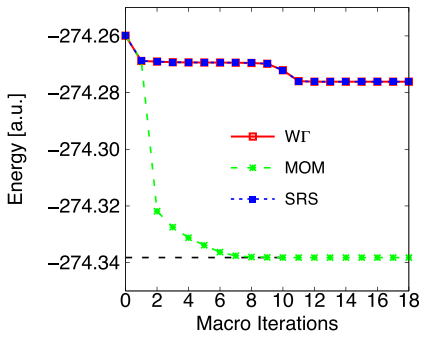


Figure 9. Changes in the total energy during SS-CASSCF optimizations for the CT1 state of MgO starting from SA-CASSCF orbitals. The horizontal dashed line indicates the energy of V1's energy stationary point.

There are of course an infinite number of possible SA weighting schemes, and it can be difficult to predict a priori what will be most effective. While this difficulty is especially concerning in a case like MgO where different states have significantly different charge distributions, here we used two SA schemes that try to achieve simple forms of balance. First, we used an equally weighted eight-state SA approach, which we denote as SA(8). Second, we attempted to minimize the number of states involved in the average by including only enough states to ensure that the desired state is present and equally weighting the states in the average. This SA(N) approach is much less straightforward because the Nth initial LDA-CASCI root does not necessarily show up in the same order (or at all) in an N-state state average. States V1, CT1, CT2, and CT3, show up as the Nth roots in equal-weight N-state SA calculations for N = 2, 3, 5, and 8, respectively. However, the near-degeneracy of V2 and CT1 at the CASSCF level prevents V2 from showing up at all in a four-state SA. The minimum number of states to include in order for V2 to show up turned out to be N = 6, in which case it appeared as the sixth root. State CT4 was even more problematic and indeed did not show up at all in equal-weight SA calculations for any $N \leq 8$, which is unsurprising in light of how much higher in energy its stationary point in Figure 6 is relative to the initial LDA-CASCI energy. Thus, in summary, CT4 was not included in the SA comparison, while the SA(N) approach for states V1, V2, CT1, CT2, and CT3 used the values N = 2, 6, 3, 5, and 8, respectively.

The LDA-CASCI, CASSCF, CASPT2, and MRCI+Q excitation energies for the six $^1A_1$ excited states for which SS-CASSCF stationary points were found are shown in Table 4. In Figure 10, we show the differences from the

Table 4. Excitation Energies (in eV) for Six $^1A_1$ Excited States of MgO[a]

method	V1	V2	CT1	CT2	CT3	CT4
LDA-CASCI	3.70	6.46	5.10	7.14	8.16	8.07
LDA-CASPT2	3.80	6.42	5.68	6.21	6.02	7.31
LDA-MRCI+Q	3.81	6.33	5.92	5.59	6.57	NC
SA(8)-CASSCF	3.97	7.24	5.10	5.67	6.94	NF
SA(8)-CASPT2	3.91	6.76	5.68	6.23	7.02	NF
SA(8)-MRCI+Q	3.81	6.31	5.92	6.46	6.66	NF
SA(N)-CASSCF	4.10	6.76	5.29	5.94	6.94	NF
SA(N)-CASPT2	3.89	6.57	5.66	6.32	7.02	NF
SA(N)-MRCI+Q	3.76	6.39	5.99	6.46	6.66	NF
SS-CASSCF	4.88	6.60	6.57	7.16	8.39	11.47
SS-CASPT2	3.76	6.55	5.73	6.32	6.91	10.44
SS-MRCI+Q	3.79	6.52	5.97	6.44	6.76	NC

[a]CASPT2 used an IPEA level shift of 0.2 au to avoid intruder states. In some cases, states were not found (NF) in the SA approach, and in others, the MRCI+Q Davidson solver did not converge (NC).

corresponding MRCI+Q excitation energies for both CASCI and CASPT2 when working in orbital bases from different methods (LDA, SA(8), SA(N), and SS-CASSCF). In every case, the SS-CASSCF orbital basis leads to the smallest difference between the CASPT2 and MRCI+Q excitation energies, which supports the hypothesis that state-specific orbital relaxations should be beneficial when relying on the perturbative assumption that the CASSCF wave function is close to the exact wave function. Unsurprisingly, working in the LDA orbital basis led to the largest differences, while the

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

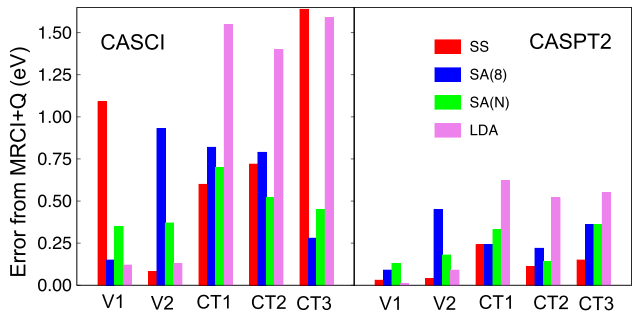


Figure 10. For different sets of molecular orbitals (SS-CASSCF, SA(8)-CASSCF, SA(*N*)-CASSCF, and LDA), we compare the (left) CASCI and (right) CASPT2 excitation energies with the MRCI+Q excitation energies. In each case, all three methods were carried out in the same basis, and the difference between the excitation energy for the method in question and the excitation energy for MRCI+Q was taken. CASPT2 used an IPEA level shift of 0.2 au to avoid intruder states.

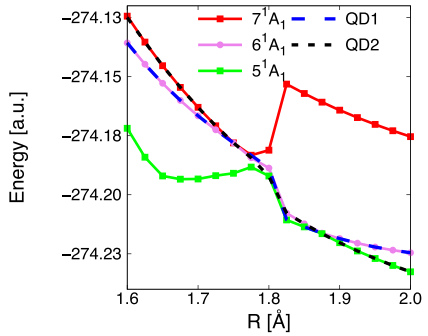


Figure 11. Energies of the SA-CASSCF adiabatic states 5^1A_1, 6^1A_1, and 7^1A_1 as functions of the MgO bond length. The dashed lines show, for each geometry, which states correspond to the quasi-diabats QD1 and QD2.

SA approaches were between these two extremes. While SS-CASSCF thus appears to offer improvements for CASPT2, it is important to note that its CASSCF excitation energies are often not closer to the corresponding MRCI+Q values compared to the situation in SA-CASSCF, as revealed by the left panel of Figure 10. This result should not be surprising, as CASSCF lacks all out-of-active-space weak correlation effects, the sizes of which are expected to differ significantly for different states. For example, neutral states have roughly 12 and eight electrons located on the Mg and O atoms, respectively, while the doubly ionic ground state has 10 and 10, so these states have different numbers of electrons in close proximity to each other. As most of the energetic effects coming from weak electron correlations are local in nature, a crude accounting of how many local electron pairs can be enumerated on each atom (12 choose 2 and 8 choose 2 compared to twice 10 choose 2) suggests that the size of weak correlation effects should be different for different states. Thus, while SS-CASSCF does not and is not expected to bring the CASSCF energetics closer to those of MRCI+Q, it does serve as a better platform for CASPT2 than either of the two SA approaches.

3.5. MgO: Potential Energy Surfaces. In many applications, including transition state searches and dynamics, it is highly desirable that an electronic state's predicted energy be a smooth function of molecular geometry. As has been observed in many other molecules,[45−47] the excited states of MgO are an example in which the SA approach can frustrate this desire. Starting at the slightly compressed bond distance of 1.6 Å and gradually stretching the bond, one finds that an equal-weight eight-state SA-CASSCF simulation predicts that the sixth and seventh $^1A_1$ states are near-degenerate and headed for an avoided crossing somewhere in the vicinity of 1.8 Å. However, as can be seen in Figure 11, this SA approach suffers from discontinuities in the PESs near this avoided crossing due to a discontinuous change in the set of molecular orbitals that give the lowest SA energy. In many contexts, dynamic weighting schemes have been successful at removing such discontinuities[45,72] by enlarging the weight of the state in question relative to those of other states on the basis of the energy differences between them. In essence, these schemes seek to (and often do) remove the discontinuity by making the SA approach more SS, which begs the question of whether a fully SS approach could provide the same benefit. In exploring this question in the context of attempting to follow the 6^1A_1 and 7^1A_1 states as the bond is stretched, we find that the energy stationary points that SS-CASSCF locates form smooth PESs and, interestingly, are diabatic in character.

To start, we note that at every bond distance there are two SA-CASSCF states whose natural orbital occupations closely match those of 6^1A_1 and 7^1A_1 at 1.6 Å. This observation is made clear in Figure 12, where the solid lines show, at each

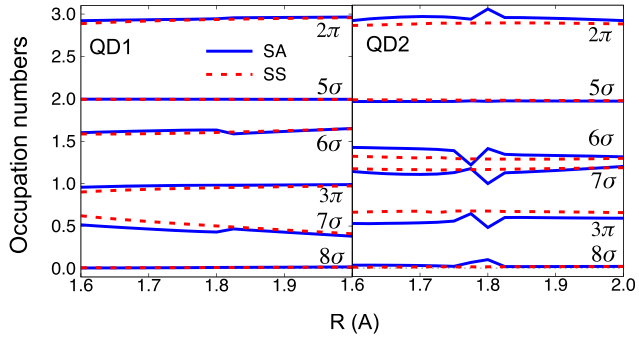


Figure 12. SA- and SS-CASSCF natural orbital occupation numbers for the (left) QD1 and (right) QD2 states as functions of the MgO bond length.

bond distance, the occupations of the SA-CASSCF states most similar to the 1.6 Å values for 6^1A_1 (on the left) and 7^1A_1 (on the right). Taking these collections of states as quasi-diabats QD1 and QD2, respectively, we find that they also maintain consistent dipole moments during the bond stretching, as shown in Figure 13, and are dominated by the same configurations. QD1 is dominated by two configurations: $5\sigma^2 2\pi^3 6\sigma^2 3\pi^1$ and $5\sigma^2 2\pi^3 6\sigma^1 3\pi^1 7\sigma^1$, whereas QD2 is dominated by three: $5\sigma^2 2\pi^3 6\sigma^1 3\pi^1 7\sigma^1$, $5\sigma^2 2\pi^4 7\sigma^2$, and $5\sigma^2 2\pi^2 6\sigma^2 7\sigma^2$. Most remarkably, using these states as the initial guesses for SS-CASSCF optimizations leads to two smooth potential energy surfaces (Figure 13, bottom) corresponding to the energies of two different CASSCF energy stationary points that move smoothly through the CI and orbital parameter space as the bond is stretched. These surfaces cross and thus violate the diatomic no-crossing rule, suggesting that these strings of CASSCF stationary points are diabatic in character, a notion that is confirmed by inspecting their dipole moments (Figure 13, top) and occupations (Figure 12). Of course, if this were full CI, these strings of

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

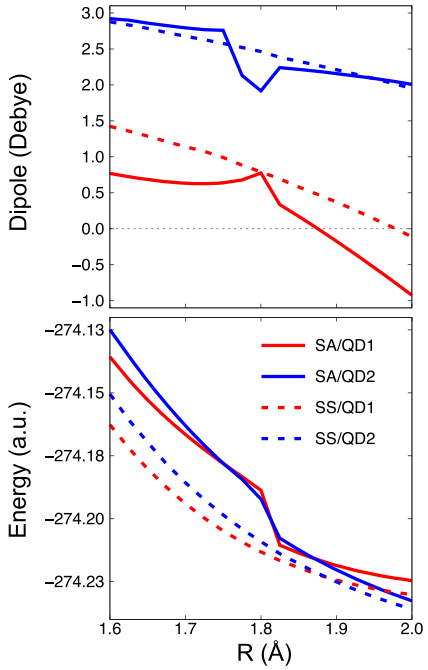


Figure 13. (top) Dipole moments and (bottom) energies of the QD1 and QD2 states as functions of the MgO bond length.

stationary points would not cross and would correspond to adiabats, but it seems that the approximate CASSCF wave function leads to stationary-point strings that are both nonorthogonal and of consistent diabatic character.

Thus, we find that like dynamic weighting schemes that lie somewhere between equal-weight SA and full state specificity, SS-CASSCF is capable of rectifying PES discontinuities related to state averaging. If one desires proper adiabats, these states would need to be rediagonalized, which could be achieved either through nonorthogonal CI methods[73,74] or variational Monte Carlo.[75] These adiabats would also be smooth, but for many applications one has to wonder whether this step is even necessary, as states with smooth diabatic PESs are often desirable, as for example in surface hopping simulations.[76,77] Either way, we see that SS-CASSCF has the potential to restore smooth PESs for highly excited and strongly correlated states.

4. CONCLUSION

We have presented an approach for locating state-specific CASSCF energy stationary points that overcomes root flipping via a metric of quality based on an excited-state variational principle and density matrix differences. This approach was inspired by both the maximum overlap method and recent progress in excited-state variational methods as well as the twin goals of controlling cost and maintaining compatibility with widely used two-step methods that treat orbital and CI coefficients in separate optimization steps. Our central finding is that this approach is highly effective in the face of both root flipping and near-degeneracies, significantly outperforming both simple root selection and a CI-vector-based adaptation of the maximum overlap method. Although the improvements can be modest, we find that compared with state averaging, state-specific CASSCF provides a superior starting point for CASPT2 in that it brings this method's excitation energies more closely in line with those of the more reliable but expensive MRCI+Q method. In addition, we find that, in MgO at least, the state-specific approach yields smooth, quasi-diabatic PESs in a case where state averaging leads to discontinuities. As the additional cost of converging each CASSCF state individually is typically small compared with the cost of post-CASSCF methods, we are happy to recommend that applications of this approach be explored across a wider spectrum of chemical systems.

Aside from post-CASSCF accuracy improvements, the ability to locate the energy stationary points of individual excited states delivers a number of important and useful properties. Unlike state-averaged CASSCF, the excited-state-specific version retains the size consistency of its ground-state counterpart. Furthermore, stationarity with respect to the energy makes the calculation of many properties, not least of which are the nuclear gradients, significantly more straightforward because it avoids the need for Lagrangian or Z-vector techniques. State-specific optimization also avoids the need to decide on how many states to average and what weights to use, choices that not only are difficult to justify a priori but also can have significant effects on CASSCF and post-CASSCF energetics. Finally, although it is less straightforward than in the case of SA-CASSCF, one can evaluate interstate properties such as spin couplings or nonadiabatic couplings for SS-CASSCF by employing either Malmqvist's biorthonormal transformation[78,79] or by relying on variational Monte Carlo's ability to work directly with multiple states built from different orbital sets.

While we have presented evidence in this study of how our approach to stationary points can be beneficial to CASPT2, other excited-state methods are likely to benefit as well. For example, the most difficult optimization stage in the application of variational Monte Carlo to excited states tends to be the step in which orbital relaxations are enabled. Thus, questions that we are very eager to answer in future work are whether orbitals optimized for an individual excited state by CASSCF could serve as a good alternative to those optimized by Monte Carlo and whether these two sources of optimized orbitals make any difference in practice when fed into the diffusion Monte Carlo method. Back in the area of quantum chemistry, it is quite possible that the recently introduced excited-state mean field method could be accelerated by two-step optimization methods, at which point it too would come face to face with root-flipping issues that the $W\Gamma$ approach presented here might alleviate. Finally, it will be exciting to pair this improved state-specific methodology with large active space methods, which already offer substantial advantages when dealing with root flipping.

Looking forward, there are numerous promising ways to extend and apply this new methodology. Although we have chosen to avoid the double-excitation components of the variational principle in the present study to keep the approach simple and inexpensive, an excellent approximation of the effect of this term should be accessible via the application of modern selective CI integral handling technology. Similarly, although the two-body reduced density matrix in the full orbital space may be impractical to employ for root comparisons because of high memory demands, one can imagine inspecting it in a second, much larger active space. As we have already identified at least one state in MgO that continues to defy all of the attempted state-specific optimizations, we are eager to explore these extensions in order to make the methodology more reliable and robust. In

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

terms of applications, the freedom to define the target state in whatever way the user chooses would appear to be a natural fit when trying to optimize the two (or more!) adiabatic states near a conical intersection. While these states mix freely at the intersection itself, making it less meaningful to differentiate between them in terms of their diabatic character, away from but near the intersection the use of diabatic states as the targets for state-specific CASSCF optimization could help identify which adiabatic states are which and also help ensure that all of the states in question are found in the state-specific optimization process. More generally, it will be interesting to investigate how a more aggressively state-specific approach to excited states in CASSCF can be applied to the ongoing challenge of handling the strong excited-state orbital relaxations common to core and charge-transfer excitations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.9b00351.

Tables S1−S3 and Figures S1 and S2 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: lantrann@berkeley.edu.

*E-mail: eneuscamman@berkeley.edu.

ORCID

Lan Nguyen Tran: 0000-0003-2649-2122

Eric Neuscamman: 0000-0002-4760-8238

Funding

This work was supported by the Early Career Research Program of the Office of Science, Office of Basic Energy Sciences, U.S. Department of Energy (Grant DE-SC0017869). J.A.R.S. acknowledges additional support from the National Science Foundation's Graduate Research Fellowship Program. Calculations were performed both on our own desktop computers and using the UC Berkeley Savio computer cluster.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Dreuw, A.; Head-Gordon, M. Single-Reference ab Initio Methods for the Calculations of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105*, 4009.

(2) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. *J. Chem. Phys.* **2003**, *119*, 2943−2946.

(3) Subotnik, J. E. Communication: configuration interaction singles has a large systematic bias against charge-transfer states. *J. Chem. Phys.* **2011**, *135*, 071104.

(4) Krylov, A. I. Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhikers Guide to Fock Space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433.

(5) Watts, J. D.; Bartlett, R. J. Iterative and non-iterative triple excitation corrections in coupled-cluster methods for excited electronic states: the EOM-CCSDT-3 and EOM-CCSD (T) methods. *Chem. Phys. Lett.* **1996**, *258*, 581−588.

(6) Neuscamman, E. Communication: Variation After Response in Quantum Monte Carlo. *J. Chem. Phys.* **2016**, *145*, 081103.

(7) Liu, J.; Matthews, D.; Coriani, S.; Cheng, L. Benchmark calculations of K-edge ionization energies for first-row elements using scalar-relativistic core-valence-separated equation-of-motion coupled-cluster methods. *J. Chem. Theory Comput.* **2019**, *15*, 1642−1651.

(8) Onida, G.; Reining, L.; Rubio, A. Electronic excitations: density-functional versus many-body Greens-function approaches. *Rev. Mod. Phys.* **2002**, *74*, 601.

(9) Dreuw, A.; Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *WIR: Comput. Mol. Sci.* **2015**, *5*, 82−95.

(10) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 41.

(11) Werner, H.-J.; Knowles, P. J. A second order multiconfiguration SCF procedure with optimum convergence. *J. Chem. Phys.* **1985**, *82*, 5053.

(12) Knowles, P. J.; Werner, H.-J. An efficient second-order MC SCF method for long configuration expansions. *Chem. Phys. Lett.* **1985**, *115*, 259.

(13) Roos, B. O. The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations. *Adv. Chem. Phys.* **2007**, *69*, 399−445.

(14) Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. A. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *J. Chem. Phys.* **1988**, *89*, 2185−2192.

(15) Malmqvist, P.-Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO2 and Cu2O2 systems. *J. Chem. Phys.* **2008**, *128*, 204109.

(16) Li Manni, G.; Ma, D.; Aquilante, F.; Olsen, J.; Gagliardi, L. SplitGAS method for strong correlation and the challenging case of Cr2. *J. Chem. Theory Comput.* **2013**, *9*, 3375−3384.

(17) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications: Mineola, N.Y., 1996.

(18) Ghosh, D.; Hachmann, J.; Yanai, T.; Chan, G. K.-L. Orbital optimization in the density matrix renormalization group, with applications to polyenes and β-carotene. *J. Chem. Phys.* **2008**, *128*, 144117.

(19) Zgid, D.; Nooijen, M. The density matrix renormalization group self-consistent field method: Orbital optimization with the density matrix renormalization group method in the active space. *J. Chem. Phys.* **2008**, *128*, 144116.

(20) Thomas, R. E.; Sun, Q.; Alavi, A.; Booth, G. H. Stochastic Multiconfigurational Self-Consistent Field Theory. *J. Chem. Theory Comput.* **2015**, *11*, 5316.

(21) Li Manni, G.; Smart, S. D.; Alavi, A. Combining the Complete Active Space Self-Consistent Field Method and the Full Configuration Interaction Quantum Monte Carlo within a Super-CI Framework, with Application to Challenging Metal-Porphyrins. *J. Chem. Theory Comput.* **2016**, *12*, 1245−1258.

(22) Smith, J. E. T.; Mussard, B.; Holmes, A. A.; Sharma, S. Cheap and near exact CASSCF with large active spaces. *J. Chem. Theory Comput.* **2017**, *13*, 5468−5478.

(23) Lan, T. N.; Zgid, D. Generalized self-energy embedding theory. *J. Phys. Chem. Lett.* **2017**, *8*, 2200−2205.

(24) Tran, L. N.; Iskakov, S.; Zgid, D. Spin-unrestricted self-energy embedding theory. *J. Phys. Chem. Lett.* **2018**, *9*, 4444−4450.

(25) Diffenderfer, R. N.; Yarkony, D. R. Use of the state-averaged MCSCF procedure: application to radiative transitions in magnesium oxide. *J. Phys. Chem.* **1982**, *86*, 5098−5105.

(26) Yarkony, D. R. Conical intersections: Diabolical and often misunderstood. *Acc. Chem. Res.* **1998**, *31*, 511−518.

(27) Levine, B. G.; Coe, J. D.; Martínez, T. J. Optimizing conical intersections without derivative coupling vectors: application to multistate multireference second-order perturbation theory (MS-CASPT2). *J. Phys. Chem. B* **2008**, *112*, 405−413.

(28) Levine, B. G.; Ko, C.; Quenneville, J.; Martínez, T. J. Conical intersections and double excitations in time-dependent density functional theory. *Mol. Phys.* **2006**, *104*, 1039−1051.

(29) Garavelli, M.; Bernardi, F.; Olivucci, M.; Vreven, T.; Klein, S.; Celani, P.; Robb, M. A. Potential-energy surfaces for ultrafast photochemistry static and dynamic aspects. *Faraday Discuss.* **1998**, *110*, 51−70.
(30) Bagus, P. S. Self-consistent-field wave functions for hole states of some Ne-like and Ar-like ions. *Phys. Rev.* **1965**, *139*, A619.
(31) Hsu, H.-l.; Davidson, E. R.; Pitzer, R. M. An SCF method for hole states. *J. Chem. Phys.* **1976**, *65*, 609−613.
(32) Naves de Brito, A.; Correia, N.; Svensson, S.; Ågren, H. A theoretical study of x-ray photoelectron spectra of model molecules for polymethylmethacrylate. *J. Chem. Phys.* **1991**, *95*, 2965−2974.
(33) Besley, N. A.; Gilbert, A. T.; Gill, P. M. Self-consistent-field calculations of core excited states. *J. Chem. Phys.* **2009**, *130*, 124308.
(34) Jensen, H. J. A.; Jørgensen, P.; Ågren, H. Efficient optimization of large scale MCSCF wave functions with a restricted step algorithm. *J. Chem. Phys.* **1987**, *87*, 451−466.
(35) Werner, H.-J.; Meyer, W. A quadratically convergent MCSCF method for the simultaneous optimization of several states. *J. Chem. Phys.* **1981**, *74*, 5794−5801.
(36) Bouabça, T.; Ben Amor, N.; Maynau, D.; Caffarel, M. A study of the fixed-node error in quantum Monte Carlo calculations of electronic transitions: The case of the singlet n π(CO) transition of the acrolein. *J. Chem. Phys.* **2009**, *130*, 114107.
(37) Malmqvist, P.-Å.; Roos, B. O. The CASSCF state interaction method. *Chem. Phys. Lett.* **1989**, *155*, 189−194.
(38) Fdez. Galván, I.; Delcey, M. G.; Pedersen, T. B.; Aquilante, F.; Lindh, R. Analytical state-average complete-active-space self-consistent field nonadiabatic coupling vectors: Implementation with density-fitted two-electron integrals and application to conical intersections. *J. Chem. Theory Comput.* **2016**, *12*, 3636−3653.
(39) Serrano-Andrés, L.; Merchán, M.; Lindh, R. Computation of conical intersections by using perturbation techniques. *J. Chem. Phys.* **2005**, *122*, 104107.
(40) Granovsky, A. A. Extended multi-configuration quasi-degenerate perturbation theory: The new approach to multi-state multi-reference perturbation theory. *J. Chem. Phys.* **2011**, *134*, 214113.
(41) Gozem, S.; Melaccio, F.; Valentini, A.; Filatov, M.; Huix-Rotllant, M.; Ferre, N.; Frutos, L. M.; Angeli, C.; Krylov, A. I.; Granovsky, A. A.; et al. Shape of multireference, equation-of-motion coupled-cluster, and density functional theory potential energy surfaces at a conical intersection. *J. Chem. Theory Comput.* **2014**, *10*, 3074−3084.
(42) Snyder, J. W., Jr; Fales, B. S.; Hohenstein, E. G.; Levine, B. G.; Martínez, T. J. A direct-compatible formulation of the coupled perturbed complete active space self-consistent field equations on graphical processing units. *J. Chem. Phys.* **2017**, *146*, 174113.
(43) Domingo, A.; Carvajal, M. À.; de Graaf, C.; Sivalingam, K.; Neese, F.; Angeli, C. Metal-to-metal charge-transfer transitions: reliable excitation energies from ab initio calculations. *Theor. Chem. Acc.* **2012**, *131*, 1264.
(44) Meyer, B.; Domingo, A.; Krah, T.; Robert, V. Charge transfer processes: the role of optimized molecular orbitals. *Dalton Trans* **2014**, *43*, 11209−11215.
(45) Deskevich, M. P.; Nesbitt, D. J.; Werner, H.-J. Dynamically weighted multiconfiguration self-consistent field: Multistate calculations for F+ H 2 O HF+ OH reaction paths. *J. Chem. Phys.* **2004**, *120*, 7281−7289.
(46) Hollas, D.; Sistik, L.; Hohenstein, E. G.; Martínez, T. J.; Slavicek, P. Nonadiabatic ab initio molecular dynamics with the floating occupation molecular orbital-complete active space configuration interaction method. *J. Chem. Theory Comput.* **2018**, *14*, 339−350.
(47) Snyder, J. W., Jr; Parrish, R. M.; Martínez, T. J. α-CASSCF: An Efficient, Empirical Correction for SA-CASSCF To Closely Approximate MS-CASPT2 Potential Energy Surfaces. *J. Phys. Chem. Lett.* **2017**, *8*, 2432−2437.
(48) Pineda Flores, S. D.; Neuscamman, E. Excited State Specific Multi-Slater Jastrow Wave Functions. *J. Phys. Chem. A* **2019**, *123*, 1487−1497.
(49) Shea, J. A.; Neuscamman, E. Communication: A mean field platform for excited state quantum chemistry. *J. Chem. Phys.* **2018**, *149*, 081101.
(50) Weinstein, D. Modified Ritz method. *Proc. Natl. Acad. Sci. U. S. A.* **1934**, *20*, 529.
(51) MacDonald, J. On the modified Ritz variation method. *Phys. Rev.* **1934**, *46*, 828.
(52) Messmer, R. P. On a variational method for determining excited state wave functions. *Theor. Chim. Acta.* **1969**, *14*, 319−328.
(53) Choi, J. H.; Lebeda, C. F.; Messmer, R. P. Variational principle for excited states: Exact formulation and other extensions. *Chem. Phys. Lett.* **1970**, *5*, 503−506.
(54) Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. Self-consistent field calculations of excited states using the maximum overlap method (MOM). *J. Phys. Chem. A* **2008**, *112*, 13164−13171.
(55) Barca, G. M.; Gilbert, A. T.; Gill, P. M. Simple Models for Difficult Electronic Excitations. *J. Chem. Theory Comput.* **2018**, *14*, 1501−1509.
(56) Umrigar, C.; Wilson, K.; Wilkins, J. Optimized trial wave functions for quantum Monte Carlo calculations. *Phys. Rev. Lett.* **1988**, *60*, 1719.
(57) Ye, H.-Z.; Welborn, M.; Ricke, N. D.; Van Voorhis, T. σ-SCF: A direct energy-targeting method to mean-field excited states. *J. Chem. Phys.* **2017**, *147*, 214104.
(58) Zhao, L.; Neuscamman, E. An efficient variational principle for the direct optimization of excited states. *J. Chem. Theory Comput.* **2016**, *12*, 3436.
(59) Zhao, L.; Neuscamman, E. A Blocked Linear Method for Optimizing Large Parameter Sets in Variational Monte Carlo. *J. Chem. Theory Comput.* **2017**, *13*, 2604−2611.
(60) Shea, J. A. R.; Neuscamman, E. Size Consistent Excited States via Algorithmic Transformations between Variational Principles. *J. Chem. Theory Comput.* **2017**, *13*, 6078.
(61) Blunt, N. S.; Neuscamman, E. Charge-transfer excited states: Seeking a balanced and efficient wave function ansatz in variational Monte Carlo. *J. Chem. Phys.* **2017**, *147*, 194101.
(62) Robinson, P. J.; Pineda Flores, S. D.; Neuscamman, E. Excitation variance matching with limited configuration interaction expansions in variational Monte Carlo. *J. Chem. Phys.* **2017**, *147*, 164114.
(63) Zhao, L.; Neuscamman, E. A. Variational Excitations in Real Solids: Optical Gaps and Insights into Many-Body Perturbation Theory. 2018, arXiv:1804.09663 [cond-mat.str-el]. arXiv.org e-Print archive. https://arxiv.org/abs/1804.09663
(64) Blunt, N. S.; Neuscamman, E. Excited-state diffusion Monte Carlo calculations: a simple and efficient two-determinant ansatz. *J. Chem. Theory Comput.* **2019**, *15*, 178−189.
(65) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483−5488.
(66) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.
(67) https://www.tensorflow.org (accessed March 22, 2019).
(68) Tamayo-Mendoza, T.; Kreisbeck, C.; Lindh, R.; Aspuru-Guzik, A. Automatic Differentiation in Quantum Chemistry with Applications to Fully Variational Hartree−Fock. *ACS Cent. Sci.* **2018**, *4*, 559.
(69) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; et al. PySCF: the Python-based simulations of chemistry framework. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1340.
(70) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 242.
(71) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184.

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803

(72) Glover, W. Communication: Smoothing out excited-state dynamics: analytical gradients for dynamically weighted complete active space self-consistent field. *J. Chem. Phys.* **2014**, *141*, 171102−171102.
(73) Thom, A. J.; Head-Gordon, M. Hartree−Fock solutions as a quasidiabatic basis for nonorthogonal configuration interaction. *J. Chem. Phys.* **2009**, *131*, 124113.
(74) Jensen, K. T.; Benson, R. L.; Cardamone, S.; Thom, A. J. Modeling Electron Transfers Using Quasidiabatic Hartree−Fock States. *J. Chem. Theory Comput.* **2018**, *14*, 4629−4639.
(75) Toulouse, J.; Assaraf, R.; Umrigar, C. J. *Adv. Quantum Chem.* **2016**, *73*, 285−314.
(76) Van Voorhis, T.; Kowalczyk, T.; Kaduk, B.; Wang, L.-P.; Cheng, C.-L.; Wu, Q. The diabatic picture of electron transfer, reaction barriers, and molecular dynamics. *Annu. Rev. Phys. Chem.* **2010**, *61*, 149−170.
(77) Wang, L.; Akimov, A.; Prezhdo, O. V. Recent progress in surface hopping: 2011−2015. *J. Phys. Chem. Lett.* **2016**, *7*, 2100−2112.
(78) Malmqvist, P. Å. Calculation of transition density matrices by nonunitary orbital transformations. *Int. J. Quantum Chem.* **1986**, *30*, 479−494.
(79) Knecht, S.; Keller, S.; Autschbach, J.; Reiher, M. A nonorthogonal state-interaction approach for matrix product state wave functions. *J. Chem. Theory Comput.* **2016**, *12*, 5881−5894.

DOI: 10.1021/acs.jctc.9b00351
J. Chem. Theory Comput. 2019, 15, 4790−4803