

# Non-orthogonal configuration interaction with single substitutions for the calculation of core-excited states

Katherine J. Oosterbaan,<sup>1,2</sup> Alec F. White,<sup>1,2</sup> and Martin Head-Gordon<sup>1,2, a)</sup>

<sup>1)</sup>*Department of Chemistry, University of California, Berkeley, California 94720, USA*

<sup>2)</sup>*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

In this paper, we present the non-orthogonal configuration interaction singles (NO-CIS) method for calculating core-excited states of closed-shell molecules. NOCIS is a black-box variant of NOCI, which uses  $A$  different core-ionized determinants for a molecule with  $A$  atoms of a given element to form single substitutions. NOCIS is a variational, spin-pure, size-consistent ab-initio method that dramatically improves on standard CIS by capturing essential orbital relaxation effects, in addition to essential configuration interaction. We apply it to the calculation of core-excitations for several smaller molecules, and demonstrate that it performs competitively with other Hartree-Fock and DFT-based methods. We also benchmark it in several basis sets.

---

<sup>a)</sup>Electronic mail: [mhg@cchem.berkeley.edu](mailto:mhg@cchem.berkeley.edu)

## INTRODUCTION

Due to the increasing accessibility of synchrotron radiation, x-ray absorption spectroscopy (XAS) methods have become far more powerful, and are able to study an increasing number of systems. However, the theoretical calculation of core-excited states, which are often needed to identify XAS peaks, still has space to improve. An ideal core-excited state calculation is one which, while still efficient, considers the interactions between the different possible core excitations and correctly treats electron localization on atoms of the same type. If possible, it should also be systematically improvable.

There are many different approaches to calculating core-excited states. Early methods used multiple scattering  $X_\alpha$ <sup>1</sup>, static exchange<sup>2</sup>, and transition potential calculations<sup>3</sup>, but several methods have evolved since then. One such method is configuration interaction singles (CIS)<sup>4</sup>, which is based on a Hartree-Fock (HF) reference for an excited state that considers all single-electron excitations. However, CIS calculations do not allow for orbital relaxation<sup>5</sup>, or include dynamic correlation<sup>6</sup>, meaning that their results are not particularly accurate. Dynamic correlation via CIS(D)<sup>7,8</sup> offers only limited improvement<sup>9</sup>.

Another method is time-dependent density functional theory (TDDFT), which calculates poles of the DFT response function from a CIS-like equation. While it allows for only very limited orbital relaxation, it does include differential electron correlation effects. Another problem is that a TDDFT potential does not have the correct long-range Coulomb tail, due to the self-interaction problem, meaning that excitation energies corresponding to states that use this tail are not accurate<sup>10</sup>, so this method has been most widely used for valence-excited states. These methods scale can scale from  $N^2$  to  $N^4$  depending on how they are implemented<sup>11</sup>. Significant effort, nonetheless, has been devoted to modifying standard density functionals to improve the performance of TDDFT for predicting core excitation energies<sup>12-16</sup>. In particular, in contrast with the success of long-range corrected functionals<sup>17</sup> for valence Rydberg and charge-transfer excited states, short-range corrected functionals<sup>18</sup> improve core-excited states.

Another set of methods is delta-self-consistent field ( $\Delta$ -SCF) and  $\Delta$ -DFT, which scale  $N^2$ - $N^3$  depending on implementation. In these methods, the excited state energy is obtained by subtracting a ground-state from an excited-state HF or Kohn-Sham energy calculation<sup>18</sup>. Since the excited state is computed separately and explicitly, this method allows for orbital

relaxation.  $\Delta$ -SCF does not include electron correlation effects, while  $\Delta$ -DFT does through its functional<sup>19</sup>. Another crucial component of this method is the functional used, as different functionals produce different results for different types of calculations. Some shortfalls of these methods are that  $\Delta$ -DFT suffers from the same self-interaction problem as TDDFT, and that both  $\Delta$  methods are not spin pure, which is important when simulating XAS spectra. Additionally, they do not consider static correlation in the case of molecules with several of the same atom.

A different method is algebraic diagrammatic construction (ADC)<sup>20</sup>, which scales as  $N^5$  for ADC(2), the most common variant used<sup>21,22</sup>. This approximation to the polarization propagator comes from Green's function theory and it has been derived with the diagrammatic perturbation expansion of the polarization propagator using the Møller-Plesset partition of the Hamiltonian. For core-excited states, there exists the core-valence separation (CVS) addition to the ADC scheme<sup>23-27</sup>, which uses the decoupling of the space of core-excitations from the valence in the ADC matrix. Orbital relaxation effects are included indirectly in ADC(2), the second-order Møller-Plesset perturbation (MP2)-corrected version. Electron correlation is also included, resulting in a relatively high degree of accuracy<sup>28</sup>.

A final set of methods to consider are those that come from equation-of-motion coupled-cluster singles and doubles (EOM-CCSD) methods<sup>29</sup>, specifically those for electronically excited states (EOM-EE) and electron affinities (EOM-EA), which scale as  $N^6$ . In these types of calculations, target states are excitations from a reference state, and they are constructed using an excitation operator. For EOM-EE, the reference state is the closed-shell CC ground state, and the excitation operator conserves the number of  $\alpha$  and  $\beta$  electrons. While EOM-EA and EOM-EE have the same reference state (typically the CCSD ground state), they differ in that EOM-EA uses electron attaching excitation operators, while EOM-EE uses electron exciting operators. Both methods include correlation effects<sup>30</sup>, but while EOM-EE does not fully include orbital relaxation<sup>31</sup>, EOM-EA does, as it is used with a core-ionized reference<sup>32</sup>.

There have been several methods developed recently that have also made important strides toward the computation of high-energy states. One such example is the DFT/ROCIS work of Roemelt et al.<sup>33</sup> in which was specifically designed for L-edge XAS and combines DFT with CI through the use of Kohn-Sham orbitals. Other work has been done by Coriani et al.<sup>34</sup> on damped coupled-cluster linear response functions to extend this method from UV

CAS spectra.

It is also important to discuss non-orthogonal configuration interaction (NOCI) methods for excited states<sup>35–38</sup>. NOCI uses different Hartree-Fock solutions as the basis for the CI calculation, which renders the basis non-orthogonal. This allows the basis to more effectively span the relevant part of the configuration space than an orthogonal calculation<sup>39</sup>. Since the creation of this method, further work has been done to increase the accuracy of NOCI calculations, for example, the spin-flip NOCI (SF-NOCI)<sup>40</sup> and NOCI-MP2<sup>41</sup> methods. In SF-NOCI, the wavefunction is defined as a linear combination of independently relaxed Slater determinants obtained from all possible spin-flipping excitations, and it improves on several less-desirable elements in the original NOCI calculation process<sup>40</sup>. NOCI-MP2 was designed to add dynamic correlation to NOCI. The formalism has recently been improved to permit size-consistency and an efficient implementation has been developed<sup>41</sup>.

In this paper, we present a novel method for the calculation of single and multi-reference core-excited states: non-orthogonal configuration interaction singles, or NOCIS. This method considers the interactions not only between singly-excited states of different core orbitals but also between localized excited states for a molecule with two or more of the same atom. While NOCI is a general framework, where the appropriate determinants are chosen on a case-by-case basis, NOCIS is designed as a black-box method that includes a precisely defined set of determinants. NOCIS allows for orbital relaxation by using as its basis the electron-attached states of independently optimized, core-ionized references, providing, as we shall demonstrate, an impressive improvement over the standard orthogonal CIS method. Thus, it can efficiently produce all core-excitation energy values for a given molecule with a high degree of accuracy. It is important to note that this method is identical to the static exchange approximation (STEX) method<sup>42</sup> for a single reference, but its invocation of non-orthogonality for multi-reference states renders it unique from other average-of-configurations methods such as 4C STEX<sup>43</sup>, which still use orthogonal orbitals.

## II. THEORY

The NOCIS wavefunction is constructed from  $R$  different core-ionized determinants corresponding to core electron removal from each identical atom of interest (eg.  $R=3$  for C in C<sub>3</sub>H<sub>4</sub>O) and ROHF optimization of those cation determinants. Electrons are then reat-

placed into the  $V$  virtual orbitals (different for each cation determinant), and the resulting core-excited configuration is projected against the closed-shell ground state determinant to ensure that the core-excited states are strongly orthogonal to the ground state. The amplitudes  $t_\alpha^a$  are unknowns to be variationally determined:

$$|\Psi_{NOCIS}\rangle = \sum_{i=1}^R \sum_{a=1}^V \mathbf{t}_i^a (1 - \hat{\mathbf{P}}) [\hat{a}_a^\dagger \hat{a}_i \pm \hat{a}_a^\dagger \hat{a}_{\bar{i}}] |\Phi_i\rangle \quad (1)$$

Here  $|\Phi_i\rangle$  is the ROHF determinant with its core hole on atom  $i$  filled.  $\hat{\mathbf{P}}$  acts by projecting against the neutral ground state reference. Since Eq. 1 is a linear trial function, application of the variational principle gives a generalized eigenvalue problem for the core-excited state energies  $E_k$  and amplitudes  $\mathbf{t}_k$ .

$$\mathbf{H}\mathbf{t}_k = \mathbf{S}\mathbf{t}_k E_k \quad (2)$$

We can at this stage note that in addition to being variational, the NOCIS core-excited state energies are also size-consistent, just as for the standard orthogonal CIS energy. For a supersystem composed of two non-interacting molecules,  $M_1$  and  $M_2$ , the NOCIS core excitation energies on  $M_1$  are unaffected by the presence of  $M_2$ . This property is essential to correctly predict substituent effects on core excitation energies.

FIG. 1. A visual of the Hamiltonian Matrix in the NOCIS scheme; the peach components are the Fock ionized elements and the blue elements are obtained with NOCI. This example is considering a potential excitation from 2 different core orbitals

The algorithm for a NOCIS calculation is as follows:

1. A self-consistent field (SCF) calculation is run on the ground state of the given molecule
2. If more than two core orbitals are to be considered, a Boys Localization<sup>44</sup> is performed on those orbitals (we restrict ourselves to K shells in this discussion though generalization is possible)
3. The molecular orbital (MO) coefficient matrices are rearranged to reflect the removal of an electron from the core orbital of interest

4. A restricted open-shell Hartree-Fock (ROHF) optimization is performed on this cation state without freezing the core orbital
5. Steps 3 and 4 are repeated as necessary for each relevant core orbital
6. The diagonal blocks of the  $H$ ,  $S$ , and  $S^2$  matrices are constructed for each  $i$  (determinants in a block are orthonormal)
7. The off-diagonal elements of the matrices are constructed using NOCI
8. The generalized eigenvalue problem is solved, the SCF ground energy is subtracted from the resulting energies to obtain the excitation energy of each state, and the  $\langle S^2 \rangle$  value for each state can be obtained

The number of core orbitals of a particular type defines the number of references in the calculation (eg. a  $1s$  calculation for  $C_2H_2$  would have 2 references). Figure 1 shows a visual example of a completed NOCIS matrix. As NOCIS calculations have only been implemented so far for closed-shell molecules, those are what are discussed here. For NOCIS, it is necessary to consider both alpha and beta ionization, but, for closed shell reference, it is only necessary to perform the optimization for either alpha or beta.

### A. Diagonal-Block Matrix Elements

In the aforementioned peach blocks of Figure 1, NOCI and CIS calculations produce the same results. These diagonal blocks represent particle states, which are formed by adding an electron to a determinant, and thus they are simply the  $vv$  block of the Fock matrix for ionized states. In the basis of these particle states, the matrix elements of  $H$  are then the matrix elements of the Fock operator, with the addition of the excited ion energy on the diagonal elements.

The  $S$  matrix is simply the identity because these matrix elements are between configurations constructed from the same set of orthonormal MOs. This orthonormality also makes the  $S^2$  matrix simple to compute.



TABLE I. Deviation of excitation energies ( $\omega_{ex}$ ) of the first three singlet core-excited states of CO (C and O 1s) and NH3 (N 1s) calculated using NOCIS and different basis sets, compared with results from the aug-cc-pCVQZ basis set.

Basis set	CO C 1s			CO O 1s			NH3 N 1s		
	States $\omega_{ex}$ (eV)			States $\omega_{ex}$ (eV)			States $\omega_{ex}$ (eV)		
	1	2	3	1	2	3	1	2	3
cc-pVDZ	-2.36	-7.88	-11.46	-2.62	-7.74	-13.30	-3.63	-4.08	-14.62
cc-pVTZ	-0.42	-3.62	-5.59	-0.54	-3.68	-6.49	-1.09	-1.58	-6.96
cc-pVQZ	-0.13	-2.25	-3.44	-0.14	-2.29	-3.79	-0.60	-1.03	-4.81
cc-pV5Z	-0.03	-1.31	-1.58	-0.02	-1.36	-1.77	-0.33	-0.61	-2.25
aug-cc-pVDZ	-2.24	-2.62	-2.71	-2.66	-2.91	-2.99	-2.37	-2.42	-2.93
aug-cc-pVTZ	-0.29	-0.35	-0.38	-0.47	-0.52	-0.56	-0.30	-0.31	-0.51
aug-cc-pVQZ	-0.11	-0.10	-0.11	-0.13	-0.12	-0.13	-0.09	-0.09	-0.09
aug-cc-pV5Z	-0.03	0.03	0.06	-0.02	0.05	0.07	-	-	-
d-aug-cc-pVDZ	-2.20	-2.29	-2.25	-2.66	-2.60	-2.53	-2.33	-2.79	-13.30
d-aug-cc-pVTZ	-0.24	-0.11	-0.06	-0.45	-0.29	-0.23	-0.29	-0.28	0.14
cc-pCVDZ	-1.04	-6.42	-10.06	-1.27	-6.32	-11.90	-2.33	-2.79	-13.30
cc-pCVTZ	-0.10	-3.27	-5.26	-0.14	-3.26	-5.96	-0.81	-1.29	-6.66
cc-pCVQZ	-0.02	-2.13	-3.31	-0.01	-2.15	-3.64	-0.50	-0.93	-4.71
aug-cc-pCVDZ	-0.95	-1.23	-1.33	-1.30	-1.50	-1.60	-1.07	-1.12	-1.63
aug-cc-pCVTZ	-0.08	-0.13	-0.16	-0.12	-0.18	-0.21	-0.07	-0.08	-0.28
d-aug-cc-pCVDZ	-0.92	-0.91	-0.86	-1.30	-1.20	-1.14	-1.06	-1.06	-0.65
d-aug-cc-pCVTZ	-0.07	0.07	0.11	-0.12	0.04	0.10	-0.07	-0.06	0.36
aug-cc-pCVQZ	289.02	294.01	294.96	534.19	538.18	539.20	401.01	402.51	403.37

## B. NOCI Matrix Elements

The main substance of NOCI can be found in References 39 and 46, but the theory is summarized here. The intention is to build and diagonalize the Hamiltonian in a basis of non-orthogonal Slater determinants, or basis states. This renders the equation a generalized

eigenvalue problem. The mathematical method for computation of the overlap and Hamiltonian between two generally non-orthogonal basis states is described in detail by Thom and Head-Gordon<sup>46</sup>.

To obtain the orbitals, we use the orbitals of Amos and Hall<sup>47,48</sup>, alternately called Löwdin-Paired orbitals<sup>49</sup>. After these biorthogonal states have been obtained, generalized Slater-Condon rules<sup>46,47</sup> are used to compute the  $H$  matrix elements and the generalized eigenvalue problem is solved.

One further element of note is that the paper of Sundstrom and Head-Gordon<sup>39</sup> contains an error in the derivation of the  $S^2$  matrix elements. This was corrected by Levine and Head-Gordon<sup>50</sup>.

### C. Computational Implementation

The code for NOCIS was implemented in a development version of Q-Chem 5.0<sup>51</sup> utilizing the Armadillo linear algebra library<sup>52</sup>. All molecular geometries were obtained from Curtiss et al.<sup>53</sup> The ROHF optimizations were performed using the Newton Minres algorithm<sup>54</sup>, and the NOCI code was a modified version of previous code developed by Sundstrom<sup>39</sup>. The maximum overlap method (MOM) is used to keep the core-hole from collapsing<sup>55</sup>.

The approximate scaling of this code is  $N^4$ . Each NOCI element requires a 2-electron integral build, which is a source of computational bottleneck. One solution, which has been implemented, is to calculate these integrals in batches instead of individually, but this also causes a significant increase in memory use. The code can be further accelerated by noting that many matrix elements share a common orbital basis, but this remains a further project. However, even in its current unoptimized state, NOCIS does not use any large intermediates that require disk storage, an advantage over ADC and EOM-CCSD methods, where the unknowns are as large as fourth-rank tensors.

## III. RESULTS

The NOCIS calculations were tested using both ROHF and UHF methods for the cation optimization. Although the UHF method provided slightly more accurate energies, they were not spin-pure. Since  $S^2$  values are vital to identifying XAS states, we decided to move



TABLE II. Comparisons of K-edge calculation errors for several molecules using CIS,  $\Delta$ -SCF,  $\Delta$ -DFT, and NOCIS with experimental values, along with the mean signed error (MSE) and root-mean-squared error (RMSE). Calculations were done in the aug-cc-pCVTZ<sup>58-60</sup> basis and all results are in eV.

Molecule	Experiment	CIS	$\Delta$ -SCF	$\Delta$ -DFT <sup>a</sup>	$\Delta$ -DFT <sup>b</sup>	TDDFT <sup>c</sup>	NOCIS
C2H4 C(1s)	284.7 <sup>61</sup>	15.58	10.44	-4.50	2.10	3.26	1.70
C2H2 C(1s)	285.9 <sup>61</sup>	16.68	9.97	-1.93	4.62	2.49	1.26
H2CO C(1s)	286 <sup>63</sup>	20.29	4.94	-0.70	0.71	-0.48	2.01
C2N2 C(1s)	286.3 <sup>73</sup>	10.07	7.07	-5.05	1.50	0.17	1.75
HCN C(1s)	286.37 <sup>68</sup>	17.07	2.77	-0.46	0.93	0.11	1.71
C2H6 C(1s)	286.9 <sup>61</sup>	12.86	7.48	-4.88	1.70	0.45	0.52
CO C(1s)	287.4 <sup>66</sup>	19.14	-0.34	-1.01	0.40	-1.34	1.70
CH3OH C(1s)	287.92 <sup>67</sup>	17.15	0.70	0.07	1.59	0.58	1.07
CH4 C(1s)	288.1 <sup>62</sup>	14.68	-1.17	-1.60	-0.13	0.03	-0.83
CO2 C(1s)	290.8 <sup>64</sup>	18.56	3.90	1.48	3.46	-1.06	2.12
C2N2 N(1s)	398.9 <sup>73</sup>	12.73	10.74	-5.45	2.26	0.38	1.13
HCN N(1s)	399.7 <sup>68</sup>	21.63	2.59	-0.71	0.64	0.21	1.04
NH3 N(1s)	400.8 <sup>69</sup>	15.29	-0.13	-0.60	0.84	0.72	0.30
N2 N(1s)	400.96 <sup>70</sup>	23.93	16.45	-1.07	1.77	-0.45	1.16
N2O Nt(1s)	401.1 <sup>65</sup>	22.04	3.09	-1.11	0.19	0.10	1.09
N2O Nc(1s)	404.8 <sup>65</sup>	22.28	5.12	-1.52	-0.19	-0.70	1.14
H2CO O(1s)	530.8 <sup>63</sup>	25.53	3.65	-0.75	0.35	0.07	0.70
H2O O(1s)	534 <sup>69</sup>	22.17	-0.38	-0.76	0.52	0.48	0.15
CH3OH O(1s)	534.07 <sup>67</sup>	21.91	-0.40	-0.62	0.64	0.73	0.22
CO O(1s)	534.2 <sup>66</sup>	25.76	-1.44	-1.18	-0.14	0.10	0.11
N2O O(1s)	535 <sup>71</sup>	22.41	1.20	-1.30	-0.30	0.58	-0.05
CO2 O(1s)	535.3 <sup>71</sup>	18.75	12.28	-0.29	1.04	1.79	0.80
HF F(1s)	682.2 <sup>72</sup>	29.52	4.20	4.16	5.20	4.40	4.87
F2 F(1s)	686.5 <sup>72</sup>	10.66	7.17	-5.32	-2.72	-5.45	-2.90
<b>MSE</b>		<b>19.03</b>	<b>4.58</b>	<b>-1.46</b>	<b>1.12</b>	<b>0.30</b>	<b>0.95</b>
<b>RMSE</b>		<b>19.64</b>	<b>6.55</b>	<b>2.62</b>	<b>1.97</b>	<b>1.77</b>	<b>1.63</b>

<sup>a</sup> B3LYP

<sup>b</sup> SRC1-R1

TABLE III. A comparison of NOCIS values with the ADC calculations of Reference 28 for the first three excited states, all done in the aug-cc-pVQZ basis.<sup>74,75</sup>

Basis set	CO C 1s			CO O 1s			NH3 N 1s		
	States $\omega_{ex}$ (eV)			States $\omega_{ex}$ (eV)			States $\omega_{ex}$ (eV)		
	1	2	3	1	2	3	1	2	3
NOCIS	289.13	294.11	295.07	534.32	538.31	539.33	401.10	402.60	403.46
<i>Error</i>	1.73	1.61	1.67	0.22	0.49	0.47	0.44	0.27	0.60
ADC-CVS(3)	287.80	293.75	294.83	538.38	544.32	545.86	403.61	405.17	406.17
<i>Error</i>	0.4	1.25	1.43	4.28	5.52	6.06	2.95	2.84	3.31
Experiment	287.4 <sup>64</sup>	292.5 <sup>64</sup>	293.4 <sup>64</sup>	534.1 <sup>64</sup>	538.8 <sup>64</sup>	539.8 <sup>64</sup>	400.66 <sup>69</sup>	402.33 <sup>69</sup>	402.86 <sup>69</sup>

forward with the spin-pure ROHF calculations.

We first assess the basis set convergence of NOCIS calculations on the 3 lowest allowed singlet core-excited states of CO (1s C and 1s O) and NH3. The results are shown in Table 1 for the standard sequence of Dunning correlation consistent basis sets<sup>56</sup> as well as the sequences augmented with diffuse and/or tight functions. The use of diffuse functions, as expected, becomes more important for states higher than the lowest, which contain some Rydberg character in the particle levels that become partly occupied in the transition. Use of the CV augmentation, on the other hand, helps accelerate basis set convergence of the core hole. The largest improvement, however, is when both an augmented and core-polarized basis set is used. This is reasonable, as including core correlated functions as well as diffuse functions should result in a more accurate result for a core excitation. Looking more in-depth at the aug-cc-pcVXZ basis sets, the balance of error reduction and efficiency suggests that a good basis set for both efficiency and accuracy is aug-cc-pCVTZ. These conclusions are generally consistent with the recent basis set assessment of Fouda and Besley<sup>57</sup>.

Having selected the aug-cc-pCVTZ basis, we next consider the comparison of the NOCIS method to three other methods, as shown in Table 2, which shows the results of CIS,  $\Delta$ -SCF,  $\Delta$ -DFT, TDDFT, and NOCIS calculations expressed as deviations from experimental values. As expected, the CIS method performs very poorly due to its lack of orbital relaxation and correlation, with a RMS error of 19.64 eV. The  $\Delta$ -SCF method, which includes orbital relaxation but not dynamic correlation, is a large improvement on CIS methods, with a

RMS error of 6.55 eV. Notably,  $\Delta$ -SCF calculations perform worse in cases such as  $N_2$  where there is a problem of static correlation as a result of core orbitals with the same energy. Static correlation will be discussed in more depth later.

The  $\Delta$ -DFT methods, which include both orbital relaxation and dynamic correlation, perform even better, as expected, with RMS errors of 2.62 eV and 1.97 eV respectively for B3LYP and SRC1-R1. The TDDFT methods perform extremely well, with an RMS error of 1.77. This is likely due to the fact that the SRC1-R1 functional<sup>76</sup> was specifically designed for core excitations in TDDFT. This result shows the importance of correlation in obtaining an accurate core-excitation energy. However, TDDFT is extremely functional dependent, and performing it with other functionals produces results that vary considerably in accuracy.

The NOCIS calculations, which include orbital relaxation and static correlation but not dynamic correlation, perform remarkably well, even when compared to the DFT methods mentioned above, outperforming both  $\Delta$  and TDDFT in RMSE. One particular area of strength for this calculation is its ability to account for static correlation, specifically excitations from core orbitals of the same type, as in  $C_2H_2$ . This is accomplished through Boys localization<sup>44</sup> on the relevant core orbitals, and the non-orthogonal interactions between them. The addition of this element significantly improved the calculation results, often by several eV. In the cc-pCVDZ basis set, the errors on  $C_2H_2$ ,  $N_2$ ,  $CO_2$ ,  $O$ , and  $F_2$  were reduced by 6.53, 8.03, 10.24, and 7.61 eV, respectively relative to non-localized NOCIS calculations. The NOCIS results together with  $\Delta$ -SCF and standard CIS give interesting insight into the relative roles of orbital relaxation and configuration interaction in core excitation energies. It is clear that both are essential in cases where there are multiple equivalent nuclei, although the orbital relaxation effect is dominant in cases where only a singly nucleus is core-excited (eg.  $H_2CO$ ). Cases where  $\Delta$ -SCF performs better than NOCIS may well be due to fortuitous error cancellation. A further note is that none of these methods include relativistic effects, and thus the larger the atom, the more poorly the method performs, as shown in the  $F_2$  and HF calculations.

Table 3 features NOCIS calculations compared with ADC calculations from Reference 28. While ADC performed better with the  $CO\ C\ 1s$ , NOCIS performed better for the  $CO\ O\ 1s$  and the  $NH_3\ N\ 1s$ , which is consistent with the fact that ADC has been shown to overestimate energies for heavier atoms<sup>28</sup>. This shows that NOCIS is at least competitive with ADC, and with its smaller scaling, more efficient. Additionally, in some smaller basis

	State 1	State 2	State 3
CIS	0.280	0.015	0.014
ADC-CVS(3) <sup>28</sup>	0.159	0.005	0.013
NOCIS	0.155	0.003	0.009
Experiment <sup>77</sup>	0.167	0.003	0.008

TABLE IV. A comparison of the oscillator strengths of the first 3 singlet excited states of the C 1s orbital in CO, all done in the aug-cc-pVQZ basis,<sup>74,75</sup> between ADC, CIS, NOCIS, and experiment.

sets, ADC does not provide the second and third excited states for these molecules, while NOCIS provides reasonable estimates for these states even in smaller basis sets. The fact that NOCIS, a method without dynamic correlation, is competitive with ADC-CVS(2), which includes it, shows the value of treating orbital relaxation via NOCI. Formulating a dynamic correlation to NOCIS looks to be potentially very worthwhile.

We also verified these k-edges by computing oscillator strengths in a method identical to that of the STEX method<sup>42</sup>. It is worth noting that this has been done in an early version of STEX in the aug-cc-pVTZ basis set<sup>78</sup>. For NOCIS, in all cases, the lowest singlet excitation energy produced the largest oscillator strength, confirming our results. The results are shown in table 4. The NOCIS results are competitive with or better than the ADC results, and vastly out-perform the CIS results. We believe that this once again illustrates the importance of orbital relaxation.

#### IV. CONCLUSIONS

This paper describes and benchmarks a new NOCI method specifically designed for core-excited states. Unlike standard NOCI, this non-orthogonal CIS method is well-defined, using a precise number of configurations that depends only on the molecular composition and the basis set. This method compares favorably to all current methods of calculating core excitations in both efficiency and accuracy. However, the implementation is still preliminary, and some work is needed before we can claim a truly efficient method for the robust identification of XAS peaks. In this vein, a further effort for this project is to construct output in such a way that it can be used to identify spectroscopic peaks. This will be done by implementing

oscillator strength calculations as well as explicitly connecting orbital transitions to their energies.

There is also further theoretical and computational work to be done on NOCIS. As previously mentioned, it does not contain dynamic correlation, so a possible next step would be implementing something like NOCI-MP2<sup>41</sup> in the construction of the matrix elements. A different direction is to include relativistic effects in NOCI calculations so that it can better calculate core-excitations for larger atoms such as silicon, which is beyond the capability of non-relativistic core-excited calculations. Computationally, the code can be sped up by implementing batching for the 2-electron integrals as described in the previous section, which reduces the prefactor and thus the total computation time.

Overall, NOCIS is a very promising method for calculating core excitations. Even without dynamic correlation, it still has a high level of accuracy, and it has been shown to deal well with static correlation. With multi-state return (in contrast to  $\Delta$  methods) and competitive scaling, NOCIS provides a convenient path to obtain accurate calculations that can be used to simulate an XAS spectrum.

## ACKNOWLEDGEMENTS

KJO thanks the members of the Head-Gordon group for fruitful discussions, particularly Daniel Levine. This work was supported by the Director, Office of Science, Office of Basic Energy Sciences of the US Department of Energy under contract No. DE-AC02-05CH11231.

## REFERENCES

- <sup>1</sup>J. A. Sheehy, T. J. Gil, C. L. Winstead, R. E. Farren, and P. W. Langhoff, *J. Chem. Phys.* **91**, 1796 (1989).
- <sup>2</sup>H. Ågren, V. Carravetta, O. Vahtras, and L. G. M. Pettersson, *Chem. Phys. Lett.* **222**, 75 (1994).
- <sup>3</sup>L. Triguero, L. G. M. Pettersson, and A. Ågren, *Phys. Rev. B* **58**, 8097 (1998).
- <sup>4</sup>J.B. Foresman, M. Head-Gordon, J.A. Pople, M.J. Frisch, *J. Phys. Chem.* **96**, 135 (1992).
- <sup>5</sup>M. Head-Gordon, A. M. Graña, D. Maurice, and C. A. White, *J. Phys. Chem.* **99**, 14261 (1995).



- <sup>6</sup>A. Douhal and J. Santamaria, *Femtochemistry and Femtobiology: Ultrafast Dynamics in Molecular Science*; World Scientific: Singapore, 2002.
- <sup>7</sup>M. Head-Gordon, R. J. Rico, M. Oumi, and T. J. Lee, *Chem. Phys. Lett.* **219**, 21 (1994).
- <sup>8</sup>M. Head-Gordon, D. Maurice, and M. Oumi, *Chem. Phys. Lett.* **246**, 114 (1995).
- <sup>9</sup>F. A. Asmuruf, and N. A. Besley, *Chem. Phys. Lett.* **463**, 267 (2008).
- <sup>10</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- <sup>11</sup>J. M. Herbert, X. Zhang, A. F. Morrison, J. Liu. *Acc. Chem. Res.* **49**, 931 (2016).
- <sup>12</sup>Stener, G. Fronzoni and M. de Simone, *Chem. Phys. Lett.* **373**, 115 (2003).
- <sup>13</sup>N. A. Besley and A. Noble, *J. Phys. Chem. C* **111**, 3333 (2007).
- <sup>14</sup>S. D. George, T. Petrenko and F. Neese, *Inorg. Chim. Acta* **361**, 965 (2008).
- <sup>15</sup>U. Ekström and P. Norman, *Phys. Rev. A: At., Mol., Opt. Phys.* **74**, 042722 (2006).
- <sup>16</sup>U. Ekström, P. Norman, V. Carravetta and H. Ågren, *Phys. Rev. Lett.* **97**, 143001 (2006).
- <sup>17</sup>H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao, *J. Chem. Phys.* **115**, 3540 (2001).
- <sup>18</sup>N. A. Besley, A. T. B. Gilbert, and P. M. W. Gill, *J. Chem. Phys.* **130**, 124308 (2009).
- <sup>19</sup>J. A. Pople, P. M. W. Gill, and B. G. Johnson, *Chem. Phys. Lett.* **199**, 557 (1992).
- <sup>20</sup>A. D. Boese and J. M. L. Martin, *J. Chem. Phys.* **121**, 3405 (2004).
- <sup>21</sup>A. Dreuw and M. Wormit, *WIREs Comput. Mol. Sci.* **5** 82 (2015).
- <sup>22</sup>J. Schirmer, *Phys. Rev. A: At., Mol., Opt. Phys.* **26** 2395 (1982).
- <sup>23</sup>S. Cederbaum, W. Domcke, and J. Schirmer, *Phys. Rev. A* **22**, 206 (1980).
- <sup>24</sup>Barth and L. S. Cederbaum, *Phys. Rev. A* **23**, 1038 (1981).
- <sup>25</sup>Barth and J. Schirmer, *J. Phys. B: At. Mol. Phys.* **18**, 867 (1985).
- <sup>26</sup>Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. C. Prince, A. B. Trofimov, E. V. Gromov, I. L. Zaytseva, and J. Schirmer, *Chem. Phys.* **347**, 360 (2008).
- <sup>27</sup>Feyer, O. Plekan, R. Richter, M. Coreno, G. Vall-llosera, K. C. Prince, A. B. Trofimov, I. L. Zaytseva, T. E. Moskovskaya, E. V. Gromov, and J. Schirmer, *J. Phys. Chem. A* **113**, 5736 (2009).
- <sup>28</sup>J. Wenzel, A. Holzer, M. Wormit, and A. Dreuw, *J. Chem. Phys.* **142**, 214104 (2015).
- <sup>29</sup>H. J. Monkhorst, *Int. J. Quantum Chem., Symp.* **11**, 421 (1977).
- <sup>30</sup>A. Krylov, *Annu. Rev. Phys. Chem.* **59**, 433 (2008).
- <sup>31</sup>E. A. Salter, H. Sekino, and R. J. Bartlett, *J. Chem. Phys.* **87**, 502 (1987).
- <sup>32</sup>M. Nooijen and R. J. Bartlett, *J. Chem. Phys.* **102**, 2629 (1995).
- <sup>33</sup>M. Roemelt, D. Maganas, S. DeBeet, and F. Neese, *J. Chem. Phys.* **138**, 204101 (2013).



- <sup>34</sup>G. Coriani, O. Christiansen, T. Fransson, and P. Norman, *Phys. Rev. A* **85**, 022507 (2012).
- <sup>35</sup>R. Broer and W.C. Nieuwpoort, *Chem. Phys.* **54**, 291 (1981).
- <sup>36</sup>R. Broer and W.C. Nieuwpoort, *Theor. Chim. Acta* **73**, 405 (1988).
- <sup>37</sup>R. Broer and W.C. Nieuwpoort, *J. Mol. Struct.-Theochem* **458**, 19 (1999).
- <sup>38</sup>R. Broer, L. Hozoi, and W. C. Nieuwpoort, *Mol. Phys.* **101** 233 (2009).
- <sup>39</sup>E. J. Sundstrom and M. Head-Gordon, *J. Chem. Phys.* **140**, 114103 (2014).
- <sup>40</sup>N. J. Mayhall, P. R. Horn, E. J. Sundstrom, and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **16**, 22694 (2014).
- <sup>41</sup>S. R. Yost and M. Head-Gordon, *J. Chem. Phys.* **145**, 054105 (2016).
- <sup>42</sup>H. Ågren, V. Carravetta, O. Vahtras, and L. G. M. Pettersson, *Theor. Chem. Acc.* **97**, 14 (1997).
- <sup>43</sup>U. Ekström, P. Norman, V. Carravetta, *Phys. Rev. A* **73**, 022501 (2006).
- <sup>44</sup>S. F. Boys, *Rev. Mod. Phys.* **32**, 296 (1960).
- <sup>45</sup>A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications: Mineola, 1996.
- <sup>46</sup>A. J. W. Thom and M. Head-Gordon, *J. Chem. Phys.* **131**, 124113 (2009).
- <sup>47</sup>A. T. Amos and G. G. Hall, *Proc. R. Soc. A* **263**, 483 (1961).
- <sup>48</sup>H. F. King, R. E. Stanton, H. Kim, R. E. Wyatt, and R. G. Parr, *J. Chem. Phys.* **47**, 1936 (1967).
- <sup>49</sup>P.-O. Löwdin, *J. Appl. Phys.* **33**, 251 (1962).
- <sup>50</sup>D. S. Levine, M. Head-Gordon, *PNAS* **114**, 12649 (2017).
- <sup>51</sup>Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, T. Kus, A. Landau, J. Liu, E. I. Proynov, R. M. Richard, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, S. A. Baeppler, Z. Benda, Y. A. Bernard, E. J. Berquist, K. B. Bravaya, D. Casanova, C. Chang, Y. Chen, A. Chien, K. D. Closser, M. P. Coons, S. Dasgupta, M. Diedenhofen, H. Do, R. G. Edgar, P. Fang, S. Fatehi, Q. Feng, Q. Ge, A. Ghysels, J. Gomes, J. Gonthier, A. Gunina, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, J. E. Herr, E. G. Hohenstein, Z. C. Holden, Kerwin Hui, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyayev, J. Kim, P. Klunzinger, K. Koh, D. Kosenkov, T. Kowalczyk, C. M. Krauter, A. Kunitsa, Ka Un Lao, A. Laurent, K. V. Lawler, J. Lee, D. Lefrancois, S. Lehtola, D. S. Levine, Y. Li,

- Y. Lin, F. Liu, E. Livshits, A. Luenser, P. Manohar, E. Mansoor, S. F. Manzer, S. Mao, Y. Mao, N. Mardirossian, A. V. Marenich, L. A. Martinez-Martinez, S. A. Maurer, N. J. Mayhall, S. C. McKenzie, J.-M. Mewes, A. F. Morrison, K. Nanda, T. S. Nguyen-Beck, R. Olivares-Amaya, J. A. Parkhill, T. M. Perrine, F. Plasser, S. Prager, A. Prociuk, D. R. Rehn, F. Rob, M. Schneider, N. Sergueev, S. M. Sharada, S. Sharma, D. W. Small, T. Stauch, T. Stein, Y. Su, A. J. W. Thom, T. Tsuchimochi, L. Vogt, O. Vydrov, M. A. Watson, J. Wenzel, M. de Wergifosse, T. A. Wesolowski, A. White, J. Witte, A. Yamada, J. Yang, K. Yao, S. Yeganeh, S. R. Yost, Z. You, A. Zech, I. Y. Zhang, X. Zhang, Y. Zhao, Y. Zhu, B. R. Brooks, G. K. L. Chan, C. J. Cramer, M. S. Gordon, W. J. Hehre, A. Klamt, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, C. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, M. Head-Gordon, *Mol. Phys.* **113**, 184 (2015).
- <sup>52</sup>C. Sanderson, *Technical Report, NICTA* 1 (2010).
- <sup>53</sup>L. A. Curtiss, K. Raghavachari, P. Redfern, J. A. Pople, *J. Chem. Phys.* **106**, 1063 (1997).
- <sup>54</sup>R. Borsdorf and N Higham, *IMA J. Numer. Anal.* **30**, 94 (2010).
- <sup>55</sup>A. T. B. Gilbert, N. A. Besley, and P. M. W. Gil, *J. Phys. Chem. A* **112**, 13164 (2008).
- <sup>56</sup>T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>57</sup>A. E. A. Fouda and N. A. Besley, *Theor. Chem. Acc.* **137**, 6 (2017).
- <sup>58</sup>T.H. Dunning, Jr. *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>59</sup>R.A. Kendall, T.H. Dunning, Jr. and R.J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- <sup>60</sup>D.E. Woon and T.H. Dunning, Jr. *J. Chem. Phys.* **98**, 1358 (1993).
- <sup>61</sup>A. P. Hitchcock and C. E. Brion, *J. Electron. Spectrosc. Relat. Phenom.* **10**, 317 (1977).
- <sup>62</sup>F. C. Brown, R. Z. Bachrach, and A. Bianconi, *Chem. Phys. Lett.* **54**, 425 (1978).
- <sup>63</sup>G. Remmers, M. Domke, A. Puschmann, T. Mandel, C. Xue, G. Kaindl, E. Hudson, and D. A. Shirley, *Phys. Rev. A* **46**, 3935 (1992).
- <sup>64</sup>M. Tronc, G. C. King and F. H. Read, *J. Phys. B: At. Mol. Phys.* **12**, 137 (1979).
- <sup>65</sup>M. Tronc, G. C. King, and F. H. Read, *J. Phys. B: At. Mol. Phys.* **13**, 999 (1980).
- <sup>66</sup>M. Domke, A. Puschmann, T. Mandel, E. Hudson, D. A. Shirley, and G. Kaindl, *Chem. Phys. Lett.* **173**, 122 (1990).

- <sup>67</sup>A. Hempelmann, M. N. Piancastelli, F. Heiser, O. Cessner, A. Rüdél, and U. Becker, *J. Phys. B: At. Mol. Opt. Phys.* **32**, 2677 (1999).
- <sup>68</sup>A. P. Hitchcock and C. E. Brion, *Chem. Phys.*, **37**, 319 (1979).
- <sup>69</sup>J. Schirmer, A. B. Trofimov, K. J. Randall, J. Feldhaus, A. M. Bradshaw, Y. Ma, C. T. Chen, and F. Sette, *Phys. Rev. A* **47**, 1136 (1993).
- <sup>70</sup>A. P. Hitchcock and C. E. Brion, *J. Electron. Spectrosc. Relat. Phenom.* **18**, 1 (1980).
- <sup>71</sup>D. M. Barrus, R. L. Blake, A. J. Burek, K. C. Chambers, and A. L. Pregonzer, *Phys. Rev. A* **20**, 1045 (1979).
- <sup>72</sup>A. P. Hitchcock and C. E. Brion, *J. Phys. B: At. Mol. Phys.* **14**, 4399 (1981).
- <sup>73</sup>A. P. Hitchcock and C. E. Brion, *Chem. Phys.* **37**, 319 (1979).
- <sup>74</sup>T.H. Dunning, Jr. *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>75</sup>D.E. Woon and T.H. Dunning, Jr. *J. Chem. Phys.* **98**, 1358 (1993).
- <sup>76</sup>N. A. Besley, M. J. G. Peach, and D. J. Tozer, *Phys. Chem. Chem. Phys.* **11**, 10350 (2009).
- <sup>77</sup>M. Tronc, G. C. King, and F. H. Read, *J. Phys. B: At. Mol. Phys.* **12**, 137 (1979).
- <sup>78</sup>L. Yang, H. Ågren, V. Carravetta, and L. G. M. Pettersson, *Phys. Script.* **54**, 614 (1996).

A 4x4 grid diagram with labels  $1\alpha$ ,  $1\beta$ ,  $2\alpha$ , and  $2\beta$  on both the top and left sides. The grid is divided into four quadrants by the first two rows and columns. The diagonal cells, representing the intersection of the same labels (e.g.,  $1\alpha$  and  $1\alpha$ ), are shaded orange. All other cells are shaded light blue. A large, semi-transparent watermark reading "ACCEPTED" is overlaid diagonally across the grid.

	$1\alpha$	$1\beta$	$2\alpha$	$2\beta$
$1\alpha$	Orange	Light Blue	Light Blue	Light Blue
$1\beta$	Light Blue	Orange	Light Blue	Light Blue
$2\alpha$	Light Blue	Light Blue	Orange	Light Blue
$2\beta$	Light Blue	Light Blue	Light Blue	Orange