Simulation of X-Ray Absorption Spectra with Orthogonality Constrained Density Functional Theory *"The history of science is rich in the example of the fruitfulness of bringing two sets of techniques, two sets of ideas, developed in separate contexts for the pursuit of new truth, into touch with one another."* J. Robert Oppenheimer Orthogonality constrained density functional theory (OCDFT) [F. A. Evangelista, P. Shushkov and J. C. Tully, *J. Phys. Chem. A*, 2013, **117**, 7378] is a variational time-independent approach for the computation of electronic excited states. In this work we extend OCDFT to compute core-excited states and generalize the original formalism to determine multiple excited states. Benchmark computations on a set of 13 small molecules and 40 excited states show that *unshifted* OCDFT/B3LYP excitation energies have a mean absolute error of 1.0 eV. Contrary to time-dependent DFT, OCDFT excitation energies for first- and second-row elements are computed with near-uniform accuracy. OCDFT core excitation energies are insensitive to the choice of the functional and the amount of Hartree–Fock exchange. We show that OCDFT is a powerful tool for the assignment of X-ray absorption spectra of large molecules by simulating the gas-phase near-edge spectrum of adenine and thymine.

1 Introduction

The advent of synchrotron light sources created a strong resurgence of spectroscopy in the X-ray region.[1] Near-edge X-ray absorption spectroscopy (NEXAS) is a useful experimental technique to probe the local electronic and geometrical structure in a variety of molecular environments. The most dominant feature of NEXAS spectra, the *near edge* (see Fig. 2.1Example of a X-ray photoabsorption spectrum (XAS). The near edge, located in the low energy region, consists of excitations of core electrons to valence orbitals. These transitions are sensitive to the chemical environment surrounding the excited atom. The high energy region of the spectrum results from excitations of core electrons to the continuumfigure.2.1), is composed of the excitations of core-electrons to unoccupied valence orbitals. Core excitations are atom-specific and sensitive to the local chemical environment, thus, NEXAS spectra can provide information about the chemical composition and the electronic structure of molecules. NEXAS has been successfully applied to large biological systems,[2] small molecules in the gas phase,[3] organic thin-films,[4] and semiconducting materials.[5] This wide range of applications is possible because synchrotron light sources can span an energy range that goes from a few electron volt (eV)[6] to hundreds of MeV.[7]

As NEXAS experiments are becoming more feasible, there is a growing need to develop accurate theoretical approaches
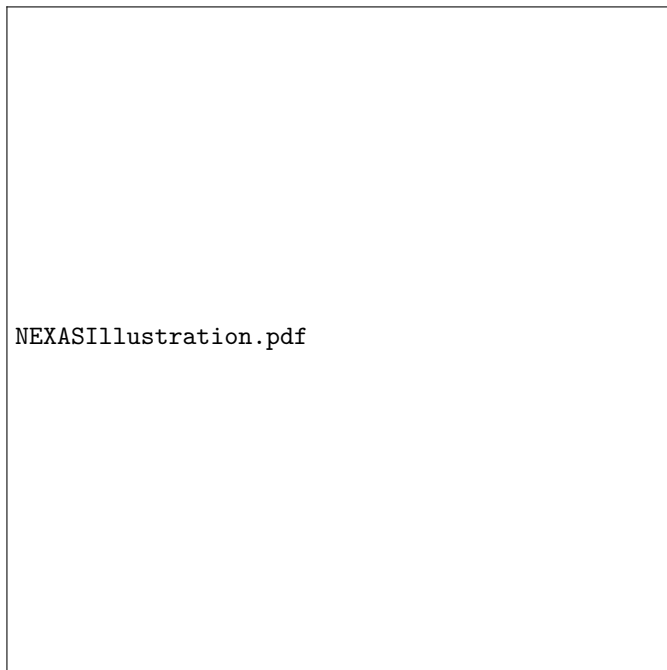

Figure 1: Example of a X-ray photoabsorption spectrum (XAS). The near edge, located in the low energy region, consists of excitations of core electrons to valence orbitals. These transitions are sensitive to the chemical environment surrounding the excited atom. The high energy region of the spectrum results from excitations of core electrons to the continuum.

to aid the interpretation of experimental spectra. Calculations of NEXAS spectra are challenging, and require computational methods that explicitly account for the excitations of core-level electrons, orbital relaxation effects, and electron correlation.[8] Several theoretical approaches have been adapted to compute core-valence excitations, including: scaled-opposite-spin configuration interaction singles with perturbative doubles [SOS-CIS(D)],[9] a restricted open-shell DFT/CIS method,[10,11] second-order algebraic digrammatic construction [ADC(2)],[12,13] multiple scattering X_α methods,[14] a maximum overlap ΔSCF approach,[15] restricted active space SCF method (RASSCF),[16] transition potential theory,[17] coupled-cluster response theory,[8] time-dependent density functional theory (TDDFT),[18] and restricted excitation window TDDFT (REW-TDDFT).[19] Among these methods, TDDFT is perhaps the most attractive option because of its reduced computational cost and ability to calculate multiple excited states.

TDDFT is a rigorous extension of the DFT ground-state formalism,[20] and it is regarded as the method of choice to treat electronic excited states within a density functional framework. When applied in conjunction with frequency-independent exchange-correlation potentials, TDDFT yields accurate excitation energies for low-lying excited states. For example, the TDDFT benchmark study of Silva-Junior and co-workers[21] on 28 organic molecules, showed that singlet and triplet excitation energies can be calculated with a mean average error (MAE) of 0.27 eV and 0.44 eV, respectively. However, Besley et al.[15] showed that TDDFT core excitations computed with conventional exchange-correlation functionals grossly underestimate experimental results, yielding a MAE of 20.2 eV. It is customary to remedy this deficiency of TDDFT by shifting the position of computed spectra by an amount that minimizes the difference between the computational and experimental peak features. For example, a study done by DeBeer, Petrenko, and Neese[22] showed that it is necessary shift the TDDFT Fe near-edge spectrum of different iron complexes by 171.3 eV in order to achieve quantitative agreement with experiment. This shift is dependent on the basis set and functional used and must be recalibrated for every basis set/functional combination applied to a given system.

Work by Peach et al.[23] provided evidence for a direct correlation between the accuracy of TDDFT and the degree of spatial overlap (Λ) between the occupied and virtual orbitals. The inaccuracy of TDDFT for excitations with low values of Λ has been attributed to the incorrect asymptotic behavior of the exchange-correlation potential and self-interaction error.[23] Low orbital overlap is a characteristic feature of both charge-transfer excitations and core excitations, and as expected, the same correlation between accuracy and orbital overlap is observed for core excitations.[24] As a consequence, TDDFT core-valence excitation energies can be improved by introducing self-interaction corrections (SIC)[25] or range separated hybrid functionals in which the amount of long and short range Hartree–Fock exchange is reparametrized.[24,26] However, the optimal value of the range separation parameter and the amount of Hartree–Fock exchange are strongly system dependent and must be tuned.[27–29]

A general method that can systematically produce accurate core-excitation energies with traditional (that is, non range corrected) hybrid density functionals is highly desirable. The maximum overlap method (MOM)[15] combined with a ΔSCF treatment of core-valence excitations is able to obtain highly accurate excitation energies using conventional functionals. However, this procedure is not guaranteed to avoid the problem of variational collapse—albeit MOM ameliorates the difficulties encountered by a straightforward ΔSCF procedure—and has not been generalized to multiple excited states of the same symmetry.

The goal of this work is to find cost-effective alternative theories to TDDFT that can be used to simulate NEXAS spectra. Orthogonality constrained density functional theory (OCDFT)[30] was rigorously derived from a variational time-independent formulation of excited state DFT. It builds upon previous successful efforts to formulate variational excited state DFT, such as: the ΔSCF procedure,[31,32] constrained DFT,[33] stationary state DFT,[34] constricted variational density functional theory (CV-DFT),[35–38] perturbative constrained excited state DFT,[39–41] ensemble DFT,[42–45] and variational time independent DFT (TI-DFT).[46,47] Formally, OCDFT may be viewed as bridging constrained and constricted variational DFT. Its main advantages are: 1) a favorable accuracy/cost ratio, similar to that of ground state DFT, 2) a numerically robust optimization procedure that avoids variational collapse, and 3) the ability to compute spin adapted excitation energies. Benchmark computations[30] show that valence excitation energies computed with OCDFT have error metrics comparable to that of TDDFT. In addition, OCDFT has the ability to accurately compute charge-transfer excitation energies regardless of the amount of Hartree–Fock exchange present in the exchange-correlation functional.

This work introduces two new developments of OCDFT that are necessary for the simulation of near-edge X-ray absorption spectra. First, we formulate an OCDFT algorithm that can be used to compute core-valence excitation energies. This new method is assessed over a test set that includes 13 molecules with 40 unique core-electron excitations. Second, we discuss one approach to extend OCDFT to multiple excited states of the same symmetry. We demonstrate the potential of this new method with computations of the gas-phase near-edge spectrum of adenine and thymine.

2 Theory

In this section we provide a brief summary of orthogonality constrained density functional theory along with the necessary extension to multiple excited states (for the full details of the OCDFT derivation we refer the reader to Ref. 30). OCDFT builds upon the time-independent variational DFT approach developed by Ayers, Levy, and Nagy.[48] Within this framework, each electronic state ($\Psi^{(n)}, n = 0, 1, \ldots$) of a N-electron system has a corresponding density functional ($E^{(n)}[\rho]$) which is a generalization of the ground-state functional of Levy. $E^{(n)}[\rho]$ minimizes the energy expectation value $\Psi \hat{H} \Psi$ imposing two constraints on to the trial wave function Ψ: 1) Ψ must be compatible with the density ρ, and 2) Ψ must be orthogonal to the first $n - 1$ *exact* electronic states, $\{\Psi^{(k)}, k = 1, \ldots, n - 1\}$:

$$E^{(n)}[\rho] = \min_{\substack{\Psi \to \rho \\ \Psi \perp \{\Psi^{(k)}\}}} \Psi \hat{H} \Psi. \tag{1}$$

OCDFT provides a practical realization of this time-independent DFT approach. The first step in the OCDFT derivation consists in defining a generalized Kohn–Sham scheme that, for each electronic state $\Psi^{(n)}$, postulates an auxiliary system of noninteracting electrons with wave function $\Phi^{(n)}$ and density $\rho_s^{(n)}$. The density of state $\Phi^{(n)}$ is assumed to be equal to the density of the exact state $\Psi^{(n)}$, which we indicate with $\rho^{(n)}$. In addition, we assume that the auxiliary wave functions are orthogonal, that is:

$$[1]\Phi^{(m)}\Phi^{(n)} = \delta_{mn}. \tag{2}$$

This condition can be imposed without loss of generality. It avoids the excited state wave functions from collapsing down to the ground state solution, and effectively transfers some of the complexity of the excited state density functionals to the kinetic energy operator. Nevertheless, this variational Kohn–Sham scheme involves energy functionals $[E_{\text{KS}}^{(n)}]$ that contain exchange-correlation contributions $[E_{\text{xc}}^{(n)}]$ *specific* for each excited state. In OCDFT, we invoke an adiabatic approximation similar to the one used in TDDFT, and replace $E_{\text{xc}}^{(n)}$ with the ground state exchange-correlation functional, $E_{\text{xc}}^{(0)}$. The resulting functional for excited state n is given by:

$$\begin{aligned} E_{\text{OCDFT}}^{(n)}[\{\phi_i^{(n)}\}] = &-\frac{1}{2} \sum_i^{\text{occ}} [1]\phi_i^{(n)} \nabla^2 [1]\phi_i^{(n)} + \int d\mathbf{r}\, v(\mathbf{r})\rho^{(n)}(\mathbf{r}) \\ &+ J[\rho^{(n)}] + E_{\text{xc}}^{(0)}[\rho^{(n)}]. \end{aligned} \tag{3}$$

Minimization of the $E^{(n)}[\{\phi_i^{(n)}\}]$ with respect to the occupied orbitals for state n can be performed with a modified self-consistent-field algorithm.[30] In the case of the first excited ($n = 1$), it is possible to show that the orthogonality condition [Eq. (2.2Theoryequation.2.2.2)] implies the existence of two special orbitals. As illustrated in Fig. 2.2Illustration of the projection operators used in OCDFT. Notice, that a consequence of Eqs. (2.4Theoryequation.2.2.4)–(2.5Theoryequation.2.2.5) is that the hole $[\phi_{\text{h}}^{(1)}]$ and particle $[\phi_{\text{p}}^{(1)}]$ orbitals are contained in the spaces spanned by $\hat{P}^{(0)}$ and $\hat{Q}^{(0)}$figure.2.2, these are the *hole* $[\phi_{\text{h}}^{(1)}]$ and *particle* $[\phi_{\text{p}}^{(1)}]$ orbitals, which are respectively unoccupied and occupied in the excited state wave function $\Phi^{(1)}$. These orbitals must satisfy the conditions:

$$\hat{Q}^{(0)}\phi_{\text{h}}^{(1)} = 0, \tag{4}$$

$$\hat{P}^{(0)}\phi_{\text{p}}^{(1)} = 0, \tag{5}$$

where $\hat{P}^{(0)} = \sum_i [1]\phi_i^{(0)}[1]\phi_i^{(0)}$ is a projector onto the occupied orbitals of $\Phi^{(0)}$, and $\hat{Q}^{(0)} = 1 - \hat{P}^{(0)}$. Eqs. (2.4Theoryequation.2.2.4) and (2.5Theoryequation.2.2.5) can be enforced via Lagrangian multipliers. Setting the variation of the Lagrangian with respect to the occupied ($\{\phi_i^{(1)}\}$), hole ($\phi_{\text{h}}^{(1)}$), and particle ($\phi_{\text{p}}^{(1)}$) orbitals to zero gives the following eigenvalue equations:

$$(1 - \hat{P}_{\text{h/p}}^{(1)})\hat{f}^{(1)}(1 - \hat{P}_{\text{h/p}}^{(1)})|\phi_i^{(1)}\rangle = \epsilon_i^{(1)}|\phi_i^{(1)}\rangle, \tag{6}$$

$$\hat{P}^{(0)}(1 - \hat{Q}_{\text{s}}^{(1)})\hat{f}^{(1)}(1 - \hat{Q}_{\text{s}}^{(1)})\hat{P}^{(0)}|\phi_{\text{h}}^{(1)}\rangle = \epsilon_{\text{h}}^{(1)}|\phi_{\text{h}}^{(1)}\rangle, \tag{7}$$

$$\hat{Q}^{(0)}(1 - \hat{P}_{\text{s}}^{(1)})\hat{f}^{(1)}(1 - \hat{P}_{\text{s}}^{(1)})\hat{Q}^{(0)}|\phi_{\text{p}}^{(1)}\rangle = \epsilon_{\text{p}}^{(1)}|\phi_{\text{p}}^{(1)}\rangle, \tag{8}$$

where $\hat{f}^{(1)}$ is the Kohn–Sham Hamiltonian of the excited state. Eq. (2.6Theoryequation.2.2.6) determines the occupied orbitals, while Eqs. (2.7Theoryequation.2.2.7) and (2.8Theoryequation.2.2.8) determine the hole and particle

Figure 2: Illustration of the projection operators used in OCDFT. Notice, that a consequence of Eqs. (2.4Theoryequation.2.2.4)–(2.5Theoryequation.2.2.5) is that the hole $[\phi_{\mathrm{h}}^{(1)}]$ and particle $[\phi_{\mathrm{p}}^{(1)}]$ orbitals are contained in the spaces spanned by $\hat{P}^{(0)}$ and $\hat{Q}^{(0)}$.

orbitals, respectively. The projection operators involved in the OCDFT equations are defined as (see Fig. 2.2Illustration of the projection operators used in OCDFT. Notice, that a consequence of Eqs. (2.4Theoryequation.2.2.4)–(2.5Theoryequation.2.2.5) is that the hole $[\phi_{\mathrm{h}}^{(1)}]$ and particle $[\phi_{\mathrm{p}}^{(1)}]$ orbitals are contained in the spaces spanned by $\hat{P}^{(0)}$ and $\hat{Q}^{(0)}$figure.2.2):

$$\hat{P}_{\mathrm{h/p}}^{(1)} = \hat{P}_{\mathrm{h}}^{(1)} + \hat{P}_{\mathrm{p}}^{(1)} = [1]\phi_{\mathrm{h}}^{(1)}[1]\phi_{\mathrm{h}}^{(1)} + [1]\phi_{\mathrm{p}}^{(1)}[1]\phi_{\mathrm{p}}^{(1)}, \tag{9}$$

$$\hat{P}_{\mathrm{s}}^{(1)} = \hat{P}^{(1)} - \hat{P}_{\mathrm{p}}^{(1)}, \tag{10}$$

$$\hat{Q}_{\mathrm{s}}^{(1)} = \hat{Q}^{(1)} - \hat{P}_{\mathrm{h}}^{(1)}, \tag{11}$$

where the subscript "s" stands for *spectator* orbitals. In OCDFT computations of valence excited states, the hole orbital is assumed to be the solution of Eq. (2.7Theoryequation.2.2.7) with the highest value of $\epsilon_{\mathrm{h}}^{(1)}$. Similarly, the particle orbital corresponds to the lowest eigenvalue of Eq. (2.8Theoryequation.2.2.8).

Our OCDFT approach for core-excited states introduces two new aspects. First, in order to compute core-excited states with OCDFT, we select hole orbitals with the smallest values of $\epsilon_{\mathrm{h}}^{(1)}$. However, this simple extension allows us only to compute one core-excited state for each irreducible representation. In principle, OCDFT can be generalized to compute an arbitrary number of excited states. For each additional excited state, it is necessary to minimize the OCDFT energy imposing orthogonality with respect to the lower energy solutions. While this appears to be a viable solution, it would undoubtedly lead to a more elaborate minimization procedure. In this work we propose simplified orthogonality conditions that are based on the orthogonality of the hole and/or particle orbitals. For example, if we choose the hole orbital for the second excited state $[\phi_{\mathrm{h}}^{(2)}]$ to be orthogonal to the first hole $[\phi_{\mathrm{h}}^{(1)}]$ and to spans the occupied space of the ground state determinant:

$$[1]\phi_{\mathrm{h}}^{(2)}\phi_{\mathrm{h}}^{(1)} = 0, \tag{12}$$

$$\hat{Q}^{(0)}\phi_{\mathrm{h}}^{(2)} = 0, \tag{13}$$

then the determinant $\Phi^{(2)}$ is guaranteed to be orthogonal to $\Phi^{(0)}$ and $\Phi^{(1)}$. The conditions Eqs. (2.12Theoryequation.2.2.12) and (2.13Theoryequation.2.2.13) are sufficient but not necessary to guarantee orthogonality among the first three electronic states.
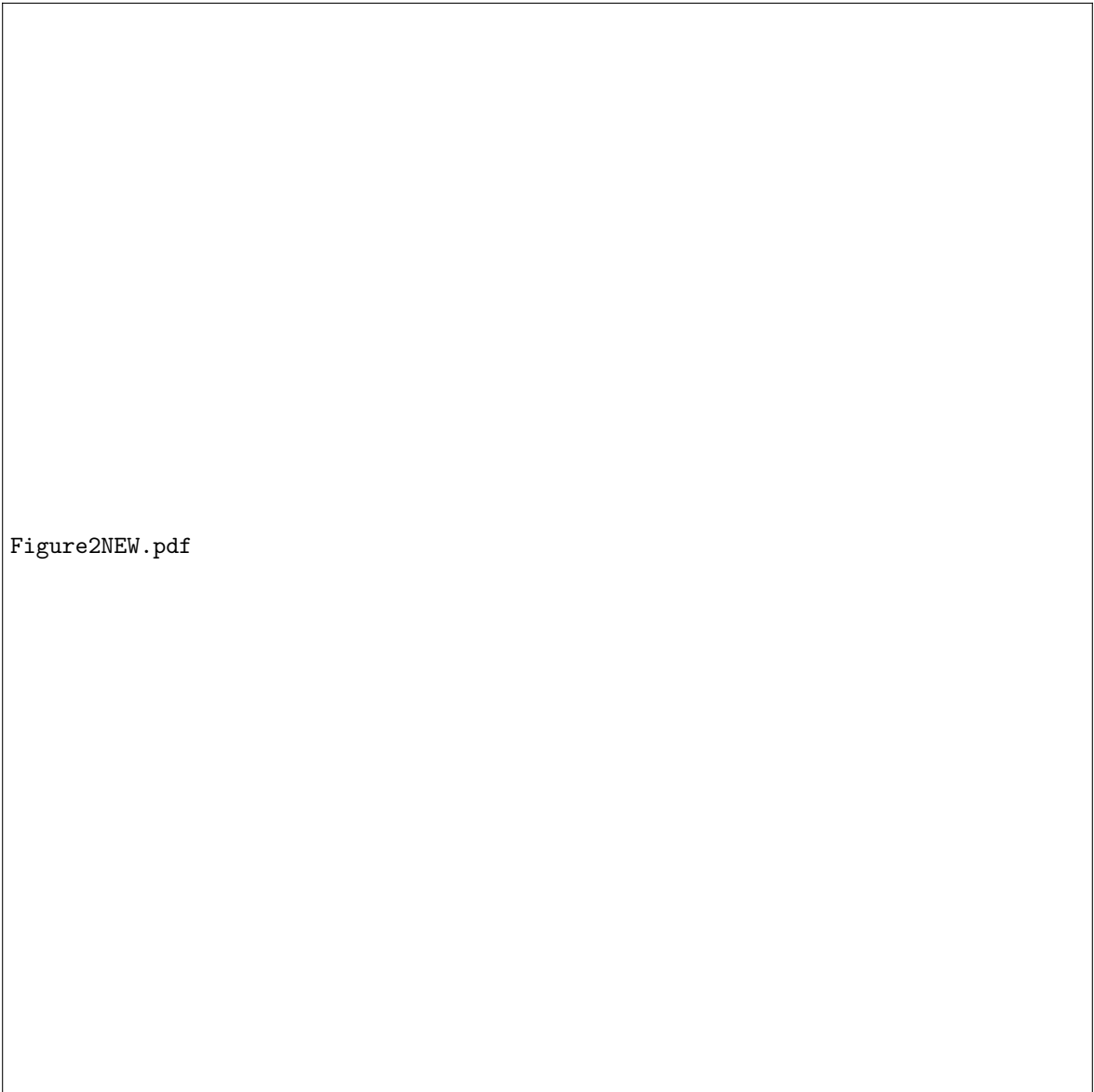
Figure 3: Constrained multiple hole/particle (CMHP) algorithm illustrated in the case of two hole and three particle orbitals.

In the following we describe our constrained multiple hole/particle (CMHP) algorithm, which generalizes Eqs. (2.12Theoryequation.2.2.12) and (2.13Theoryequation.2.2.13) to the case of n electronic states. Due to the complexity of this algorithm, we recommend the reader to follow its description with the help of Fig. 2.3Constrained multiple hole/particle (CMHP) algorithm illustrated in the case of two hole and three particle orbitalsfigure.2.3. Suppose we are interested in the excited states that result from the excitation of a given number of core orbitals (n_c) and unoccupied orbitals (n_u). For convenience we will label the excited state Kohn–Sham determinants [$\Phi^{(i,a)}$] with two indices, i and a, which stand respectively for the core and unoccupied orbital that are involved in an excited state. Our algorithm starts with a ground-state DFT computation, which yields the determinant $\Phi^{(0,0)}$. Next, we perform a sweep of n_u OCDFT computations, which produces the series of solutions:

$$\Phi^{(1,1)}, \Phi^{(1,2)}, \ldots, \Phi^{(1,n_u)}. \tag{14}$$

These solutions are characterized by hole (particle) orbitals that span the occupied (unoccupied) space of $\Phi^{(0,0)}$:

$$\begin{cases} \hat{Q}^{(0)} \phi_h^{(1,a)} = 0 \\ \hat{P}^{(0)} \phi_p^{(1,a)} = 0 \end{cases}, \qquad \forall a \leq n_u, \tag{15}$$

and particle orbitals that form an orthogonal set:

$$[1]\phi_p^{(1,a)} \phi_p^{(1,b)} = \delta_{ab}, \qquad \forall a, b \leq n_u. \tag{16}$$

In the following iteration of the CMHP algorithm we increase the core index by one and sweep again through a series of solutions:

$$\Phi^{(2,1)}, \Phi^{(2,2)}, \ldots, \Phi^{(2,n_u)}, \tag{17}$$

such that the hole and particle orbitals span respectively the virtual and occupied space of the ground-state determinant:

$$\begin{cases} \hat{Q}^{(0)} \phi_h^{(2,a)} = 0 \\ \hat{P}^{(0)} \phi_p^{(2,a)} = 0 \end{cases}, \qquad \forall a \leq n_u, \tag{18}$$

and the particle orbitals are orthogonal:

$$[1]\phi_p^{(2,a)} \phi_p^{(2,b)} = \delta_{ab}, \qquad \forall a, b \leq n_u. \tag{19}$$

In addition, we impose orthogonality between the first core orbital of the first sweep, $\phi_h^{(1,1)}$, and the second core orbital of each of the states determined during the current sweep, $\phi_h^{(2,a)}$:

$$[1]\phi_h^{(1,1)} \phi_h^{(2,a)} = 0, \qquad \forall a \leq n_u. \tag{20}$$

This condition guarantees that the hole orbital for the second series of computations is different from the first one. The CMHP algorithm proceeds in a similar way until $i = n_c$. After each sweep over the particle orbitals, a new set of orthogonality conditions is imposed onto the hole orbitals. We notice that the CMHP algorithm does *preserve strict orthogonality among all states*, but it is not equivalent to imposing the strictly minimum orthogonality conditions.

In summary, the CMHP algorithm simplifies the optimization of a series of mutually orthogonal determinants by imposing separate orthogonality conditions onto the core (hole) and valence (particle) orbitals. It is instructive to compare the computational cost of OCDFT and TDDFT. The cost to compute d excited state in OCDFT is proportional to dN^3 and dN^4 for pure and hybrid functional, respectively, since for each excited states one has to solve a set of constrained Kohn–Sham equations. In the case of TDDFT, the computational cost is dominated by the solution of a pseudo-eigenvalue problem, which scales as dN^4. TDDFT certainly provides a more direct approach to the computation of multiple excited states, but in our experience we observe that the OCDFT/CMHP algorithm has a cost comparable to that of TDDFT and in most cases can be applied in a black-box way.

3   Computational Details

Our OCDFT for core-excited states was implemented as a plugin for the PSI4 *ab initio* quantum chemistry package.[49] The OCDFT excitation energies for a set of 13 small molecules reported in this study were computed using the

B3LYP,[50–53] PBE0[54], and BLYP[53,55] functionals, using the correlation-consistent polarized core-valence basis sets (cc-pCVXZ, X = T,Q)[56] and the Karlsruhe valence polarized basis sets (def2-XZVP, X = T,Q).[57,58] All geometries were optimized at the same level of theory as the given excitation energy. The optimized geometries and the NEXAS spectrum of adenine and thymine were computed at the def2-TZVP/B3LYP level of theory (cartesian geometries are provided in the supplementary material). Benchmark TDDFT excitation energies were computed at the B3LYP/def2-QZVP level of theory using the ORCA software package.[59] All excitation energies reported in this work, including TDDFT, are for singlet states. In the case of OCDFT, the singlet excitation energies were computed via the spin-raising approach described in Ref. 30.

It is mandatory to consider relativistic effects when studying excitations of core electrons.[60–63] For excitations involving 1s orbitals, we approximate the relativistic excitation energy (ω^{R}) as the sum of the nonrelativistic excitation energy (ω^{NR}) minus a correction $\Delta\epsilon_{\mathrm{1s}}$:

$$\omega^{\mathrm{R}} = \omega^{\mathrm{NR}} - \Delta\epsilon_{\mathrm{1s}}, \tag{21}$$

where $\Delta\epsilon_{\mathrm{1s}}$ is the energy difference between the ground-state nonrelativistic (NR) and relativistic (R) Kohn–Sham energies of the 1s orbital:

$$\Delta\epsilon_{\mathrm{1s}} = \epsilon_{\mathrm{1s}}^{\mathrm{R}} - \epsilon_{\mathrm{1s}}^{\mathrm{NR}}. \tag{22}$$

In this work, the excitation energies computed in OCDFT and TDDFT utilize relativistic orbital energies calculated with first-order Douglass–Kroll–Hess (DKH) Hamiltonian.[64–66] Relativistic corrections for the 1s orbitals of the second row nuclei range from 3.8 eV (Si) to 10.1 eV (Cl). In the case of first row 1s core orbitals, $\Delta\epsilon_{\mathrm{1s}}$ is negligible (C, N, and O 1s corrections are about 0.1, 0.2, and 0.3 eV, respectively). Similarly, excitations from 2p orbitals of second row elements are negligible (max 0.05 eV) and were not applied to the final results.

The treatment of core-excited states in molecules with symmetry equivalent atoms becomes problematic for both pure and hybrid functionals due to the approximate treatment of exchange and correlation which introduces a self-interaction error.[67,68] In this case, the symmetry restricted solution produces core holes distributed evenly amongst the symmetry equivalent atoms. Instead, the symmetry unrestricted solution may consists of core holes localized on each individual atom. For all molecules with symmetry equivalent atoms (N_2, C_2H_2, and Cl_2) we studied both the symmetry restricted and unrestricted solutions. To obtain a state where the core hole is localized, we utilize a wave function with broken spatial and spin symmetry mixing the coefficients of the alpha and beta orbitals.

Peak intensities for the transition $\Psi^{(n)} \leftarrow \Psi^{(0)}$ are based on the oscillator strength (f_{osc}):

$$f_{\mathrm{osc}} = \frac{2}{3}|\mu_{n0}|^2\omega_n, \tag{23}$$

which is calculated from the excitation energies (ω_n) and transition dipole moments ($\boldsymbol{\mu}_{n0}$). Although OCDFT does not provide a direct way to compute transition dipole moments, these can be approximated using the Kohn–Sham determinants as:

$$\boldsymbol{\mu}_{n0} = [1]\Phi^{(n)}\hat{\mathbf{r}}[1]\Phi^{(0)}, \tag{24}$$

where $\Phi^{(n)}$ is a generic excited state and $\hat{\mathbf{r}}$ is the position vector. Eq. (2.23Computational Detailsequation.2.3.23) yields the absolute oscillator strength (f_{abs}) for a given transition. The intensity of the spectrum is then scaled relative to the most intense peak, we will refer to these scaled values as the *relative oscillator strength* (f_{rel}). Natural spectroscopic broadening effects are simulated by convoluting the OCDFT peaks with a Gaussian function whose exponent was calibrated to best fit the experimental spectrum.

4 Results and Discussion

4.1 Calibration of OCDFT core-excitation energies

The accuracy of OCDFT was benchmarked using a test set that comprises molecules containing first-row (CO, H_2CO, N_2O, N_2, HCN, CH_4, C_2H_4) and second-row elements (SiH_4, PH_3, H_2S, SO_2, HCl, Cl_2), and a total of 40 excited states. Table 2.1OCDFT core-excitation energies for a benchmark set composed of 13 diatomic molecules. Mean absolute error (in eV) computed using various combinations of basis sets and density functionals. These statistics refer to a subset of the benchmark set comprised of 35 core-excited statestable.2.1 summarizes the performance of OCDFT by reporting the mean absolute error (MAE) in the excitation energy for 12 unique combinations of basis set and
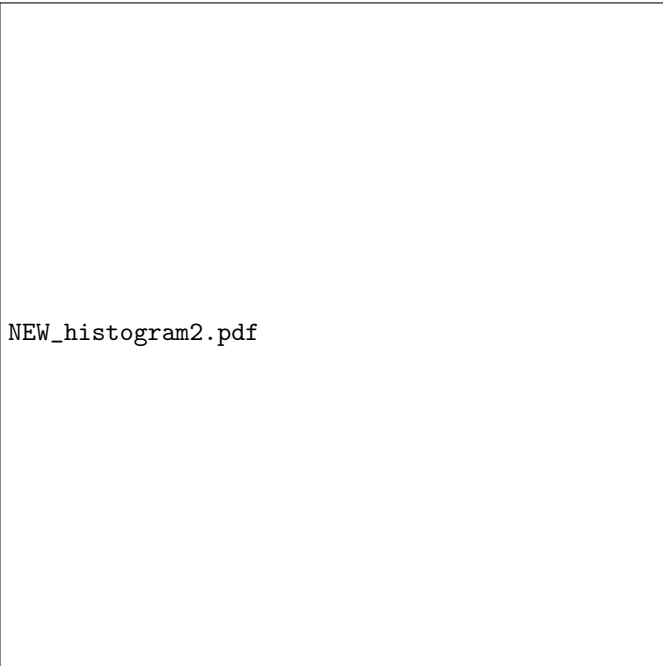
density functionals. Individual excitation energies computed with OCDFT and TDDFT are compared to values from gas-phase NEXAS experiments[69–83] in Tables 2.2Core excitation energies for molecules containing first-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 69– 77table.2.2 and 2.3Core excitation energies for molecules containing second-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 78– 83table.2.3. Table 2.2Core excitation energies for molecules containing first-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 69– 77table.2.2 displays excitations from first-row elements, while Table 2.3Core excitation energies for molecules containing second-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 78– 83table.2.3 shows excitations from second-row elements. For molecules with symmetry-equivalent atoms (N_2, C_2H_2, and Cl_2), we examined both the symmetrical and symmetry-broken solutions and found that the former have a MAE of 13.7 eV, while the latter yield a MAE of 0.9 eV. Thus all results for highly symmetric molecules reported in Tables 2.2Core excitation energies for molecules containing first-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 69– 77table.2.2 and 2.3Core excitation energies for molecules containing second-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 78– 83table.2.3 are based on broken symmetry solutions. The three functionals considered in Table 2.1OCDFT core-excitation energies for a benchmark set composed of 13 diatomic molecules. Mean absolute error (in eV) computed using various combinations of basis sets and density functionals. These statistics refer to a subset of the benchmark set comprised of 35 core-excited statestable.2.1 produce MAEs ranging from 0.9–1.7 eV. Interestingly, there is no dramatic difference in the accuracy of OCDFT regardless of the amount of Hartree–Fock (HF) exchange present in the functional. Even the BLYP functional, which contains no HF exchange, yields a MAE (1.0 eV using the def2-QZVP basis set) comparable to that of the MAEs produced by its hybrid counterparts. The Karlsruhe family of basis sets yields results that are in better agreement with the experimental excitation energies than the correlation-consistent basis sets.

The average error of OCDFT is commensurate to that of wave function methods for core-excited states. For example, Asmuruf and Besley[9] reported an average error of 1.2 eV for SOS-CIS(D) applied to a set of excitations similar to the ones used in the present study. While Coriani et al.[8] reported absolute errors of less than 0.9 eV when applying coupled cluster response theory to a set of carbon, nitrogen, and neon core excitations.

A full comparison of the accuracy of OCDFT and TDDFT core-excitations computed at the B3LYP/def2-QZVP level of theory is shown in Fig. 2.4Histogram showing the distribution of the error in the computed core excited states. All calculations were done using the B3LYP functional and the def2-QZVP basis set. The red filled bar is OCDFT while the empty blue bar is TDDFT. Notice that the two distributions do not overlapfigure.2.4. The contrast between the two error distributions is striking. As expected, TDDFT performs rather poorly underestimating the excitation energies, on average, by 15 eV and a maximum error of -53.6 eV. On the contrary, OCDFT yields an error distribution peaked near zero and a maximum error of -3.7 eV. The TDDFT error distribution has a peculiar shape, displaying two

Table 1: OCDFT core-excitation energies for a benchmark set composed of 13 diatomic molecules. Mean absolute error (in eV) computed using various combinations of basis sets and density functionals. These statistics refer to a subset of the benchmark set comprised of 35 core-excited states.

Basis Set	Mean Absolute Error (eV)		
	BLYP	B3LYP	PBE0
def2-TZVP	1.0	1.0	0.9
def2-QZVP	1.0	1.0	1.4
cc-pCVTZ	1.3	1.3	1.6
cc-pCVQZ	1.5	1.5	1.7

Figure 4: Histogram showing the distribution of the error in the computed core excited states. All calculations were done using the B3LYP functional and the def2-QZVP basis set. The red filled bar is OCDFT while the empty blue bar is TDDFT. Notice that the two distributions do not overlap.

distinct groups of excited states. The first is a narrow distribution that exists in the range -4 eV to -15 eV, while the second one is broader and ranges from -54 eV to -38 eV. An analysis of the group of excited states with the largest errors reveals that these consist solely of 1s core-excitations of second-row elements. This finding is in agreement with previous studies by Nakata[84] and Besley.[24] Since our excitation energies are corrected for relativistic effects (albeit with a crude approximation), the bulk of the error observed for second-row elements must be attributed to a deficiency of the exchange-correlation functional.[85] This dramatic difference in the performance of TDDFT suggests that it is helpful to separately analyze first-row and second-row core excitations to highlight their distinctive features.

Table 2.2Core excitation energies for molecules containing first-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 69– 77table.2.2 reports valence and Rydberg excitation energies of 1s orbital localized on first row elements (C, N, or O). Compared to experiment, TDDFT produces a mean absolute error of 11.6 eV, and as discussed earlier, this result is consistent with previous studies that used TDDFT with conventional hybrid functionals.[15] OCDFT calculations yield a mean absolute error of 0.4 eV. To put this error into perspective, it can be compared to the performance of TDDFT with the $BH^{0.58}LYP$ functional,[24] a reparameterization of the BHLYP functional[50,86] that has been augmented to include 58% HF exchange, 39% B88 exchange, and 8% Slater exchange. When applied to a test set of first row core excitations similar to those in Figure 2.2Core excitation energies for molecules containing first-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 69– 77table.2.2, $BH^{0.58}LYP$ yielded a mean average error of 0.8 eV. It is encouraging that OCDFT can achieve a higher level of accuracy without altering the amount of Hartree–Fock exchange included in the functional. It is also gratifying to see that the OCDFT MAE for this set of first row core excited states is comparable to the MAE obtained for valence excited states (0.3 eV reported in Ref. 30). When computing core excited states of second-row nuclei, TDDFT becomes highly inaccurate, producing an average error larger than 30 eV. Previous work by Tozer and coworkers[23] showed that there is a correlation between the level of accuracy of TDDFT excitation energies and the amount of overlap between the orbitals involved. We expect this correlation to also be observed in core electron excitations, where the core hole and valence particle orbitals have little overlap. Following Tozer et al.,[23] in OCDFT

we define the overlap between the hole and particle orbital (Λ_{hp}) for any excited state n as the integral:

$$\Lambda_{\text{hp}} = \int |\phi_{\text{h}}^{(n)}(\mathbf{r})||\phi_{\text{p}}^{(n)}(\mathbf{r})| \; \mathrm{d}\mathbf{r}. \tag{25}$$

 Multiple studies have concluded that the shortcoming of traditional density functionals is not the treatment of long-range interaction, but instead it is the incorrect description of the core region that leads to inaccuracy.[87][?][−89] This problem was specifically addressed by the SRC-BLYP functional[24] which is a flexible functional that increases the amount of Hartree–Fock exchange at short range and is identical to the long range corrected CAM-BLYP at long range. It has been concluded that varying HF exchange parameters can vastly improve the description of core-excited states,[26,87–90] however, here we approached the problem in a different way. With a slight augmentation to DFT derived from first principles, an accurate description of core excited states is possible without relying on varying the amount of HF exchange.   Figure 2.5Scatterplot displaying the excitation energy error as a function of orbital overlap for a benchmarking test set of 40 core excitations from 13 different molecules. Excitation energies were calculated using the B3LYP functional and def2-QZVP basis set. The overlap integrals were computed with numerical grid integration making use of the Gaussian cube files produced by in Psi4 OCDFT calculations. Grids were calculated with a double zeta basis set and 0.1 grid spacingfigure.2.5 reports the distribution of OCDFT and TDDFT excited states as a function of the energy error and the hole/particle orbital overlap. The scatterplot clearly shows that OCDFT is less sensitive to variations in the overlap. When calculating core excited states with $\Lambda_{\text{hp}} < 0.12$, the MAE for OCDFT increases by only 1.5 eV, while in the case of TDDFT the absolute error increases drastically by 35.9 eV.

Tables 2.2Core excitation energies for molecules containing first-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 69– 77table.2.2 and 2.3Core excitation energies for molecules containing second-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 78– 83table.2.3 also report oscillator strengths computed with OCDFT at the B3LYP level of theory. An extensive quantitative comparison with experimental line intensities is not practical, however, some qualitative analysis is possible and provides insight

Figure 5: Scatterplot displaying the excitation energy error as a function of orbital overlap for a benchmarking test set of 40 core excitations from 13 different molecules. Excitation energies were calculated using the B3LYP functional and def2-QZVP basis set. The overlap integrals were computed with numerical grid integration making use of the Gaussian cube files produced by in Psi4 OCDFT calculations. Grids were calculated with a double zeta basis set and 0.1 grid spacing.

Table 2: Core excitation energies for molecules containing first-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 69– 77.

| Molecule | Excitation | Exp. (eV) | Error (eV) | | $f_{abs} \times 10^{-2}$ | | |
			TDDFT	OCDFT	OCDFT		
CO	C 1s → π*	287.4	−11.3	−0.8	4.23	−10.9	0.9
	C 1s → 3s	292.4	−10.5	0.9	0.26	−10.8	1.2
	O 1s → π*	534.2	−13.4	−1.2	1.74	−14.4	−1.4
	O 1s → 3s	538.9	−13.0	0.2	0.01	−13.9	0.3
H$_2$CO	C 1s → π*	286.0	−10.7	−0.6	3.33	−10.7	−0.8
	C 1s → 3s	290.2	−10.7	−0.2	0.62	−10.7	−0.3
	O 1s → 3s	535.4	−14.1	−0.6	0.06	−14.1	−0.7
	O 1s → π*	530.8	−14.0	−0.8	2.04	−14.1	−1.1
N$_2$O†	O 1s → π*	534.8	−14.3	−1.0	1.08	−14.3	−1.3
	O 1s → 3s	536.7	−13.6	−0.4	1.67	−13.6	−0.6
	N$_c$ 1s → 3s	407.5	−12.1	0.6	3.11	−12.6	0.6
	N$_t$ 1s → π*	401.1	−12.2	−0.9	2.41	−14.6	2.5
	N$_t$ 1s → 3s	404.0	−11.5	−0.4	0.90	−11.0	−0.5
N$_2$	N 1s → π*	401.0	−12.4	−0.9	2.90	−12.4	−1.1
	N 1s → 3s	406.2	−8.5	1.7	0.00	−7.2	2.7
HCN	C 1s → π*	286.4	−10.6	−0.5	2.57	−10.6	−0.7
	C 1s → 3s	289.1	−9.9	−0.1	0.81	−9.9	−0.3
	N 1s → π*	399.7	−12.0	−0.8	2.46	−12.0	−1.0
	N 1s → 3s	401.8	−10.4	0.2	0.28	−10.4	0.0
CH$_4$	C 1s → 3p	288.0	−10.1	0.1	1.85	−10.1	0.0
	C 1s → 3s	287.1	−10.8	−0.5	0.00	−10.8	−0.6
C$_2$H$_2$	C 1s → π*	285.8	−10.5	−0.6	2.27	−10.4	−0.7
	C 1s → 3s	287.7	−9.1	−0.1	0.11	−8.9	−0.3
MAE			11.6	0.4		11.7	0.6

†The subscripts c and t stand for the center and tail nitrogen of N$_2$O.

into the reliability of the computed oscillator strengths. In general, the lower energy core-valence excitations are more intense than the higher energy Rydberg excitations. For example, an analysis of the experimental K-edge spectrum of carbon monoxide obtained by Domke et al.[77] shows that the C 1s → π* transition is a sharp, very intense peak at 287.4 eV. While the C 1s → 3s transition is a peak of significantly lower intensity at 292.4 eV. OCDFT produces an oscillator strength for the C 1s → π* transition that is an order of magnitude greater than that of the C 1s → 3s transition, which is consistent with the observed experimental trend. The same comparison can be done with the oxygen K-edge of carbon monoxide and OCDFT shows similar agreement with the experimental spectrum.

To understand why OCDFT outperforms TDDFT, we will consider a model consisting of two electrons in two orbitals (ϕ_h, ϕ_p) of different symmetry.[30,38,91] This model makes it possible to compare the TDDFT and OCDFT excitation energies to that of CIS. As illustrated in Fig. 2.5Scatterplot displaying the excitation energy error as a function of orbital overlap for a benchmarking test set of 40 core excitations from 13 different molecules. Excitation energies were calculated using the B3LYP functional and def2-QZVP basis set. The overlap integrals were computed with numerical grid integration making use of the Gaussian cube files produced by in Psi4 OCDFT calculations. Grids were calculated with a double zeta basis set and 0.1 grid spacingfigure.2.5, core excitations are characterized by a very small overlap between the hole and particle orbital. Therefore, our analysis considers the limit of zero overlap ($\Lambda_{hp} = 0$). For a functional containing a given fraction (a) of Hartree–Fock exchange, the TDDFT and CIS singlet excitation energies (ω_s) for our model differ approximately by:

$$\omega_s^{\text{TDDFT}} - \omega_s^{\text{CIS}} \cong (1-a)[v_p^x - v_h^x] + (1-a)[J_{ph} - J_{hh}], \tag{26}$$

where $v_i^x = (\phi_i|v_x|\phi_i)$ is an exchange potential integral and J_{ij} is the Coulomb repulsion integral $(\phi_i\phi_i|r_{12}^{-1}|\phi_j\phi_j)$. When $a \ll 1$, the right-hand side of Eq. 2.26Calibration of OCDFT core-excitation energiesequation.2.4.26 contains three local integrals v_h^x, v_p^x, and J_{hh}. However, the Coulomb repulsion integral between the hole and particle orbitals

Table 3: Core excitation energies for molecules containing second-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 78– 83.

Molecule	Excitation	Exp. (eV)	Error (eV)		$f_{abs} \times 10^{-2}$		
			TDDFT	OCDFT	OCDFT		
SiH_4	Si $1s \to \sigma^*$	1842.5	−38.4	−1.8	0.18	−38.9	−2.3
	Si $2p \to \sigma^*$	102.8	−4.8	0.6	0.33	−4.1	1.5
PH_3	P $1s \to \sigma^*$	2145.8	−44.1	−2.9	0.23	−44.1	−3.2
	P $2p \to \sigma^*$	132.3	−5.1	0.7	0.36	−5.1	0.0
H_2S	S $1s \to \sigma^*$	2473.1	−48.3	−3.0	0.14	−48.3	−3.8
	S $1s \to 4p$	2476.3	−52.1	−1.5	0.24	−52.1	−0.6
	S $2p \to \sigma^*$	164.5	−5.1	0.8	1.98	−5.1	0.4
	S $2p \to 4s$	166.5	−7.1	−0.7	0.51	−7.1	−1.0
SO_2	S $1s \to \pi^*$	2473.8	−50.1	−3.7	0.39	−51.3	−4.5
	S $1s \to 4p$	2478.4	−49.3	−2.4	0.24	−50.4	−3.2
	S $2p \to 4s$	171.3	−8.3	−1.5	0.05	−8.2	−1.9
HCl	Cl $1s \to \sigma^*$	2823.9	−53.8	−2.3	0.21	−55.4	−4.7
	Cl $1s \to 4p$	2827.8	−52.1	−0.7	0.22	−52.7	−1.7
	Cl $2p \to \sigma^*$	201.0	−6.1	0.8	0.00	−6.1	0.3
Cl_2	Cl $1s \to \sigma^*$	2821.3	−53.6	−1.6	0.35	−55.3	−3.6
	Cl $1s \to 4p$	2828.5	−51.7	0.9	0.14	−53.5	−0.2
	Cl $2p \to \sigma^*$	198.7	−5.7	−0.8	0.00	−5.8	−0.3
MAE			31.6	1.6			

(J_{ph}) is nonlocal and causes TDDFT to incorrectly describe the physics of the hole/particle pair. On the contrary, when $a = 1$ there is exact cancellation of the nonlinear terms and TDDFT is equivalent to CIS. This is consistent with the observation that increasing the amount of Hartree–Fock exchange improves the description of core-excited states. [26,87–90]

The same analysis finds that the OCDFT and CIS excitation energies differs by a sum of local self-interaction terms:

$$\omega_s^{OCDFT} - \omega_s^{CIS} \cong (1-a)[v_p^x - v_h^x + \frac{1}{2}J_{pp} - \frac{1}{2}J_{hh}$$
$$+ \frac{1}{2}(hh|\hat{f}_x|hh) + \frac{1}{2}(pp|\hat{f}_x|pp)], \tag{27}$$

where $(ii|\hat{f}_x|ii)$ is an exchange kernel integral. As observed by Ziegler and co-workers in the case of charge-transfer excitations computed via the constricted variational DFT, [38] there is partial cancellation between the local integrals that appear in Eq. (2.27Calibration of OCDFT core-excitation energiesequation.2.4.27) and the excitation energy has the correct asymptotic behavior. Our observation that the amount of Hartree–Fock exchange has little influence on the computed core excitation energies, suggests that even in the case of OCDFT there is partial cancellation of the terms in Eq. (2.27Calibration of OCDFT core-excitation energiesequation.2.4.27) that appear in square brackets. Thus, although charge-transfer and core-excited states are very different in nature, this simple model suggests a formal connection between the two.

4.2   Application to Nucleobases: Thymine and Adenine Near-Edge Spectra

Nucleobases play a key biological role as the building blocks of DNA and recently show promise as potential materials for electronic/technological applications. [92–95] Early X-ray studies of nucleobases consisted of scattering and diffraction experiments. [96–99] The first near-edge absorption experiments were performed in the 90s by Mitra-Kirtley et al. [100] These authors specifically targeted the nitrogen core electrons and probed the sensitivity of the $1s \to \pi^*$ resonances to the surrounding chemical environment. More recent experiments have moved beyond simple characterization of the intramolecular environment and aim to probe intermolecular interactions of nucleobases with metal surfaces. [101–104]

At the same time a wide array of computational methods have been used to compute the near-edge structure of nucleobases, including: restricted active space SCF (RASSCF),[105] improved virtual orbital SCF (IVO-SCF)[106], a complex polarization propagator method (CPP)[107], SIC-DFT,[108] a DFT transition potential method (DFTTP),[106] the equivalent core approximation (ECA) method,[109] and the second-order algebraic diagrammatic construction [ADC(2)] method.[110,111]

Here we present an OCDFT simulation of the gas-phase NEXAS spectrum of thymine and adenine. We first give an overview of the performance of OCDFT relative to the gas-phase NEXAS experiments done by Plekan et al.[110] This is followed by an in-depth analysis of the spectral features simulated in OCDFT. Last we compare our work to previous studies employing ADC(2) theory. The numbering schemes used for adenine and thymine are shown in Tables 2.4Calculated (B3LYP/def2-TZVP) and experimental thymine oxygen, nitrogen, and carbon core excitation energies (ω_{fi}, in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_{h}) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_{p})table.2.4 and 2.6Calculated (B3LYP/def2-TZVP) and experimental adenine nitrogen and carbon core excitation energies (ω_{fi}, in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_{h}) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_{p})table.2.6. We follow a widely used convention of numbering the atoms according to the Hartree–Fock core orbital energy. The relevant ground-state virtual orbitals for thymine and adenine are shown in Figs. 2.6Relevant virtual orbitals for thymine numbered in according to the orbital energy. Orbitals with obvious π^* character are labeled as such, while orbitals where electron density is diffused are labeled as Dfigure.2.6 and 2.8Relevant virtual orbitals for adenine numbered according to the orbital energy. Orbitals with obvious π^* character are labeled as such, while orbitals where electron density is mostly diffused on the outside of the molecule are labeled as $\phi_D$figure.2.8, respectively. The simulated and experimental NEXAS spectra of thymine and adenine are reported in Figs. 2.7Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.7 and 2.9Core excited states for adenine computed using B3LYP functional and def2-TZVP basis set. The OCDFT carbon and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.9.

4.2.1 Overall Performance

The simulated thymine and adenine NEXAS spectra shown in Figs. 2.7Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.7 and 2.9Core excited states for adenine computed using B3LYP functional and def2-TZVP basis set. The OCDFT carbon and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.9 agree well with the experimental data. Tables 2.4Calculated (B3LYP/def2-TZVP) and experimental thymine oxygen, nitrogen, and carbon core excitation energies (ω_{fi}, in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_{h}) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_{p})table.2.4 and 2.6Calculated (B3LYP/def2-TZVP) and experimental adenine nitrogen and carbon core excitation energies (ω_{fi}, in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_{h}) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_{p})table.2.6 report the dominant contributions to the NEXAS spectra (full tables showing all contributions are provided in the supplementary material), along with excitation energies and relative oscillator strengths. We also report the nature of each excited state. That is, for each transition we specify the core 1s electron excited (X_i where X = O, N, C and i is the label of the atom in our numbering scheme), and the ground-state virtual orbital that has the most overlap with the particle orbital, together with its weight. The experimental energies reported are the peak maxima for each spectral feature, and can be approximated by the OCDFT transition in that region with the strongest oscillator strength. When using peak maxima as a comparison, OCDFT represents the thymine spectrum with an average error of 0.3 eV, and that of adenine with an average error of 0.1 eV. A common feature of NEXAS spectra is the appearance of multiple low-intensity transitions in the higher energy regions. This is represented well by OCDFT as evidenced by the stick spectrum shown in Figs. 2.7Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function
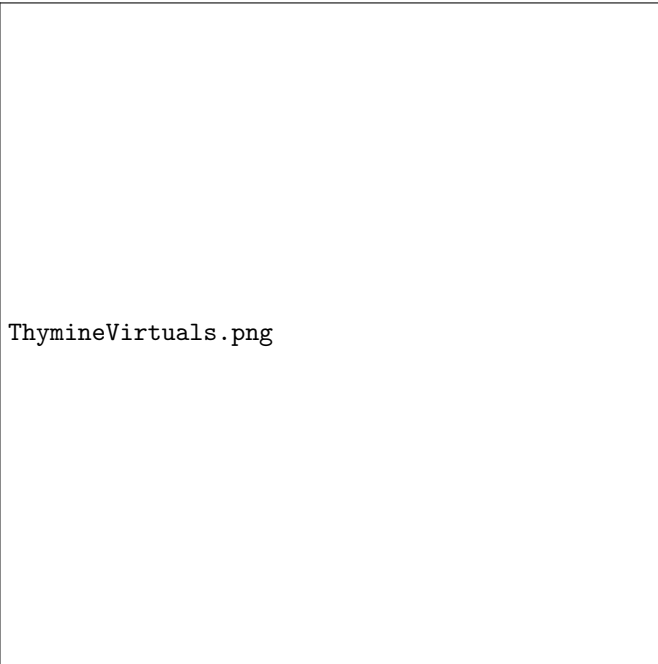
Figure 6: Relevant virtual orbitals for thymine numbered in according to the orbital energy. Orbitals with obvious π^* character are labeled as such, while orbitals where electron density is diffused are labeled as D.

with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.7 and 2.9Core excited states for adenine computed using B3LYP functional and def2-TZVP basis set. The OCDFT carbon and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.9 where the higher energy regions are populated by multiple low intensity transitions. We emphasize that the computed OCDFT spectra are obtained from unshifted excitation energies.

4.2.2 Thymine Oxygen K-Edge

Figure 2.7Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.7a displays the experimental and theoretical oxygen K-edge Spectra resulting from inner-shell ionizations are commonly referred to as K, L, or M edges depending on which core-shell is involved in the excitation. This nomenclature is simply derived from the Barkla labeling scheme for the X-ray series, and we will routinely employ it to refer to specific spectra. spectra of thymine. The low energy regime of the oxygen K-edge is dominated by two high intensity peaks. Peak A results from the transition $O_2 \rightarrow \phi_{\pi_2^*}$, while peak B results from the transition $O_1 \rightarrow \phi_{\pi_1^*}$ with a peak intensity that is roughly equal to that of peak A. Experimentally, A and B are centered at 531.4 eV and 532.3 eV, respectively, and are predicted by OCDFT to within 0.3 eV. The $O_2 \rightarrow \phi_{\pi_2^*}$ transition is predicted as the peak of highest intensity with $f_{abs} = 0.02$, and this result is consistent with the ADC(2) analysis performed by Plekan et al.[110] which predicts $f_{abs} = 0.03$. According to OCDFT, the shoulder feature B′ is composed of $O_1 \rightarrow \phi_{\pi_2^*}$ and $O_2 \rightarrow \pi_1^*$ transitions predicted to have a fairly strong oscillator strength ($f_{rel} \approx 0.2$), which is in discrepancy with the low intensity peaks observed in the experimental spectrum. As stated earlier, excitations of weaker intensity are characteristic of the higher energy regime of the K-edge and have been attributed to strong mixing of core-valence excited states with Rydberg excited states of similar energy.[112] This strong mixing causes excitations to be spread out over several different final states, resulting in transitions of weak intensity. The mixing in this region of the spectrum makes it difficult to classify specific transitions experimentally. OCDFT results show that peak C is largely composed of a mixture of diffuse Rydberg excitations within the energy interval of 534.7 to 536.4 eV. While the majority of the contributions to peak D, are excitations to ϕ_{D_2} and ϕ_{D_5}, with $f_{rel} < 0.1$, peaks C and D both have $\phi_{\pi_3^*}$ character. In both cases, these resonances are weak and overshadowed by multiple Rydberg

transitions in both cases. .D..-1

Table 4: Calculated (B3LYP/def2-TZVP) and experimental thymine oxygen, nitrogen, and carbon core excitation energies (ω_{fi}, in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_{h}) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_{p}).

ThymineNumbering.png

ϕ_{h}	ϕ_{p}	OCDFT ω_{fi}	f_{rel}	Experiment Peak	ω_{fi}
Oxygen K-Edge					
O_2	81.8% π_1^*	531.1	1.000	A	531.4
O_1	64.0% π_2^*	532.1	0.995	B	532.3
O_1	71.2% π_1^*	533.4	0.130	2*B$'$	2*533.8
O_2	78.3% π_2^*	533.8	0.143		
O_1	69.5% D_3	535.5	0.092	3*C	3*535.7
O_1	76.0% π_3^*	536.1	0.212		
O_2	69.0% π_3^*	536.2	0.090		
O_1	83.5% D_2	537.1	0.040	3*D	3*537.1
O_1	35.6% D_5	537.5	0.049		
O_2	44.0% D_5	537.7	0.065		
Nitrogen K-Edge					
N_4	81.8% π_1^*	401.2	0.998	2*A	2*401.7
N_3	64.0% π_2^*	401.8	0.802		
N_4	78.3% π_2^*	402.5	0.087	3*B	3*402.7
N_3	77.0% D_1	403.1	0.895		
N_4	65.9% D_1	403.3	0.790		
N_4	44.1% D_3	404.2	0.894	C	404.1
N_3	69.5% D_3	405.1	0.370	3*D	3*405.5
N_4	86.3% D_2	405.3	0.852		
N_3	71.2% π_1^*	405.7	0.502		
Carbon K-Edge					
C_8	92.1% π_1^*	284.9	0.328	A	284.9
C_7	95.9% π_1^*	286.0	0.635	B	285.9
C_8	97.6% π_2^*	287.3	0.140	3*C	3*287.8
C_6	81.8% π_1^*	287.7	0.770		
C_9	89.9% π_1^*	287.9	0.108		
C_9	87.7% D_1	288.5	0.205	D	288.4
C_5	64.0% π_2^*	289.1	1.000	3*E	3*289.4
C_9	63.1% π_3^*	289.3	0.294		
C_9	49.5% D_2	289.3	0.343		
C_8	33.3% D_4	290.3	0.098	3*F	3*290.7
C_9	30.5% D_4	290.4	0.042		
C_7	71.6% D_3	290.4	0.041		

4.2.3   Thymine Nitrogen K-Edge

The K-edge pictured in Figure 2.7Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.7b is characterized by four distinct spectral peaks. These lowest energy contributions to peak A are excitations from N_3 and N_4 to $\phi_{\pi_2^*}$. OCDFT predicts their excitation energies to be 401.8 eV and 401.2 eV, respectively. The experimental peak maximum is at 401.7 eV,

which agrees well with the Gaussian profile shown in the OCDFT spectrum. Rydberg transitions from the N_3 and N_4 to ϕ_{D_1} are the dominant resonances contributing to the character of peak B, along with a valence excitation $N_4 \rightarrow \phi_{\pi_1^*}$ predicted at 402.5 eV. This agrees well with the experimental peak assignment at 402.7 eV. A very intense $N_4 \rightarrow \phi_{D_3}$ transition accounts for the peak at 404.1 eV. OCDFT simulates this peak perfectly with a Gaussian centered at 404.2 eV. peak D is the amalgamation of two π^* resonances and multiple Rydberg states with the π^* resonances being the transitions of strongest intensity. Excitation energies of these π^* resonances agree well with the experimental peak assignment at 405.5 eV, with $N_4 \rightarrow \phi_{\pi_3^*}$ at 405.3 eV and $N_3 \rightarrow \phi_{\pi_1^*}$ at 405.7 eV.

4.2.4   Thymine Carbon K-Edge

The shape of the carbon K-edge displayed in Figure 2.7Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.7c is governed by four strong π^* resonances. Unique to the carbon K-edge is the fact that the strongest transition is not the lowest energy π^* resonance, the $C_5 \rightarrow \phi_{\pi_1^*}$ transition is a relatively high energy excitation and produces the strongest peak intensity, despite close proximity to several Rydberg states. Peak A at 284.9 eV is the result of the transition $C_8 \rightarrow \phi_{\pi_2^*}$, the position of this peak is predicted exactly by OCDFT. A slightly stronger transition at 285.9 eV is mostly due to a $C_7 \rightarrow \phi_{\pi_2^*}$ excitation, with small contribution from another $\phi_{\pi_2^*}$ transition resulting from an excitation from the C_9 core. Peaks C and D have experimental peak maxima at 287.8 eV and 288.4 eV respectively, and blend together to form one band. Both contributions to the spectral band are represented well by OCDFT. Peak E is mostly composed of core excitations to diffuse orbitals, however these transitions have relatively weak intensities compared to the strong $\phi_{\pi_2^*}$ transition predicted at 289.1 eV. Excitations to diffuse orbitals are dominant contributors to the remaining spectral features F and F′.

4.2.5   Adenine Nitrogen K-Edge

Figure 2.9Core excited states for adenine computed using B3LYP functional and def2-TZVP basis set. The OCDFT carbon and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.9a displays the nitrogen k-edge of adenine computed with OCDFT. The most dominant feature of the spectrum, peak A, is experimentally classified at 399.5 eV, OCDFT predicts that this feature is a result of core-valence excitations originating from three of the four nitrogens on the purine ring system (N_4, N_3, N_5). The largest contributor being a transition from $N_5 \rightarrow \phi_{\pi_2^*}$ predicted at 399.4 eV, which agrees well with the experimental classification for the peak. An apparent shoulder feature with fairly weak oscillator strength is present in the experimental spectrum in the region from 399.8 – 400.4 eV. OCDFT represents this feature well, with two weak π^* resonances resulting from the two nitrogens on the six-membered ring. A few rising shoulder features are shown in the experimental spectrum in the region from 401.0 eV – 401.3 eV, OCDFT predicts a $N_3 \rightarrow \phi_{\pi_2^*}$ transition in this region with a relative oscillator strength of 0.094. Peak B is a mixture of transitions to π^* orbitals as well as diffuse orbitals, with the π^* resonances being the prominent contributors. The experimental spectrum shows a relatively weak resonance around 403.0 (peak C). We predict that the dominant contributor to peak C is a transition from $N_2 \rightarrow \phi_{D_3}$, the intensity of which, that rivals the strongest transition (peak A $f_{rel} = 0.925$). This peak intensity is contrary to the experimental spectrum which shows peak C as a superposition of weak transitions. The highest energy transitions are all weak transitions to mostly orbitals of diffuse character, with transitions getting more intense as they approach 405.0 eV.

4.2.6   Adenine Carbon K-Edge

Fig. 2.9Core excited states for adenine computed using B3LYP functional and def2-TZVP basis set. The OCDFT carbon and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.9b compares the OCDFT and experimental carbon K-edge of adenine. The experimental carbon K-edge for adenine is dominated by a large single band with three distinct resonances in the low energy regime. The theoretical spectrum shows peaks A and B, blending together into a single band, the experimental spectrum shows these peaks at similar intensities, with peak B being slightly more intense, this is represented well in our calculated spectrum. Spectral positions for peaks A, B, and C are all in good agreement with experiment, however, the oscillator strength is inconsistent with

Table 5: Comparison of the most intense transitions at each edge of the thymine and adenine spectra. OCDFT assignments and energies (in eV) are shown and are compared to the most intense transitions from the theoretical ADC(2) spectra reported in Ref 110.

K-Edge	OCDFT	ADC(2)
Thymine		
Oxygen	$O_2 \to \pi_1^*$ (531.1)	$O_2 \to \pi_1^*$ (531.4)
Nitrogen	$N_4 \to \pi_1^*$ (401.2)	$N_4 \to \pi_1^*$ (401.5)
Carbon	$C_5 \to \pi_2^*$ (289.1)	$C_5 \to \pi_2^*$ (289.7)
Adenine		
Nitrogen	$N_5 \to \pi_2^*$ (399.4)	$N_5 \to \pi_2^*$ (399.6)
Carbon	$C_6 \to \pi_1^*$ (287.3)	$C_6 \to \pi_1^*$ (287.4)

experimental peak intensities. According to the experimental results, peak C should be roughly 50% more intense than the adjacent peak B. OCDFT predicts that the $C_6 \to \phi_{\pi_1^*}$ that dominates peak C is only slightly more intense than the dominant contributions to peaks A and B. A very weak shoulder feature is present on the falling edge of peak C, we denote this feature as C' and it is represented well by OCDFT. This shoulder results from π^* transitions with hole orbitals located on the bridge carbons (C_8 and C_{10}) and a carbon located on the six-member ring (C_9). Peaks D-G were not assigned due to their low intensities. OCDFT reveals that peak D is a very weak spectral feature with dominant contributions from excitations to diffuse orbital ϕ_{D_3}. Peaks E and F are also weak spectral features resulting mostly from transitions to diffuse orbitals. Every transition with energy higher than peak F is extremely weak

4.2.7 Comparison with ADC(2) Calculations

Previous theoretical studies performed using ADC(2) allow us to assess the accuracy of our adenine and thymine OCDFT spectra. Three key differences in the spectra are noted here. The shoulder feature B' in the thymine oxygen K-Edge shown in Fig. 2.7Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.7 is absent from the ADC(2) spectrum.[110] However, a more recent study by Wenzel, Wormit, and Dreuw[111] uses a core-valence separation (CVS) approximation to the ADC(2) working equations [CVS-ADC(2)] and predicts three excitations in this shoulder region B', all with $f_{abs} < 0.001$. These weaker oscillator strengths predicted by CVS-ADC(2) are more consistent with the experimental peak profile.

The overall shape of the OCDFT thymine nitrogen K-Edge is more consistent with the experimental excitation manifold than the ADC(2) spectrum. The two peaks A and B in Fig. 2.7Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.7b have clear, distinct maxima which are produced well quantitatively with ADC(2) (after applying a uniform shift of -2.59 eV to the spectrum), strong π^* resonances are reported near both experimental peak maxima. However, the contour of the peak is inconsistent with the experimental manifold. The ADC(2) spectrum blends into one large spectral band over the interval from 401.0 eV to 404.5 eV encompassing very closely spaced transitions, all with relatively high oscillator strengths. The extremely tight spacings and high intensities of these transitions seem to be present even in the updated CVS-ADC(2) results. The OCDFT spectrum doest not suffer from this single band issue, as the π^* resonances in peak A are well separated from the strong ϕ_{D_1} transitions in peak B by more than 1.0 eV. The agreement of these results with experiment, suggest that well-separated π^* and Rydberg resonances are more congruous with reality. However, a more detailed study of the nitrogen core excitation manifold of thymine is required to verify this observation. ADC(2) was unable to fully resolve peaks B', B, and C in the adenine nitrogen edge shown in Fig. 2.9Core excited states for adenine computed using B3LYP functional and def2-TZVP basis set. The OCDFT carbon and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV. The experimental spectra are reproduced with permission from Ref. 110figure.2.9a. On the contrary, the OCDFT spectrum represents these peaks well, as separated spectral features, in compliance with the experimental result.

Table 2.5Comparison of the most intense transitions at each edge of the thymine and adenine spectra. OCDFT assignments and energies (in eV) are shown and are compared to the most intense transitions from the theoretical

Table 6: Calculated (B3LYP/def2-TZVP) and experimental adenine nitrogen and carbon core excitation energies (ω_{fi}, in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_{h}) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_{p}).

ϕ_{h}	ϕ_{p}	ω_{fi}	f_{rel}	Peak	ω_{fi}
	OCDFT			Experiment	
				Nitrogen K-Edge	
N_4	81.0% π_1^*	399.1	0.825	3*A	3*399.5
N_3	63.7% π_1^*	399.3	0.915		
N_5	92.6% π_2^*	399.4	1.000		
N_5	92.4% π_1^*	399.7	0.001	2*A$'$	2*400.4
N_4	98.9% π_2^*	400.4	0.019		
N_3	81.6% π_2^*	401.2	0.094	B$'$	401.3
N_2	82.1% π_1^*	401.4	0.351	3*B	3*401.9
N_1	69.3% π_1^*	401.8	0.587		
N_2	78.4% D_2	402.3	0.204		
N_1	74.8% π_2^*	403.1	0.109	3*C	3*403.0
N_1	87.2% D_1	403.3	0.418		
N_2	57.1% D_3	403.3	0.925		
				Carbon K-Edge	
C_{10}	92.6% π_2^*	286.3	0.270	2*A	2*286.4
C_9	81.0% π_1^*	286.5	0.924		
C_{10}	92.4% π_1^*	286.7	0.242	2*B	2*286.8
C_7	82.1% π_1^*	286.9	0.878		
C_6	69.3% π_1^*	287.3	1.000	2*C	2*287.4
C_8	63.7% π_1^*	287.4	0.965		
C_{10}	34.7% π_3^*	287.9	0.083	2*C$'$	2*288.0
C_9	98.9% π_2^*	288.0	0.021		
C_6	74.8% π_2^*	288.8	0.006	3*D	3*289.0
C_{10}	36.4% D_2	289.2	0.033		
C_8	66.2% π_3^*	289.2	0.012		
C_9	77.4% π_3^*	289.4	0.010	3*E	
C_7	87.1% π_2^*	289.4	0.029		
C_7	57.1% D_3	289.7	0.291		
C_9	80.2% D_1	290.0	0.225	3*F	
C_8	90.2% D_1	290.1	0.038		
C_9	83.4% D_2	290.1	0.140		
C_7	94.8% π_3^*	290.4	0.073	3*G	
C_6	77.2% π_3^*	290.7	0.052		
C_7	56.6% D_1	290.8	0.047		

ADC(2) spectra reported in Ref 110table.2.5 shows a direct comparison of the OCDFT peak assignment with those obtained from ADC(2). The comparison is restricted to the transitions of highest intensity at each K-edge. All peak assignment are consistent between the two methods and we observe only small deviations in the computed excitation energy (max error 0.54 eV). The excellent agreement of OCDFT and ADC(2) for the highest intensity transitions is encouraging, and suggests that OCDFT could be a very useful tool to aid the assignment of NEXAS spectra.

5   Conclusions

Simulating NEXAS spectra is a challenge for TDDFT in tandem with conventional density functionals. Many of the

efforts to overcome this problem have centered around the development of new functionals that compensate for the shortcomings of the standard GGA functionals by varying the amount of Hartree–Fock exchange and other semi-empirical parameters. Orthogonality constrained density functional theory represents a new frontier in simulating NEXAS spectra using density functional methods.

In this work we have extended OCDFT in order to calculate multiple core-excited states of first- and second-row elements. We present two developments in OCDFT theory: 1) we show that core excitations can be easily targeted by selecting *hole* orbitals that correspond to core electrons, and 2) we proposed a generalized set of orthogonality conditions that can be used to compute multiple excited states. Our benchmark computations on core excitations from 1s and 2p orbitals of first- and second-row elements using conventional pure and hybrid functionals yield excitation energies with a mean absolute error (MAE) of 1.0 eV. OCDFT excitation energies are slightly more accurate for first-row elements (MAE = 0.4 eV) than second-row elements (MAE = 1.6 eV).

There are a few potential sources of the remaining error in the OCDFT excitation energies. First, is the fact that the (in principle) exact OCDFT excited state functional is approximated with the ground-state exchange-correlation functional. Second, in our comparison with experiment we use vertical excitation energies, which neglect vibrational effects. Third, our treatment of relativistic effects is rather simplistic. It can only approximately account for the relaxation of the core-hole and neglects relaxation effects of valence orbitals. Last, OCDFT cannot properly handle states that have a multideterminantal character. We expect that the use of an approximate excited state functional is the most important source of error, and that the relativistic and configuration mixing errors are perhaps the least relevant in the case of first- and second-row elements.

The results of our formal analysis and benchmark computations suggests that there are two factors contributing to the success of OCDFT. As in the case of other variational and constrained DFT approaches, OCDFT can properly describe charge-transfer excitations. In addition, because the wave function of the auxiliary system is only subject to orthogonality constraints, OCDFT can fully describe orbital relaxation. To demonstrate this, we consider the NEXAS spectrum of CO_2, which was suggested as a test case by one of the reviewers. Previous computational studies of CO_2[113] showed that B3LYP/TDDFT underestimates the oxygen K-edge by 14 eV. MRCI/CIS and MRCI/CISD underestimate the same feature by 8 and 4.2 eV, respectively. In the case of MRCI/CIS and MRCI/CISD, the residual error was attributed to the lack of orbital relaxation effects. A broken-symmetry OCDFT computation predicts the dominant feature of the oxygen K-edge to be at 534.4 eV, which corresponds to a deviation from experiment of 0.6 eV. This result suggests that relaxation effects do indeed play an important role in the oxygen K-edge of CO_2 and that OCDFT can fully account for them.

Excitation energies computed with TDDFT are significantly less accurate and degrade as one goes from first-row (MAE = 11.6 eV) to second-row elements (MAE = 31.6 eV). Moreover, we show that in OCDFT the choice of the functional and the amount of Hartree–Fock exchange has little effect on the accuracy of the computed core excitation energies. We perform a formal comparison of OCDFT, TDDFT, and CIS excitation energies for an excitation in which the core and virtual orbitals have zero overlap. This analysis suggests that OCDFT's superior performance can be ascribed to the local nature of the integrals that appear in the expression for the excitation energy.

Our gas-phase OCDFT X-ray absorption spectra of thymine and adenine are in excellent agreement with experiments. OCDFT reproduces all the characteristic features of the NEXAS spectra of these molecules,[110,111] including the distinct π^* transitions in the lower energy regime and the significant mixing between the π^* and diffuse orbitals in the higher energy regime of the spectra. In addition, OCDFT assignments of the spectral features are in excellent agreement with those made using the ADC(2) method.[110] This study shows that our OCDFT approach for core-excited states is a practical and useful tool for the interpretation of NEXAS experiments. From the computational point of view, the scaling of OCDFT vs. the number of electrons (N) is identical to that of ground state DFT (N^3 and N^4 for pure and hybrid functionals) and it is lower than the second-order approximate coupled cluster method (CC2)[114] and [ADC(2)], which scale as N^5.

Two classes of systems that are worth further exploration are transition metal complexes and open-shell molecules. Core excitations in transition metal complexes present a number of additional challenges that will require further extension of the present theory. One of the biggest improvements that will be necessary is a proper treatment of relativistic effects. While the scheme used in this work was sufficient for first- and second-row elements, it will likely prove ineffective for the treatment of transition metals where both scalar and spin-orbit relativistic effects play an important role. Scalar relativistic effects can be accounted for by combining OCDFT with spin-free approximate relativistic Hamiltonians. One of our immediate goals is to combine OCDFT with the one-electron spin-free version of the exact two-component approach.[115–120] This improvement will provide a more consistent way to introduce scalar relativistic effects, and will be essential to compute accurate K-edge spectra of elements past the second row. The

simulation of L-edge spectra presents additional challenges [11] due to the strong mixing of excitations from degenerate 2p core orbitals and the necessity to account for the coupling of molecular multiplets that experience strong spin-orbit coupling. In this respect, the current formulation of OCDFT—which is ideal for excitations that are dominated by a single Slater determinant—cannot properly treat multideterminantal electronic states that arise in L-edge excitations. One way to overcome this limitation is to employ the basis of non-orthogonal determinants that are generated in an OCDFT computation in a subsequent configuration interaction procedure, as described in the original formulation of OCDFT. This solution will certainly present some challenges, but if successful, could be used to account for the coupling of various molecular multiplets via spin-orbit interactions. These considerations also apply to species with a high-spin open-shell ground state and whose excited states cannot be represented by a single Slater determinant. A case that is more problematic is that of molecules with a low-spin open-shell ground state. In this situation, if the ground state DFT calculation yields an unphysical result, then it will be unlikely for OCDFT to yield accurate excitation energies.

Perhaps one of the aspects of our approach that needs further refinement is the treatment of relativistic effects. The simple additive correction used in this work can at most account for the relativistic contraction of the 1s orbital and omits relaxation effects that affect all other orbitals. In the future we plan to combine OCDFT with an approximate relativistic Hamiltonian. A more robust treatment of relativistic effects would likely improve the accuracy of the computed OCDFT core-excited states, in particular for third- and higher-row elements. Other interesting direction for future studies is the application of OCDFT to transition metal complexes, molecules in solution, and solid state problems.

6   Acknowledgements

Appendix A: OCDFT Equations for the Constrained Multiple Hole/Particle Approach

In this appendix we report details of the algorithm used to compute multiple solutions of the OCDFT equations via the constrained multiple hole/particle method. The OCDFT equations are solved following the sequence:

$$
\begin{aligned}
(i,a): (0,0) \rightarrow (1,1) \rightarrow (1,2) \rightarrow \cdots \rightarrow (1,n_{\mathrm{u}}) \rightarrow \\
\rightarrow (2,1) \rightarrow (2,2) \rightarrow \cdots \rightarrow (2,n_{\mathrm{u}}) \rightarrow \\
\vdots \\
\rightarrow (n_{\mathrm{c}},1) \rightarrow (n_{\mathrm{c}},2) \rightarrow \cdots \rightarrow (n_{\mathrm{c}},n_{\mathrm{u}}),
\end{aligned}
\tag{28}
$$

where n_{c} and n_{u} are the number of core and unoccupied orbitals, respectively, and $n_{\mathrm{c}}n_{\mathrm{u}}$ is the total number of excited states computed.

The OCDFT equations consist of a set of three coupled eigenvalue equations:

$$
\begin{aligned}
(1 - \hat{P}_{\mathrm{h/p}}^{(i,a)})\hat{f}^{(i,a)}(1 - \hat{P}_{\mathrm{h/p}}^{(i,a)})|\phi_k^{(i,a)}\rangle &= \epsilon_k^{(i,a)}|\phi_k^{(i,a)}\rangle, \\
\hat{P}^{(0)}(1 - \hat{Q}_{\mathrm{s}}^{(i,a)})\hat{f}^{(i,a)}(1 - \hat{Q}_{\mathrm{s}}^{(i,a)})\hat{P}^{(0)}|\phi_{\mathrm{h}}^{(i,a)}\rangle &= \epsilon_{\mathrm{h}}^{(i,a)}|\phi_{\mathrm{h}}^{(i,a)}\rangle, \\
\hat{Q}^{(0)}(1 - \hat{P}_{\mathrm{s}}^{(i,a)})\hat{f}^{(i,a)}(1 - \hat{P}_{\mathrm{s}}^{(i,a)})\hat{Q}^{(0)}|\phi_{\mathrm{p}}^{(i,a)}\rangle &= \epsilon_{\mathrm{p}}^{(i,a)}|\phi_{\mathrm{p}}^{(i,a)}\rangle,
\end{aligned}
\tag{29}
$$

where $\hat{f}^{(i,a)}$ is the Kohn–Sham Hamiltonian operator computed using the density corresponding to the state $\Phi^{(i,a)}$.

The projection operators that enter the OCDFT equations are defines as:

$$\hat{P}_{\text{h/p}}^{(i,a)} = \hat{P}_{\text{h}}^{(i,a)} + \hat{P}_{\text{p}}^{(i,a)} \tag{30}$$

$$\hat{P}_{\text{h}}^{(i,a)} = \sum_{j<i}^{\text{holes}} [1]\phi_{\text{h}}^{(j,1)}[1]\phi_{\text{h}}^{(j,1)} \tag{31}$$

$$\hat{P}_{\text{p}}^{(i,a)} = \sum_{b<a}^{\text{particles}} [1]\phi_{\text{p}}^{(i,b)}[1]\phi_{\text{p}}^{(i,b)}, \tag{32}$$

$$\hat{P}_{\text{s}}^{(i,a)} = \hat{P}^{(i,a)} - \hat{P}_{\text{p}}^{(i,a)}, \tag{33}$$

$$\hat{Q}_{\text{s}}^{(i,a)} = \hat{Q}^{(i,a)} - \hat{P}_{\text{h}}^{(i,a)}. \tag{34}$$

References

[1] McMillan, E. M. *Phys. Rev.* **1945**, *68*, 143–144.

[2] Hua, W.; Gao, B.; Li, S.; Ågren, H.; Luo, Y. *J. Phys. Chem. B* **2010**, *114*, 13214–13222.

[3] Contini, G.; Carravetta, V.; Di Castro, V.; Stranges, S.; Richter, R.; Alagia, M. *J. Phys. Chem. A.* **2001**, *105*, 7308–7314.

[4] Hähner, G. *Chem. Soc. Rev.* **2006**, *35*, 1244–1255.

[5] Guo, Q. X.; Senda, H.; Saito, K.; Tanaka, T.; Nishio, M.; Ding, J.; Fan, T. X.; Zhang, D.; Wang, X. Q.; Liu, S. T.; Shen, B.; Ohtani, R. *Appl. Phys. Lett.* **2011**, *98*, 181901.

[6] Feneberg, M.; Röppischer, M.; Esser, N.; Cobet, C.; Neuschl, B.; Meisch, T.; Thonke, K.; Goldhahn, R. *Appl. Phys. Lett.* **2011**, *99*, 021903.

[7] Nakazato, T.; Oyamada, M.; Niimura, N.; Urasawa, S.; Konno, O.; Kagaya, A.; Kato, R.; Kamiyama, T.; Torizuka, Y.; Nanba, T.; Kondo, Y.; Shibata, Y.; Ishi, K.; Ohsaka, T.; Ikezawa, M. *Phys. Rev. Lett.* **1989**, *63*, 1245–1248.

[8] Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. *Phys. Rev. A.* **2012**, *85*, 022507.

[9] Asmuruf, F. A.; Besley, N. A. *Chem. Phys. Lett.* **2008**, *463*, 267–271.

[10] Roemelt, M.; Neese, F. *J. Phys. Chem. A* **2013**, *117*, 3069–3083.

[11] Roemelt, M.; Maganas, D.; DeBeer, S.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 204101.

[12] Schirmer, J. *Phys. Rev. A.* **1982**, *26*, 2395.

[13] Trofimov, A. B.; Schirmer, J. *J. Phys. B: At. Mol. Opt. Phys.* **1995**, *28*, 2299.

[14] Sheehy, J. A.; Gil, T. J.; Winstead, C. L.; Farren, R. E.; Langhoff, P. W. *J. Chem. Phys.* **1989**, *91*, 1796–1812.

[15] Besley, N. A.; Gilbert, A. T. B.; Gill, P. M. W. *J. Chem. Phys.* **2009**, *130*, 124308.

[16] Ågren, H.; Jensen, H. J. A. *Chem. Phys.* **1993**, *172*, 45–57.

[17] Triguero, L.; Pettersson, L. G. M.; Ågren, H. *Phys. Rev. B* **1998**, *58*, 8097–8110.

[18] Stener, M.; Fronzoni, G.; de Simone, M. *Chem. Phys. Lett.* **2003**, *373*, 115–123.

[19] Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. *J. Chem. Theory Comput.* **2012**, *8*, 3284–3292.

[20] Runge, E.; Gross, E. K. *Phys. Rev. Lett.* **1984**, *52*, 997.

[21] Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, 104103.

[22] DeBeer George, S.; Petrenko, T.; Neese, F. *J. Phys. Chem. A* **2008**, *112*, 12936–12943.

[23] Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.

[24] Besley, N. A.; Peach, M. J. G.; Tozer, D. J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10350.

[25] Tu, G.; Carravetta, V.; Vahtras, O.; Ågren, H. *J. Chem. Phys.* **2007**, *127*, 174110.

[26] Nakata, A.; Imamura, Y.; Otsuka, T.; Nakai, H. *J. Chem. Phys.* **2006**, *124*, 094105.

[27] Capano, G.; Penfold, T.; Besley, N.; Milne, C.; Reinhard, M.; Rittmann-Frank, H.; Glatzel, P.; Abela, R.; Rothlisberger, U.; Chergui, M.; Tavernelli, I. *Chem. Phys. Lett.* **2013**, *580*, 179–184.

[28] Besley, N. A.; Noble, A. *J. Phys. Chem. C.* **2007**, *111*, 3333–3340.

[29] Besley, N. A.; Asmuruf, F. A. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024.

[30] Evangelista, F. A.; Shushkov, P.; Tully, J. C. *J. Phys. Chem. A* **2013**, *117*, 7378–7392.

[31] Kowalczyk, T.; Yost, S. R.; Van Voorhis, T. *J. Chem. Phys.* **2011**, *134*, 054128.

[32] Ziegler, T.; Rauk, A.; Baerends, E. J. *Theor. Chim. Acta.* **1977**, *43*, 261–271.

[33] Wu, Q.; Van Voorhis, T. *J. Chem. Theory Comput.* **2006**, *2*, 765–774.

[34] Görling, A. *Phys. Rev. A.* **1999**, *59*, 3359.

[35] Ziegler, T.; Seth, M.; Krykunov, M.; Autschbach, J.; Wang, F. *J. Chem. Phys.* **2009**, *130*, 154102.

[36] Ziegler, T.; Krykunov, M.; Cullen, J. *J. Chem. Theory Comput.* **2011**, *7*, 2485–2491.

[37] Krykunov, M.; Ziegler, T. *J. Chem. Theory Comput.* **2013**, *9*, 2761–2773.

[38] Ziegler, T.; Krykunov, M.; Cullen, J. *J. Chem. Phys.* **2012**, *136*, 124107.

[39] Baruah, T.; Pederson, M. R. *J. Chem. Theory Comput.* **2009**, *5*, 834–843.

[40] Olguin, M.; Zope, R. R.; Baruah, T. *J. Chem. Phys.* **2013**, *138*, 074306.

[41] Zope, R. R.; Olguin, M.; Baruah, T. *J. Chem. Phys.* **2012**, *137*, 084317.

[42] Theophilou, A. K. *J. Phys. C.* **1979**, *12*, 5419.

[43] Fritsche, L. *Phys. Rev. B.* **1986**, *33*, 3976.

[44] Gross, E. K. U.; Oliveira, L. N.; Kohn, W. *Phys. Rev. A.* **1988**, *37*, 2805.

[45] Gross, E. K. U.; Oliveira, L. N.; Kohn, W. *Phys. Rev. A.* **1988**, *37*, 2809.

[46] Levy, M.; Nagy, A. *Phys. Rev. Lett.* **1999**, *83*, 4361.

[47] Nagy, A.; Levy, M. *Phys. Rev. A.* **2001**, *63*, 052502.

[48] Ayers, P. W.; Levy, M.; Nagy, A. *Phys. Rev. A.* **2012**, *85*, 042518.

[49] Turney, J. M. et al. *WIREs Comput Mol Sci* **2012**, *2*, 556–565.

[50] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

[51] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

[52] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.

[53] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

[54] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.

[55] Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200–206.

[56] Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1995**, *103*, 4572.

[57] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

[58] Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

[59] Neese, F. *WIREs Comput Mol Sci* **2012**, *2*, 73–78.

[60] Maganas, D.; Roemelt, M.; Weyhermüller, T.; Blume, R.; Hävecker, M.; Knop-Gericke, A.; DeBeer, S.; Schlögl, R.; Neese, F. *Phys. Chem. Chem. Phys.* **2014**, *16*, 264.

[61] DeBeer George, S.; Neese, F. *Inorg. Chem.* **2010**, *49*, 1849–1853.

[62] Bauer, M. *Phys. Chem. Chem. Phys.* **2014**, *16*, 13827.

[63] Ankudinov, A. L.; Rehr, J. J.; Low, J. J.; Bare, S. R. *J. Chem. Phys.* **2002**, *116*, 1911.

[64] Douglas, M.; Kroll, N. M. *Ann. Phys.* **1973**, *82*, 89.

[65] Hess, B. A. *Phys. Rev. A* **1985**, *32*, 756–763.

[66] Hess, B. A. *Phys. Rev. A.* **1986**, *33*, 3742.

[67] Bally, T.; Sastry, G. N. *J. Phys. Chem. A.* **1997**, *101*, 7923–7925.

[68] Lundberg, M.; Siegbahn, P. E. M. *J. Chem. Phys.* **2005**, *122*, 224103.

[69] Püttner, R.; Dominguez, I.; Morgan, T. J.; Cisneros, C.; Fink, R. F.; Rotenberg, E.; Warwick, T.; Domke, M.; Kaindl, G.; Schlachter, A. S. *Phys. Rev. A.* **1999**, *59*, 3415.

[70] Remmers, G.; Domke, M.; Puschmann, A.; Mandel, T.; Xue, C.; Kaindl, G.; Hudson, E.; Shirley, D. A. *Phys. Rev. A.* **1992**, *46*, 3935.

[71] Chen, C. T.; Ma, Y.; Sette, F. *Phys. Rev. A.* **1989**, *40*, 6737.

[72] Tronc, M.; King, G. C.; Read, F. H. *J. Phys. B: At., Mol. Phys.* **1980**, *13*, 999.

[73] Tronc, M.; King, G. C.; Read, F. H. *J. Phys. B: At., Mol. Phys.* **1979**, *12*, 137.

[74] Francis, J. T.; Enkvist, C.; Lunell, S.; Hitchcock, A. P. *Can. J. Phys.* **1994**, *72*, 879–884.

[75] Adachi, J.-i.; Kosugi, N.; Shigemasa, E.; Yagishita, A. *Chem. Phys. Lett.* **1999**, *309*, 427–433.

[76] Hitchcock, A. P.; Brion, C. E. *J. Electron. Spectrosc. Relat. Phenom.* **1979**, *15*, 201–206.

[77] Domke, M.; Xue, C.; Puschmann, A.; Mandel, T.; Hudson, E.; Shirley, D. A.; Kaindl, G. *Chem. Phys. Lett.* **1990**, *173*, 122–128.

[78] Nayandin, O.; Kukk, E.; Wills, A.; Langer, B.; Bozek, J.; Canton-Rogan, S.; Wiedenhoeft, M.; Cubaynes, D.; Berrah, N. *Phys. Rev. A.* **2001**, *63*, 062719.

[79] Bodeur, S.; Millié, P.; Nenner, I. *Phys. Rev. A.* **1990**, *41*, 252.

[80] Gedat, E.; Püttner, R.; Domke, M.; Kaindl, G. *J. Chem. Phys.* **1998**, *109*, 4471.

[81] Hudson, E.; Shirley, D. A.; Domke, M.; Remmers, G.; Kaindl, G. *Phys. Rev. A* **1994**, *49*, 161–175.

[82] Cavell, R. G.; Jürgensen, A. *J. Electron. Spectrosc. Relat. Phenom.* **1999**, *101–103*, 125–129.

[83] Bodeur, S.; Esteva, J. M. *Chem. Phys.* **1985**, *100*, 415–427.

[84] Nakata, A.; Imamura, Y.; Nakai, H. *J. Chem. Theory Comput.* **2007**, *3*, 1295–1305.

[85] Saue, T. *ChemPhysChem* **2011**, *12*, 3077–3094.

[86] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

[87] Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207–8215.

[88] Song, J.-W.; Watson, M. A.; Nakata, A.; Hirao, K. *J. Chem. Phys.* **2008**, *129*, 184113.

[89] Henderson, T. M.; Izmaylov, A. F.; Scuseria, G. E.; Savin, A. *J. Chem. Phys.* **2007**, *127*, 221103.

[90] Henderson, T. M.; Izmaylov, A. F.; Scuseria, G. E.; Savin, A. *J. Chem. Theory Comput.* **2008**, *4*, 1254–1262.

[91] Casida, M. E.; Gutierrez, F.; Guan, J.; Gadea, F.-X.; Salahub, D.; Daudey, J.-P. *J. Chem. Phys.* **2000**, *113*, 7062.

[92] Di Mauro, E.; Hollenberg, C. P. *Adv. Mater.* **1993**, *5*, 384–386.

[93] Niemeyer, C. M. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 585–587.

[94] Niemeyer, C. M. *Angew. Chem. Int. Ed.* **2001**, *40*, 4128–4158.

[95] Song, B.; Cuniberti, G.; Sanvito, S.; Fang, H. *Appl. Phys. Lett.* **2012**, *100*, 063101.

[96] Langridge, R.; Marmur, J. *Science* **1964**, *143*, 1450–1451.

[97] Sundaralingam, M. *Ann. N. Y. Acad. Sci.* **1975**, *255*, 3–42.

[98] Camerman, N.; Camerman, A. *Science* **1968**, *160*, 1451–1452.

[99] Davies, D. R. *Annu. Rev. Biochem.* **1967**, *36*, 321–364.

[100] Kirtley, S. M.; Mullins, O. C.; Chen, J.; van Elp, J.; George, S. J.; Chen, C. T.; O'Halloran, T.; Cramer, S. P. *Biochim. Biophys. Acta.* **1992**, *1132*, 249–254.

[101] Seifert, S.; Gavrila, G. N.; Zahn, D. R. T.; Braun, W. *Surf. Sci.* **2007**, *601*, 2291–2296.

[102] Yamada, T.; Shirasaka, K.; Takano, A.; Kawai, M. *Surf. Sci.* **2004**, *561*, 233–247.

[103] Fujii, K.; Akamatsu, K.; Muramatsu, Y.; Yokoya, A. *Nucl. Instr. Meth. Phys. Res.* **2003**, *199*, 249–254.

[104] Fujii, K.; Akamatsu, K.; Yokoya, A. *J. Phys. Chem. B* **2004**, *108*, 8031–8035.

[105] Mochizuki, Y.; Koide, H.; Imamura, T.; Takemiya, H. *J. Synchrotron Radiat.* **2001**, *8*, 1003–1005.

[106] MacNaughton, J.; Moewes, A.; Kurmaev, E. Z. *J. Phys. Chem. B.* **2005**, *109*, 7749–7757.

[107] Ekström, U.; Norman, P.; Carravetta, V.; Ågren, H. *Phys. Rev. Lett.* **2006**, *97*, 143001.

[108] Bolognesi, P.; Mattioli, G.; O'Keeffe, P.; Feyer, V.; Plekan, O.; Ovcharenko, Y.; Prince, K. C.; Coreno, M.; Amore Bonapasta, A.; Avaldi, L. *J. Phys. Chem. A* **2009**, *113*, 13593–13600.

[109] Healion, D. M.; Schweigert, I. V.; Mukamel, S. *J. Phys. Chem. A* **2008**, *112*, 11449–11461.

[110] Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K. C.; Trofimov, A. B.; Gromov, E. V.; Zaytseva, I. L.; Schirmer, J. *Chem. Phys.* **2008**, *347*, 360–375.

[111] Wenzel, J.; Wormit, M.; Dreuw, A. *J. Chem. Theory Comput.* **2014**, *10*, 4583–4598.

[112] Robin, M. B. *Chem. Phys. Lett.* **1975**, *31*, 140–144.

[113] Maganas, D.; Kristiansen, P.; Duda, L. C.; Gericke, A. K. *J. Phys. Chem. C* **2014**, *118*, 20163–20175.

[114] Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409.

[115] Dyall, K. G. *J. Chem. Phys.* **1997**, *106*, 9618–9626.

[116] Dyall, K. G. *J. Chem. Phys.* **2001**, *115*, 9136–9143.

[117] Kutzelnigg, W.; Liu, W. *J. Chem. Phys.* **2005**, *123*, 241102.

[118] Liu, W.; Peng, D. *J. Chem. Phys.* **2009**, *131*, 031104.

[119] Zou, W.; Filatov, M.; Cremer, D. *Theor. Chim. Acta* **2011**, *130*, 633–644.

[120] Cheng, L.; Gauss, J. *J. Chem. Phys.* **2011**, *135*, 084114.

ThymineOKexperiment.png

ThymineNKexperiment.png

ThymineCKexperiment.png

Figure 8: Relevant virtual orbitals for adenine numbered according to the orbital energy. Orbitals with obvious π^* character are labeled as such, while orbitals where electron density is mostly diffused on the outside of the molecule are labeled as ϕ_D.
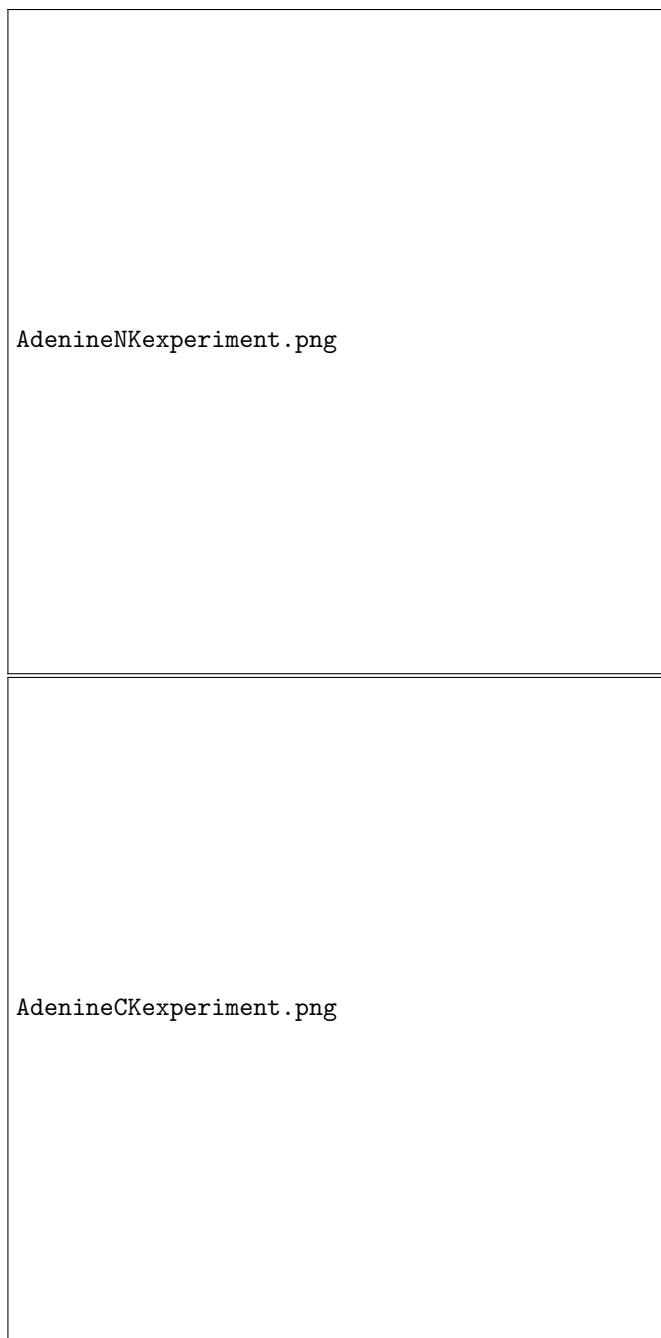
Figure 9: Core excited states for adenine computed using B3LYP functional and def2-TZVP basis set. The OCDFT carbon and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV. The experimental spectra are reproduced with permission from Ref. 110.