

# Accurate time-dependent density functional theory calculations of the near edge X-ray absorption fine structure of large systems

Stephen T. Skowron · Nicholas A. Besley

Received: 17 June 2012 / Accepted: 15 August 2012 / Published online: 2 September 2012  
© Springer-Verlag 2012

**Abstract** It is shown that by using a numerical integration grid of low quality and large two-electron pre-screening threshold, the computational cost of computing near-edge X-ray absorption fine structure (NEXAFS) spectra within time-dependent density functional theory can be reduced significantly with a very small loss in accuracy. This allows accurate NEXAFS spectra to be computed for relatively large molecules involving excitations from a large number of core orbitals using short-range corrected exchange-correlation functionals. The approach is illustrated by calculations of the carbon *K*-edge NEXAFS spectra of coronene and two semi-conducting polymers, where the calculations give good agreement with experiment and allow the origin of the different spectral features to be assigned.

**Keywords** NEXAFS · TDDFT · Large molecules

## 1 Introduction

Near-edge X-ray absorption fine structure (NEXAFS) spectroscopy is used extensively in many fields. For example, in surface science, it provides information on the structure and orientation of adsorbed molecules and the nature of the bonding to the surface [1, 2], bioinorganic chemistry where it can differentiate between different oxidation states [3–5], and it has also been used to probe the structure of liquids [6]. NEXAFS spectra arise from the excitation of a core electron to give a bound state below the ionization continuum. The technique provides information

on the unoccupied orbitals and has the advantage of being element-specific, that is, it is possible to study excitations from the core orbitals of different elements separately. Often the interpretation of experimental data is aided by theoretical calculations, making accurate calculations of NEXAFS spectra important for many applications.

There has been a number of different methods developed for calculating NEXAFS spectra. Early calculations used the multiple scattering  $X_\alpha$  method [1]. In the NEXAFS region, the muffin-tin approximation at the heart of the method leads to inaccuracies in the computed spectra, and this has stimulated the search for more accurate methods. One widely used method is the static exchange (STEX) method [7–9]. In this approach, the calculation of the absorption spectrum comprises a number of steps. A calculation of the core hole state is performed with the valence orbitals frozen, followed by optimization of the valence orbitals with the core hole frozen. The STEX Hamiltonian is diagonalized, and the excitation energies are obtained by summing the core ionization potential to the eigenvalues of the STEX Hamiltonian. In an effort to improve the STEX approach, the transition potential method was introduced [10, 11]. In this approach, the orbital binding energy is computed as the derivative of the total energy with respect to the orbital occupation number. To take into account the relaxation of the orbitals, the energy is approximated by calculating the derivative at the point corresponding to the occupation 0.5. Formally, this corresponds to a core orbital with half an electron removed which captures a balance between final and initial states.

Density functional theory (DFT) is the most widely used quantum chemistry method, and core-excited states and NEXAFS spectra can be computed within a DFT framework. Within Kohn-Sham DFT, NEXAFS spectra can be computed using a  $\Delta$ Kohn-Sham approach. In this approach the core

S. T. Skowron · N. A. Besley (✉)  
School of Chemistry, University of Nottingham, University Park,  
Nottingham NG7 2RD, UK  
e-mail: nick.besley@nottingham.ac.uk

excitation energy is the difference in the expectation values of the neutral and core-excited Kohn-Sham Hamiltonians, where the orbitals have been variationally optimized for the different states. However, obtaining a core-excited state within a Kohn-Sham formalism is not straightforward, and usually some constraints, overlap criterion or intermediate optimization with a frozen core hole is used to prevent the collapse of the core hole during the self-consistent field procedure [12–15]. An advantage of the  $\Delta$ Kohn-Sham approach is that the relaxation of the core hole is included, and a recent study showed that core excitation energies computed with the B3LYP functional were in good agreement with experiment provided that uncontracted basis functions were used [15]. The principal disadvantage of  $\Delta$ Kohn-Sham calculations is that a separate calculation is required for each core-excited state. Computing NEXAFS spectra for even relatively small molecular systems requires many different core-excited states to be computed, and the calculations can become expensive and tedious. Consequently, time-dependent density functional theory (TDDFT) in which the excited states are obtained within a single calculation becomes an attractive option for computing NEXAFS spectra [16].

It has been shown that TDDFT with standard generalized gradient approximation (GGA) and hybrid functionals result in a large underestimation of core excitation energies [17]. This has motivated the development of exchange-correlation functionals that can predict correctly core excitation energies. Recent developments include the CV-B3LYP [18] and CVR-B3LYP [19, 20] functionals of Nakai and co-workers, optimized hybrid functional for core excitations [21], the LCgau-core-BOP of Hirao et al. [22] and short-range corrected (SRC) functionals [23]. These functionals can provide core excitation energies and NEXAFS spectra that are in close agreement with experiment. However, the application of TDDFT for the calculation of NEXAFS spectra of large systems remains problematic due to the high density of states in the NEXAFS region requiring many states to be evaluated, making the calculations computationally expensive. This problem is particularly acute when excitations from many core orbitals need to be considered, for example at the carbon *K*-edge of organic molecules, and is worse for SRC-type functionals where additional integrals need to be evaluated. In this paper, approximations in the calculation of NEXAFS spectra within the framework of TDDFT with a SRC functional are investigated in order to achieve accurate calculations for large systems.

## 2 Computational details

In the Tamm-Dancoff approximation (TDA) [24] of linear response TDDFT, excitation energies and intensities are determined from the eigenvalue problem

$$\mathbf{A}\mathbf{X} = \omega\mathbf{X} \quad (1)$$

The matrix  $\mathbf{A}$  is given by

$$A_{ia\sigma,jb\tau} = \delta_{ij}\delta_{ab}\delta_{\sigma\tau}(\epsilon_a - \epsilon_i) + K_{ia\sigma,jb\tau} \quad (2)$$

where

$$K_{ia\sigma,jb\tau} = \int \int \psi_{i\sigma}^*(\mathbf{r}_1)\psi_{a\sigma}^*(\mathbf{r}_1) \left( \frac{1}{r_{12}} + \frac{\delta^2 E_{XC}}{\delta\rho_\sigma(\mathbf{r}_1)\delta\rho_\tau(\mathbf{r}_2)} \right) \times \psi_{j\tau}(\mathbf{r}_2)\psi_{b\tau}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (3)$$

$E_{XC}$  is the exchange-correlation functional, and  $\epsilon_i$  and  $\epsilon_a$  are the orbital energies of the Kohn-Sham orbitals  $\psi_i$  and  $\psi_a$ . Standard implementations of TDDFT use the iterative subspace algorithm of Davidson [25] which is inefficient for core-excited states. By their nature, core excitations are high in energy, and consequently, a very large number of roots are required to obtain core-excited states making the calculations prohibitively expensive even for small molecular systems. A number of groups have proposed different methods for overcoming this problem. The Sakurai-Sugiura projection method can be used to find excitation energies in a specified range, and this has been implemented within TDDFT and shown to be an efficient approach for core excitations [26], and a resonant converged complex polarization propagator has been implemented to study NEXAFS [27, 28]. A simple solution, which is used in this work, is to restrict the single excitation space to include only excitations from the relevant core orbital(s) [21, 29, 30]. This approximation is remarkably accurate, and it has been shown that for a range of core excitations from 1s orbitals, the largest error observed was 0.01 eV in the excitation energy and 0.01 in the oscillator strength [31]. This is a consequence of the large energy separation between core orbitals localized on nuclei with different atomic charge making the mixing between excitations included and those excluded negligible.

SRC exchange-correlation functionals are based on a reversal of the standard long-range partitioning scheme [23]. For the SRC exchange-correlation functional used in this work, the electron repulsion operator is partitioned in the evaluation of the exchange energy using the error function according to

$$\frac{1}{r_{12}} = C_{SHF} \frac{\text{erfc}(\mu_{SR}r_{12})}{r_{12}} - C_{SHF} \frac{\text{erfc}(\mu_{SR}r_{12})}{r_{12}} + C_{LHF} \frac{\text{erf}(\mu_{LR}r_{12})}{r_{12}} - C_{LHF} \frac{\text{erf}(\mu_{LR}r_{12})}{r_{12}} + \frac{1}{r_{12}}. \quad (4)$$

Treating the first and third terms of Eq. 4 with HF exchange and the remaining terms with DFT exchange leads to the following functional

$$E_{xc}^{SRC1} = C_{SHF}E_x^{SR-HF}(\mu_{SR}) - C_{SHF}E_x^{SR-DFT}(\mu_{SR}) + C_{LHF}E_x^{LR-HF}(\mu_{LR}) - C_{LHF}E_x^{LR-DFT}(\mu_{LR}) + E_x^{DFT} + E_c^{DFT} \quad (5)$$

where

$$E_x^{LR-HF} = -\frac{1}{2} \sum_{\sigma} \sum_{i,j}^{occ} \int \int \psi_{i\sigma}^*(\mathbf{r}_1) \psi_{j\sigma}^*(\mathbf{r}_1) \frac{\text{erf}(\mu_{LR}r_{12})}{r_{12}} \times \psi_{i\sigma}(\mathbf{r}_2) \psi_{j\sigma}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (6)$$

and

$$E_x^{SR-HF} = -\frac{1}{2} \sum_{\sigma} \sum_{i,j}^{occ} \int \int \psi_{i\sigma}^*(\mathbf{r}_1) \psi_{j\sigma}^*(\mathbf{r}_1) \frac{\text{erfc}(\mu_{SR}r_{12})}{r_{12}} \times \psi_{i\sigma}(\mathbf{r}_2) \psi_{j\sigma}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (7)$$

respectively. The long and short-range DFT exchange is computed from modifying the usual exchange energy [32]

$$E_x = -\frac{1}{2} \sum_{\sigma} \int \rho_{\sigma}^{4/3} K_{\sigma} d\mathbf{r} \quad (8)$$

to give

$$E_x^{LR-DFT} = -\frac{1}{2} \sum_{\sigma} \int \rho_{\sigma}^{4/3} K_{\sigma} \frac{8}{3} a_{\sigma} \times \left[ \sqrt{\pi} \text{erf}\left(\frac{1}{2a_{\sigma}}\right) + 2a_{\sigma}(b_{\sigma} - c_{\sigma}) \right] d\mathbf{r} \quad (9)$$

and

$$E_x^{SR-DFT} = -\frac{1}{2} \sum_{\sigma} \int \rho_{\sigma}^{4/3} K_{\sigma} \times \left\{ 1 - \frac{8}{3} a_{\sigma} \left[ \sqrt{\pi} \text{erf}\left(\frac{1}{2a_{\sigma}}\right) + 2a_{\sigma}(b_{\sigma} - c_{\sigma}) \right] \right\} d\mathbf{r} \quad (10)$$

where

$$b_{\sigma} = \exp\left(-\frac{1}{4a_{\sigma}^2}\right) - 1 \quad (11)$$

and

$$c_{\sigma} = 2a_{\sigma}^2 b_{\sigma} + \frac{1}{2} \quad (12)$$

For the short-range component

$$a_{\sigma} = \frac{\mu_{SR}}{6\sqrt{\pi}} \rho_{\sigma}^{-1/3} K_{\sigma}^{1/2} \quad (13)$$

and for the long-range component

$$a_{\sigma} = \frac{\mu_{LR}}{6\sqrt{\pi}} \rho_{\sigma}^{-1/3} K_{\sigma}^{1/2} \quad (14)$$

This functional is combined with the LYP correlation functional [33] to give the full exchange-correlation functional. There are four parameters introduced,  $C_{SHF}$ ,  $C_{LHF}$ ,  $\mu_{SR}$  and  $\mu_{LR}$ . These are optimized to minimize the mean absolute deviation (MAD) for a set of core excitation energies using the 6-311(2+, 2+)G\*\* basis set, yielding values of 0.50, 0.17,  $0.56a_0^{-1}$  and  $2.45a_0^{-1}$  for  $C_{SHF}$ ,  $C_{LHF}$ ,  $\mu_{SR}$  and  $\mu_{LR}$ , respectively, for the *K*-edge of first-row nuclei, and values of 0.87, 0.25,  $2.20a_0^{-1}$  and  $1.80a_0^{-1}$  for the *K*-edge of second-row nuclei. This functional is implemented within the Q-Chem software package [34].

The majority of time in a TDDFT calculation is spent in the evaluation of the integrals over the exchange-correlation functional and the two-electron integrals (Eq. 3). For hybrid functionals, additional exchange integrals are required, and for the SRC functional integrals involving the error function are also required. The default integration grid in Q-Chem is the SG-1 grid [35] which is a pruned Euler–Maclaurin–Lebedev (50,194) grid (i.e., 50 radial points, and 194 angular points per radial point). Within the Q-Chem integrals package, two-electron integrals are pre-screened and only evaluated if they satisfy conditions such as

$$\kappa_{ab}\kappa_{cd} \geq \tau \quad (15)$$

where

$$\kappa_{ab} = \sqrt{(ab|ab)} \quad (16)$$

and  $\tau$  is a threshold parameter that is typically set to  $1 \times 10^{-11}$ . Here, we explore reducing the quality of the integration grid and increasing the two-electron integral threshold in order to speed the calculation of NEXAFS spectra. A test set of excitations comprising the lowest eight states for the molecules and *K*-edges shown in Fig. 1, to give a total of 48 excitation energies, was used to assess the different integration grids and integral thresholds. The structure of the molecules was optimized using Møller–Plesset perturbation theory (MP2) [36] with the cc-pVTZ basis set [37, 38], and the SRC1 exchange-correlation functional [23] in conjunction with the 6-311G\* basis set was used to compute the NEXAFS spectra. Table 1 shows the numerical integration grids and notation used in this work. These grids use atom-centered Euler–Maclaurin radial grids with Lebedev angular grids. The use of different integration grids for the Kohn–Sham and TDDFT parts of the calculation is considered. This is denoted GX/GY, where GX is the grid used in the Kohn–Sham calculation and GY is the grid used for the TDDFT part of the calculation. Subsequently, the NEXAFS spectra for the larger systems are computed and compared with experiment; for these calculations, the structures were optimized using B3LYP/6-31G\*. Computed spectra are generated by convoluting the computed excitation energies

and intensities with gaussian functions with a full width at half maximum of 0.4 eV.

### 3 Results and discussion

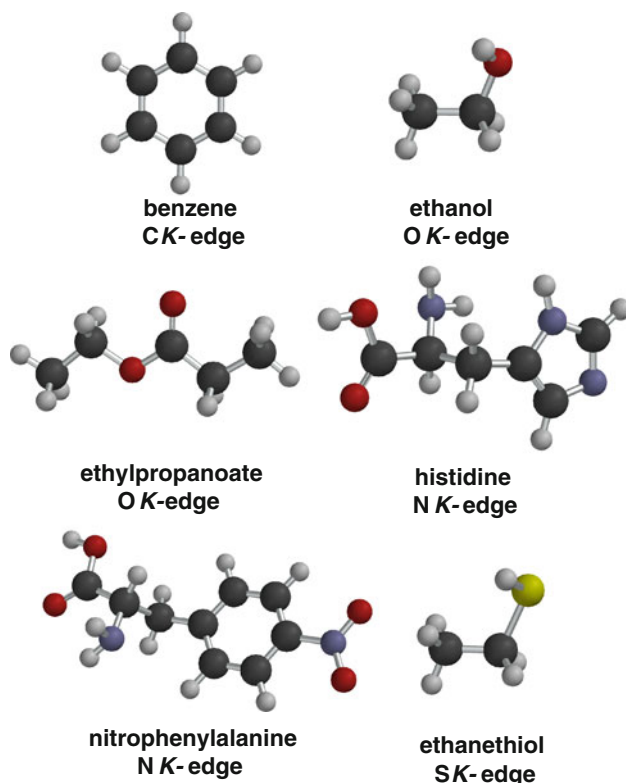
#### 3.1 Variation of integration grid and integral threshold

Table 2 summarizes the effect on the excitation energy of reducing the quality of the integration grid. If the size of the integration grid is reduced in both the Kohn-Sham and TDDFT parts of the calculation, the error in the computed excitation energies, compared to the full G1 grid, soon grow. For the G3/G3 grid, a maximum error of 0.2 eV is observed, and for the G4/G4 grid, there is a maximum error of over 1 eV. However, if the larger integration grid is maintained for the Kohn-Sham calculation, and the quality of the integration grid is reduced just for the TDDFT part of the calculation, the effect on the computed excitation energies is much smaller. For the G1/G4 grid, only a very small error introduced. This is surprising since the (10,18) grid is very small compared to the usual quality integration grids used in DFT calculations. Even reducing the size of the grid further does not lead to errors that are particularly large. One reason for this behavior is that for core

excitations, the elements in **A** are dominated by the energy difference between the Kohn-Sham orbital energies (Eq. 2).

For TDDFT calculations with the SRC1 exchange-correlation functional, two-electron integrals involving the error function and the Coulomb operator are required. For large systems, the number of these integrals grow rapidly and can become the most time-consuming part of the calculation. The number of these integrals can be reduced by increasing the threshold used for the two-electron screening. By default, this threshold is  $1 \times 10^{-11}$  a.u. Table 2 also shows the effect of increasing this value to  $1 \times 10^{-4}$  and  $1 \times 10^{-3}$  a.u. in conjunction with the G1/G4 DFT integration grid. Again, this threshold is only changed in the TDDFT part of the calculation and is kept at the default value in the Kohn-Sham part of the calculation. Even with a threshold of  $1 \times 10^{-3}$  a.u., there is little additional error, and the maximum error observed is insignificant compared to the overall accuracy of the calculations. To simulate NEXAFS spectra, the predicted oscillator strength is also important. For the test set of excitations, the difference in the predicted oscillator strengths for the G1/G1 and G1/G4 with a threshold of  $1 \times 10^{-3}$  a.u. is less than 0.001 for all of the excitations.

Table 3 shows the time in seconds for the TDDFT calculation for the series of linear polyaromatic hydrocarbons benzene  $\rightarrow$  pentacene for a standard calculation (G1/G1 grid, threshold  $1 \times 10^{-11}$  a.u.) and a faster calculation (G1/G4 grid, threshold  $1 \times 10^{-3}$  a.u.). For these calculations, the lowest 25 roots were computed using the SRC1 functional with 6-311G\* basis set using a single processor. It should be noted that computational savings are only gained in the TDDFT part of the calculation and that the Kohn-Sham DFT calculation will be the same in both case. However, for large systems where many roots are required, the TDDFT calculation is much more expensive than the Kohn-Sham DFT calculation. For pentacene, the TDDFT calculation for 25 roots is approximately eight times longer than the preceding Kohn-Sham DFT calculation, and for most NEXAFS calculations, considerably, more than 25 roots are required. For these calculations, the time for the faster calculations is about 25–30 % of the standard calculation, which represents a significant saving. One reason for the reduction in the computational time not being greater is that the time for diagonalization of **A** becomes significant, and this is not affected by the approximations considered here. However, it may be possible to reduce the time for the calculations further by imposing additional restrictions on the size of the virtual orbital space. Another important factor of using a larger integral threshold and a smaller integration grid is that the memory required by the calculations is significantly reduced, and for larger systems, even within a direct implementation, this can be a



**Fig. 1** Molecules and edges used to test the grid quality and integral threshold

**Table 1** Numerical integration grids used

Label	Radial	Angular	Total points
G1	50	194	9,700
G2	30	38	1,140
G3	15	38	570
G4	10	18	180
G5	5	6	30

**Table 2** Error in the computed excitation energies (in eV) relative to the SG1 grid for various grid combinations

Grid	Max. error	MAD <sup>a</sup>
G2/G2	0.047	0.016
G3/G3	0.205	0.060
G4/G4	1.021	0.374
G1/G2	0.001	0.000
G1/G3	0.008	0.002
G1/G4	0.016	0.001
G1/G5	0.034	0.012
G1/G4 ( $\tau = 4$ )	0.016	0.002
G1/G4 ( $\tau = 3$ )	0.020	0.006

<sup>a</sup> Mean absolute deviation

**Table 3** Time in seconds for the TDDFT calculation for the calculation of the lowest 25 states using a 2.0GHz Intel Xenon processor

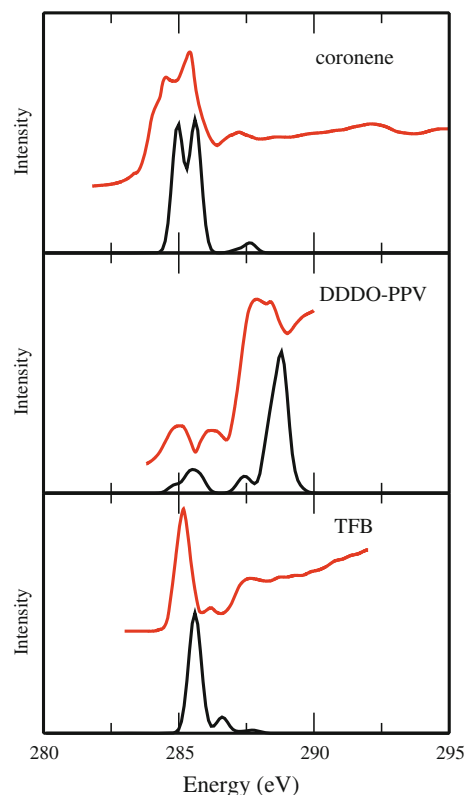
Molecule	Standard calculation	G1/G4 ( $\tau = 3$ )
Benzene	496	143
Naphthalene	1,528	478
Anthracene	4,203	1,249
Tetracene	9,550	2,618
Pentacene	15,771	4,300

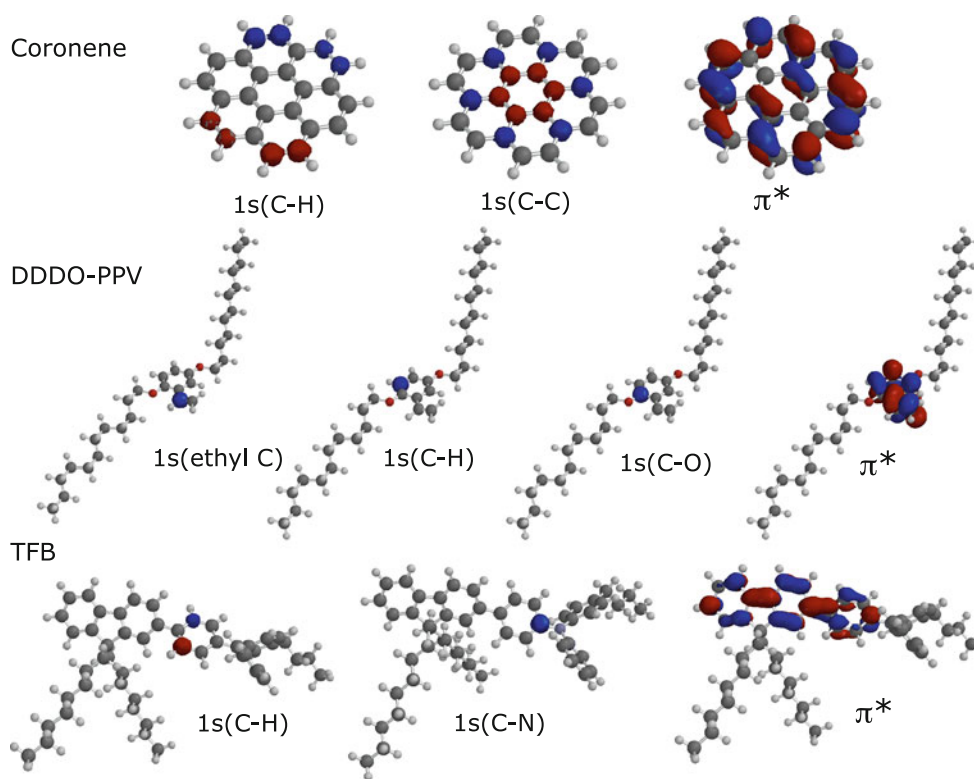
key factor in making the calculations tractable. For naphthalene, the number of two-electron integrals is reduced by 50 %, and for anthracene, this grows to over 60 % and continues to grow to over 70 % for pentacene.

### 3.2 Large molecules

Figure 2 shows the computed and experimental carbon *K*-edge NEXAFS spectra for coronene and two conjugated polymers poly[2,5-didodecyloxy-1,4-phenylenevinylene] (DDDO-PPV) and poly[(9,9-dioctylfluorenyl-2,7-diyl)-co-(4,4'-(4-sec-butyl phenyl)diphenylamine)] (TFB). For the calculations of the polymers, only one repeating unit is considered with the ends capped with hydrogen, that is, molecular formulae  $C_{32}O_2H_{36}$  and  $C_{51}NH_{63}$  (see Fig. 3).

For coronene, the experimental spectrum has a broad band with two distinct peaks at 284.5 and 285.4 eV and a smaller band at 287.2 eV. The theoretical spectrum includes 100 excited states and reproduces the features of the experimental spectrum well, with two intense bands predicted at 285.0 and 285.6 eV, with a weak band at 287.6 eV. While the theoretical values are slightly higher than the experiment, they remain within the typical range of errors observed with the SRC-type functionals. The calculations allow the origin of the spectral bands to be examined in more detail. The intense peaks correspond to excitation to the lowest unoccupied orbitals which are two degenerate  $\pi^*$  orbitals. Within the DFT calculations, the core orbitals comprise linear combinations of the individual carbon 1s orbitals. Excitations from a number of these core molecular orbitals contribute to the peaks observed. However, the two peaks can be identified to arise from excitations from core orbitals associated with different types of carbon atom. The lower energy peak computed at 285.0 eV corresponds to core molecular orbitals on the carbon atoms bonded to hydrogen, and the higher energy peak computed at 285.6 eV corresponds to core molecular orbitals on the carbon atoms not bonded to hydrogen. Representative orbitals that lead to the two different peaks are shown in Fig. 3. This is in agreement with the findings

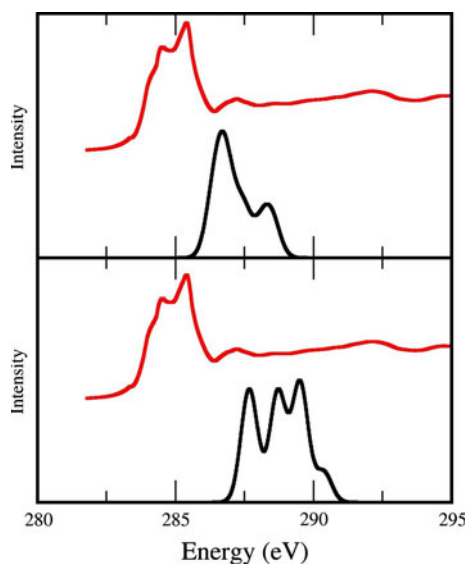
**Fig. 2** Calculated (*lower line*) and experimental (*upper line*) NEXAFS spectra, experimental data adapted from references [39, 40]



**Fig. 3** Relevant molecular orbitals for the NEXAFS of coronene, DDDO-PPV and TFB

of previous work [39]. The weaker band at 287.6 eV arises from excitation to a higher energy  $\pi^*$  orbital.

For the polymers, the experimental spectrum for DDDO-PPV has two weaker peaks at 285.0 and 286.2 eV



**Fig. 4** Calculated (*lower line*) and experimental (*upper line*) NEXAFS spectra, experimental spectrum adapted from reference [39]. *Upper panel* neglect of coupling between different core orbitals, *lower panel*  $K = 0$

and an intense peak at 288.1 eV. The computed spectrum reproduces the spectral profile well, with bands calculated to lie at 285.5, 287.4 and 288.8 eV. The spectrum for TFB is considerably different with an intense band at 285.2 eV with a weaker band at higher energy, 286.2 eV. Again the calculated spectrum reproduces the spectrum well with bands predicted at 285.6 and 286.6 eV. All of the calculated excitation energies for these systems are a little too high. However, it should be noted that the calculations consider one repeating unit in the gas phase and are not directly comparable to the experiments which correspond to a thin film of the polymer. The calculations do provide a basis for rationalizing the spectral features observed. The bands in the DDDO-PPV spectrum arise from excitations to the  $\pi^*$  orbital localized on the ring, and the three peaks correspond to excitations from the core molecular orbitals associated with three different types of carbon atom. The lowest energy peaks corresponds to the ethyl carbon bonded to the ring, the second peak to the carbon atoms of the ring bonded to hydrogen and the highest energy peak to the carbon atoms in ring bonded to oxygen. Representative orbitals are illustrated in Fig. 3. Similarly, for TFB, the bands also correspond to excitations to  $\pi^*$  orbitals, with the lower energy band arising from core molecular orbitals associated with the aromatic carbon atoms bonded to hydrogen and the smaller feature at higher energy to the core orbitals of the three carbon atoms bonded to nitrogen.

Coronene provides a good system to examine additional approximations in the calculation of NEXAFS spectra within a TDDFT framework that lead to further computational savings. Figure 4 shows two additional computed spectra for coronene. In the first of these spectra, the excitations from each of the 24 molecular orbitals arising from the carbon 1s orbitals are considered separately, and in the second spectrum, the coupling matrix  $K$  is neglected and  $\mathbf{A}$  is simply taken to be

$$A_{ia\sigma,jb\tau} = \delta_{ij}\delta_{ab}\delta_{\sigma\tau}(\epsilon_a - \epsilon_i). \quad (17)$$

Neglecting  $K$  leads to a large shift to higher energy, and the distinct two peaks observed in the intense  $\pi^*$  band are not reproduced, and overall, there is a poor agreement with the spectrum from experiment. Neglecting the coupling between the excitations from the different core orbitals leads to a spectrum that is in better agreement with experiment, although clearly poorer than the fully coupled TDDFT spectrum. While the overall shape of the experimental spectrum is reproduced, there is a small shift to higher energy, and the relative intensities of the bands are not as accurate as the fully coupled TDDFT spectrum. However, this approximation does make the cost of the calculations linear with respect to the number of core orbitals from which excitations are being considered and can be readily applied to much larger systems.

#### 4 Conclusions

Calculations of NEXAFS spectra are an important tool to aid the interpretation of experimental data. TDDFT in conjunction with short-range corrected functionals provide an accurate approach for the computation of NEXAFS at the  $K$ -edge. However, these calculations become computationally expensive for large systems, particularly when excitations from a large number of core orbitals are required, for example at the carbon  $K$ -edge in organic molecules. We have shown that the cost of such calculations can be reduced significantly through the use of a coarse numerical integration grid and large two-electron integral threshold within the TDDFT calculation, with a negligible error being introduced. Calculations on coronene and two polymers, DDDO-PPV and TFB, illustrate the approach and allow the origin of the spectra bands observed in experiment to be assigned.

**Acknowledgments** The authors would like to thank the University of Nottingham for access to its High Performance Computing facility.

#### References

1. Stöhr J (1996) NEXAFS spectroscopy, Springer series in surface science. Springer, Heidelberg
2. Nilsson A, Pettersson LGM (2004) Surf Sci Rep 55:49
3. DuBois JL, Mukherjee PM, Stack TDP, Hedman B, Solomon EI, Hodgson KO (2000) J Am Chem Soc 122:5775
4. Solomon EI, Szilagyí RK, George SD, Basumallick L (2004) Chem Rev 104:419
5. Penner-Hahn JE (2005) Coord Chem Rev 249:161
6. Wernet P, Nordlund D, Bergmann U, Cavalleri M, Ogasawara H, Näslund LA, Hirsch TK, Ojamae L, Glatzel P, Pettersson LGM, Nilsson A (2004) Science 304:995
7. Hunt WJ, Goddard WA III (1969) Chem Phys Lett 3:414
8. Ågren H, Carravetta V, Vahtras O, Pettersson LGM (1994) Chem Phys Lett 222:75
9. Ågren H, Carravetta V, Vahtras O, Pettersson LGM (1997) Theor Chem Acta 97:14
10. Stener M, Lisini A, Decleva P (1995) Chem Phys 191:141
11. Triguero L, Pettersson LGM, Ågren H (1998) Phys Rev B 58:8097
12. Hsu H, Davidson ER, Pitzer RM (1976) J Chem Phys 65:609
13. Navesde Brito A, Correlá N, Svensson S, Ågren H (1991) J Chem Phys 95:2965
14. Gilbert ATB, Besley NA, Gill PMW (2008) J Phys Chem A 112:13171
15. Besley NA, Gilbert ATB, Gill PMW (2009) J Chem Phys 130:124308
16. Besley NA, Asmuruf FA (2009) Phys Chem Chem Phys 12:12024
17. Imamura Y, Otsuka T, Nakai H (2006) J Comp Chem 28:2067
18. Nakata A, Imamura Y, Ostuka T, Nakai H (2006) J Chem Phys 124:094105
19. Nakata A, Imamura Y, Nakai H (2006) J Chem Phys 125:064109
20. Nakata A, Imamura Y, Nakai H (2007) J Chem Theory Comput 3:1295
21. Besley NA, Noble A (2007) J Phys Chem C 111:3333
22. Song J-W, Watson MA, Nakata A, Hirao K (2008) J Chem Phys 129:184113
23. Besley NA, Peach MJG, Tozer DJ (2009) Phys Chem Chem Phys 11:10350
24. Hirata S, Head-Gordon M (1999) Chem Phys Lett 314:291
25. Davidson DR (1975) J Computat Phys 17:87
26. Tsuchimochi T, Kobayashi M, Nakata A, Imamura Y, Nakai H (2008) J Comput Chem 29:2311
27. Ekström U, Norman P (2006) Phys Rev A 74:042722
28. Ekström U, Norman P, Carravetta V, Ågren H (2006) Phys Rev Lett 97:143001
29. Stener M, Fronzoni G, de Simone M (2003) Chem Phys Lett 373:115
30. George SD, Petrenko T, Neese F (2008) Inorg Chim Acta 361:965
31. Asmuruf FA, Besley NA (2008) J Chem Phys 129:064705
32. Song J-W, Hirose T, Tsuneda T, Hirao K (2007) J Chem Phys 126:154105
33. Lee C, Yang W, Parr RG (1998) Phys Rev B 57:785
34. Shao Y, Molnar LF, Jung Y, Kussmann J, Ochsenfeld C, Brown ST, Gilbert ATB, Slipchenko LV, Levchenko SV, O'Neill DP, DiStasio RA Jr, Lochan RC, Wang T, Beran GJO, Besley NA, Herbert JM, Lin CY, Voorhis TV, Chien S-H, Sodt A, Steele RP, Rassolov VA, Maslen PE, Korambath PP, Adamson RD, Austin B, Baker J, Byrd EFC, Dachsel H, Doerksen RJ, Dreuw A, Dunietz BD, Dutoi AD, Furlani TR, Gwaltney SR, Heyden A, Hirata S, Hsu C-P, Kedziora G, Khalliulin RZ, Klunzinger P, Lee AM, Lee MS, Liang W, Lotan I, Nair N, Peters B, Proynov EI, Pieniazek PA, Rhee YM, Ritchie J, Rosta E, Sherrill CD, Simmonett AC, Subotnik JE, Woodcock HL III, Zhang W, Bell AT, Chakraborty AK, Chipman DM, Keil FJ, Warshel A, Hehre WJ, Schaefer HF III, Kong J, Krylov AI, Gill PMW, Head-Gordon M (2006) Phys Chem Chem Phys 8:3172

35. Gill PMW, Johnson BG, Pople JA (1993) *Chem Phys Lett* 209: 506
36. Møller C, Plesset MS (1934) *Phys Rev* 46:618
37. Dunning TH Jr (1989) *J Chem Phys* 90:1007
38. Woon DE, Dunning TH Jr (1993) *J Chem Phys* 98:1358
39. Oji H, Mitsumoto R, Ito E, Ishii H, Ouchi Y, Seki K, Yokoyama T, Ohta T, Kosugi N (1998) *J Chem Phys* 109:10409
40. Watts B, Swaraj S, Nordlund D, Lüning J, Ade H (2011) *J Chem Phys* 134:024702