Cost and sensitivity of restricted active space calculations of metal L-edge X-ray absorption spectra

Rahul V. Pinjari[1,‖], Mickaël G. Delcey[1], Meiyuan Guo[1], Michael Odelius[2], and Marcus Lundberg[1,**]

[1]Department of Chemistry - Ångström laboratory, Uppsala University, SE-751 20 Uppsala, Sweden.
[2]Department of Physics, Stockholm University, AlbaNova University Center, SE-106 91 Stockholm, Sweden.

October 7, 2015

Abstract

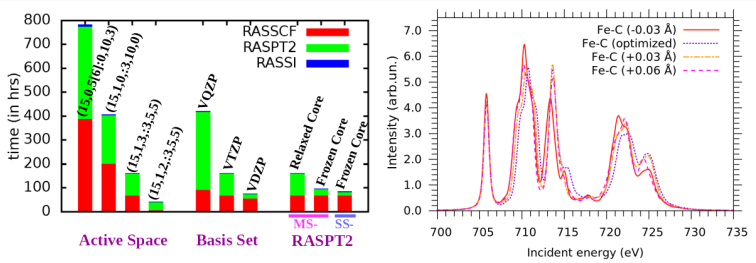
The restricted active space (RAS) approach can accurately simulate metal L-edge X-ray absorption spectra of first-row transition metal complexes without the use of any fitting parameters. These characteristics provide a unique capability to identify unknown chemical species and to analyze their electronic structure. To find the best balance between cost and accuracy, the sensitivity of the simulated spectra with respect to the method variables have been tested for two models, $[FeCl_6]^{3-}$ and $[Fe(CN)_6]^{3-}$. For these systems the reference calculations give deviations, compared to experiment, of $\leqslant$1 eV in peak positions, $\leqslant$30% for the relative intensity of major peaks, and $\leqslant$50% for minor peaks. Compared to these deviations, the simulated spectra are sensitive to the number of final states, the inclusion of dynamical correlation and the ionization potential-electron affinity (IPEA) shift, in addition to the selection of the active space. The spectra are less sensitive to the quality of the basis set and even a double-ζ basis gives reasonable results. Inclusion of dynamical correlation through second-order perturbation theory can be done efficiently using the state-specific formalism without correlating the core orbitals. Although these observations are not directly transferable to other systems, they can, together with a cost analysis, aid in the design of RAS models and help extend the use of this powerful approach to a wider range of transition-metal systems.

Keywords: transition metals, X-ray absorption spectroscopy, multiconfigurational wavefunction, spin-orbit coupling, charge transfer

‖Present address: School of Chemical Sciences, Swami Ramanand Teerth Marathwada University, Nanded 431606, India

**Corresponding author: marcus.lundberg@kemi.uu.se

TABLE OF CONTENTS



With an appropriate choice of active space, basis set and computational procedure the restricted active space approach can be used to simulate metal L-edge X-ray absorption spectra with reasonable accuracy and computational cost. The sensitivity of the simulated results with respect to geometrical changes opens up for analysis of dynamical processes.

INTRODUCTION

Metal L-edge (2p $\rightarrow$ 3d) X-ray absorption spectroscopy (XAS) is a powerful technique to study first-row transition metal systems. It directly targets the unoccupied or partially unoccupied 3d orbitals involved in electron transfer during homogeneous and heterogeneous catalysis. However, due to the complex electronic structure of the final states, determined by electron-electron repulsion and spin-orbit coupling (SOC) in 2p and 3d orbitals, advanced theoretical methods are required to correlate the spectral shape with the electronic structure of the system.

A variety of methods have been designed to model L-edge XAS spectra viz. the Bethe-Salpeter equation,[1,2] the multichannel multiple scattering method,[3,4] and different configuration interaction (CI) methods.[5–14] The CI-based methods differ in the way they describe electron correlation, in the selection of the electronic configurations, and in the treatment of SOC.

The restricted active space (RAS) method is an *ab initio* method based on the multiconfigurational self-consistent field (MCSCF) approach. In the complete active space (CAS) method,[15,16] a full CI is performed among the active orbitals, but for simulations of X-ray spectra it is more convenient to use the restricted active space (RAS) method,[17] and restrict the number of excitations from the core orbitals to one.[14,18–27] Results can be improved further by including correlation with the occupied inactive and empty virtual orbitals using second-order perturbation theory (PT2).[28–31] SOC is then added as a perturbation on top of the multiconfigurational spin-free states.[32]

The RAS approach has been used to calculate both L-edge XAS and resonant inelastic X-ray scattering (RIXS) spectra of transition-metal coordination complexes,[14,20–24,33] as well as X-ray photoelectron spectra (XPS) of Uranium and Ytterbium compounds.[18] To evaluate the applicability of the approach, it was recently applied to a series of ferric systems; Fe^{3+} ions with varying ligand fields, the high-spin complex $[FeCl_6]^{3-}$ (ferric chloride) and the low-spin complex $[Fe(CN)_6]^{3-}$ (ferricyanide).[26] All these systems have $3d^5$ electron configurations in the ground state. This gives the largest number of possible valence excitations for a transition metal complex, and they should therefore be among the most challenging systems

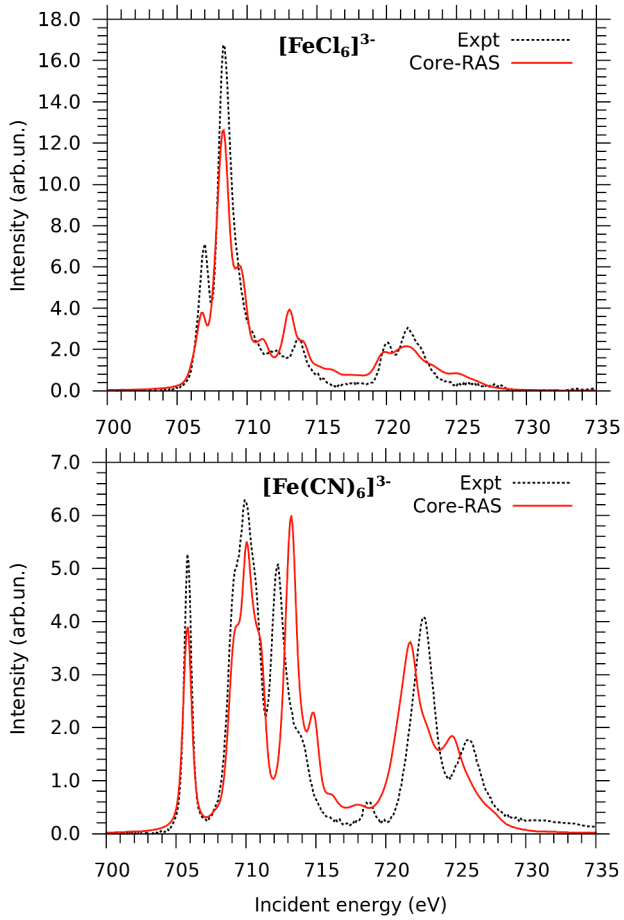


Figure 1: The experimental (black-dotted line) L-edge XAS spectra of $[FeCl_6]^{3-}$ (Wasinger *et al.*[34]) and $[Fe(CN)_6]^{3-}$ (Hocking *et al.*[35]) compared to those calculated from RASPT2-cc(15,1,0;3,10,0)/ANO-RCC-VTZP with 120 states (red-solid line).

to describe. For these three systems the accuracy of the RAS approach is comparable to that of the semi-empirical charge-transfer multiplet (CTM) model.[6,7] Compared to experiment, the RAS simulations include all the major and minor peaks with errors in energies of less than 1.0 eV, see Figure 1. These are encouraging results, especially as they are obtained without any fitting parameters.

The combination of reasonable accuracy with an *ab initio* approach provides a powerful tool for the analysis of X-ray spectra. It can be used to analyze the electronic structure of transition metal complexes without *a priori* assumptions, both at distorted ground state geometries and in excited states. The possibility to model both L-edge XAS and RIXS with an

accurate description of chemical bonding opens up for many applications in material science, biochemistry and photochemistry. The major limitation is the relatively high computational cost and it is therefore essential to find good compromises between accuracy and timing.

The RAS method is relatively flexible in the model design, e.g., the choice of active space, the number of final states, the size of the basis set, the system geometry, and the choice of RAS procedure. The sensitivity to these variables can be evaluated relative to the difference between the experiments and the reference simulations. Effects of the active-space selection on the L-edge XAS spectra of ferric chloride and ferricyanide have already been analyzed,[26] but none of the other variables have been systematically investigated, neither in terms of sensitivity nor the computational timing. As both systems have well-defined ground states, it is not clear how transferable the observations will be e.g., for systems with low-lying valence excited states. However, detailed initial analysis of these effects should aid the design of more efficient core RAS models and help to extend the approach to larger systems in the future.

METHODOLOGY

Calculations of $[FeCl_6]^{3-}$ have been performed using an O_h symmetric structure with Fe−Cl distances of 2.40 Å taken from X-ray crystallography.[26] The $[Fe(CN)_6]^{3-}$ geometry is obtained through a CASPT2 optimization, with four Fe−C distances of 1.939 Å and two of 1.916 Å; the C−N distances are 1.178 (1.180) Å.[26] The Jahn-Teller distortion to D_{4h} has only a minor effect on the results and for simplicity the orbitals are discussed using terms from the O_h point group.[26] All calculations are effectively performed in the highest symmetry Abelian point group (D_{2h}). To take full advantage of the high symmetry, counter ions were not included, leading to anionic model systems. The potential challenges of modeling highly anionic systems are not expected to significantly affect the sensitivity of the results with the respect to the different parameters.

All RASSCF/RASPT2 calculations have been performed with MOLCAS 7.9.[36] The active space is designated RAS(n,l,m;i,j,k), where, i, j and k are the number of orbitals in the RAS1, RAS2 and RAS3 spaces respectively, n is the total number of electrons in the active space, l

	RAS1	RAS2	RAS3
Ferric chloride			
RAS(15,1,0:3,10,0)	2p orbitals with ≤ 1 hole	σ, t_{2g}, e_g and t'_{2g} orbitals	—
Ferricyanide			
RAS(9,0,0:0,10,0)	—	σ, t_{2g}, e_g and t'_{2g} orbitals	—
RAS(15,1,0:3,10,0)	2p orbitals with ≤ 1 hole	σ, t_{2g}, e_g and t'_{2g}/π^* orbitals	—
RAS(15,1,3:3,5,5)	2p orbitals with ≤ 1 hole	σ and t_{2g} orbitals	e_g and t'_{2g}/π^* orbitals with ≤ 3 electrons
RAS(15,1,2:3,5,5)	2p orbitals with ≤ 1 hole	σ and t_{2g} orbitals	e_g and t'_{2g}/π^* orbitals with ≤ 2 electrons
RAS(15,0,5[6]:0,10,3)	—	σ, t_{2g}, e_g and t'_{2g}/π^*	2p orbitals with ≤ 5 (or 6) electrons

Table 1: Description of the active spaces used in the calculations. Orbital diagrams are shown in Figure SI 1 and orbital pictures are available in reference 26.

the maximum number of holes allowed in RAS1, and m the maximum number of electrons in RAS3. The default RAS2 space consists of ten orbitals of metal 3d character, see Figure SI 1. The shape and character of these orbitals are described in detail in reference[26]. The RAS1 space contains the 2p orbitals, allowing a maximum of one hole, giving the (15,1,0;3,10,0) active space. Detailed descriptions of the different active spaces used in the present study are given in Table 1. As the RAS2 orbitals all have "gerade" symmetry in the D_{2h} point group, the 2p core excited states are easy to reach as they are simply the lowest states with "ungerade" symmetry. RASSCF orbital optimizations have been performed using the state average (SA) formalism, i.e., all the final states of a specific spin and symmetry are calculated using the same set of orbitals. To ensure that the RAS1 hole stays in the low-lying metal 2p orbitals, rotations between these and other occupied orbitals are not allowed in the RASSCF optimizations. A relatively large level shift of 1.0 hartree has been applied to achieve convergence for all different combinations of RAS spaces and number of states.

For systems that lack centrosymmetry, core and valence excited states appear in the same irreducible representations. To avoid calculating a large number of states one possible solution is to put the 2p orbitals in the RAS3 space and allowing a maximum of six electrons in the ground state and five electrons in the excited state, RAS(15,0,5[6]:0,10,3). This makes the single core excited states lowest in energy regardless of the symmetry, but the drawback is that it adds configuration state functions with two to six core holes. As these new configurations have a very high energy compared to those with a single core excitation,

they do not have any significant effect on the relative energies of the relevant final states. Despite the increase in computational cost associated with the additional configurations, this can be beneficial compared to the RAS1 approach that could require a very large number of states.

By default calculations are performed by including scalar relativistic effects by using a Douglas-Kroll (DK) Hamiltonian[37,38] in combination with a relativistic atomic natural orbital basis set of triple-zeta quality (ANO-RCC-VTZP).[39,40] Due to an error in the ANO-RCC basis set definition in Molcas, calculations have been performed with carbon atoms using coefficients optimized for nitrogen. Changing to the correct basis set only has a small effect, see Figure SI 2, and the method sensitivity should not be affected by the basis set problem. A density-fitting approximation has been made for the electron repulsion integrals.[41–43] SOC is included from a one-electron spin-orbit Hamiltonian based on atomic mean field integrals.[44] The spin-orbit eigenstates are obtained by diagonalizing the spin-orbit interaction matrix.[32] These states are used to calculate the strength of the electric dipole transitions with the restricted active space state interaction (RASSI) approach.[45,46]

All possible $(2p)^5(3d)^6$ spin multiplicities (doublet, quartet and sextet) have been included in the calculations. For systems where the ground state is not a mixture of spin states, excited states with $\Delta S = \pm 1$ are sufficient to reproduce the experimental spectra. For ferricyanide, which has a doublet ground state, doublet and quartet excited states are required, while sextet states have small effects on the final spectrum through their interactions with the quartets, see Figure SI 3. For ferric chloride, which has a sextet ground state, doublet states have only small effects. Octet states would require both an excitation from 2p to an orbital above the metal-centered t_{2g} and e_g, see Figure SI 1, as well as a spin-flip. They do not have any significant effect on the spectrum. Dectet states would require an additional excitation with associated spin flip and can be ignored.

Unless stated otherwise, RASPT2 calculations are performed with multi-state (MS) RASPT2,[30] an imaginary shift of 0.3 hartree,[47] and the default value of the ionization-potential-electron-affinity (IPEA) shift (0.25 hartree).[48] The default option in the RASPT2 calculations is to freeze the deep core orbitals, i.e. they are not correlated. In this paper most calculations have instead been performed with correlated core orbitals, indicated by

RASPT2-cc.

For both systems the reference calculations have been performed at the RASPT2-cc (15,1,0;3,10,0)/ANO-RCC-VTZP level with 120 states per spin multiplicity and final state symmetry (A_{1u}, B_{1u}, B_{2u}, and B_{3u}). Considering four different symmetries and three different spin states, this represents 1440 spin-free final states. There are ten active valence orbitals, between which all possible excitations are allowed, in addition to the single excitation from the 2p orbitals. To save computational time, some comparisons for ferricyanide have been performed with RASPT2-cc(15,1,3;3,5,5) and 80 states. This calculation allows only three excitations to the e_g and π^* orbitals, which significantly reduces the number of configurations, while still reproducing the major features of the spectrum.[26]

In the experimental data sets the absorption has been normalized to an edge jump of 1.0.[34,35] The core RAS L-edge XAS spectra are obtained from the transition intensities using a Lorentzian broadening with a full-width-at-half-maximum (FWHM) of 0.4 and 0.8 eV respectively for the L_3 and L_2 edges,[49,50] and thereafter convoluted with an experimental Gaussian broadening of 0.4 eV.[26,35] The absorption intensity is scaled so that the integrated intensity matches the experimental value. The exception is the comparisons with different number of states, where the same scaling factor is used for all systems. The energies of the calculated spectra for ferric chloride are aligned to the most intense peak (708 eV) and the spectra for ferricyanide are aligned to the first intense transition (705.8 eV).

RESULTS AND DISCUSSION

Ferric chloride has a high-spin $(t_{2g})^3(e_g)^2$ ground state configuration (^6A_g), see Figure SI 1. The experimental L-edge XAS spectrum has a main peak at 708 eV preceded by a minor peak at 707 eV. Higher in the L_3 edge a resonance at 714 eV can also be distinguished, see Figure 1. The first two peaks includes excitations to t_{2g} and e_g orbitals. The feature at 714 eV is a shake-up transition, where the 2p $\rightarrow$ 3d transition is coupled to a ligand-to-metal charge transfer (LMCT), an assignment originally proposed from charge-transfer multiplet modeling.[34] The reference RAS calculation includes all these features although the intensity of the first feature is underestimated by 40 % and the energy of the shake-up feature is too

low by 1.0 eV, see Figure 1.

Ferricyanide has a low-spin $(t_{2g})^5(e_g)^0$ ground state configuration ($^2T_{2g}$), see Figure SI 1. The experimental L-edge spectrum has a narrow t_{2g} peak at 705.8 eV and a broad feature corresponding to e_g excitations, see Figure 1. The intense peak at 713 eV is a signature of backbonding, which borrows intensity from the main e_g feature. As the peak is more sensitive to the amount of π^* than σ^* backbonding, it is often labeled as a π^* feature.[35] The simulated core RAS spectrum overestimates the t_{2g}-e_g splitting by 0.2 eV and the relative energy of the π^* peak by 1.0 eV. The relative intensity of the π^* peak is also overestimated by 30%. Finally, the splitting between L_3 and L_2 edges is underestimated by 1.5 eV.

Number of states

Compared to calculations of valence-excited states, simulations of the L-edge XAS spectrum are more expensive, not only due to the addition of the 2p orbitals in the active space but also due to the large number of final states that must be included. When using a minimal valence active space with only the five metal-dominated orbitals, there are 450 spin-free states for the $2p^53d^6$ electron configuration and they can all be included in the calculations. For an active space with ten valence orbitals and nine electrons (ten in the core excited state), there are more than 300 000 possible spin-free states, and it is only possible to include a small subset of these states. It is not straightforward to determine how many states are required to get a converged L-edge spectrum. In addition, a correct application of the spin-orbit coupling scheme requires that all directly coupled states of different multiplicity are included in the calculations.

As an approximation, an ad hoc procedure has been chosen with an equal number of states for each symmetry and spin multiplicity. To analyze the sensitivity of the spectra they have been calculated with between 40 and 120 states for each combination of spin and symmetry. For $[FeCl_6]^{3-}$ there is no clear convergence with increasing number of states, see Figure 2. The overall shape of the L_3 edge is relatively constant, although the intensity of the high-energy feature of the main peak and the LMCT peak increases by 30%. However, the shape of the L_2 edge is very sensitive to the number of final states and the L_3-L_2 splitting

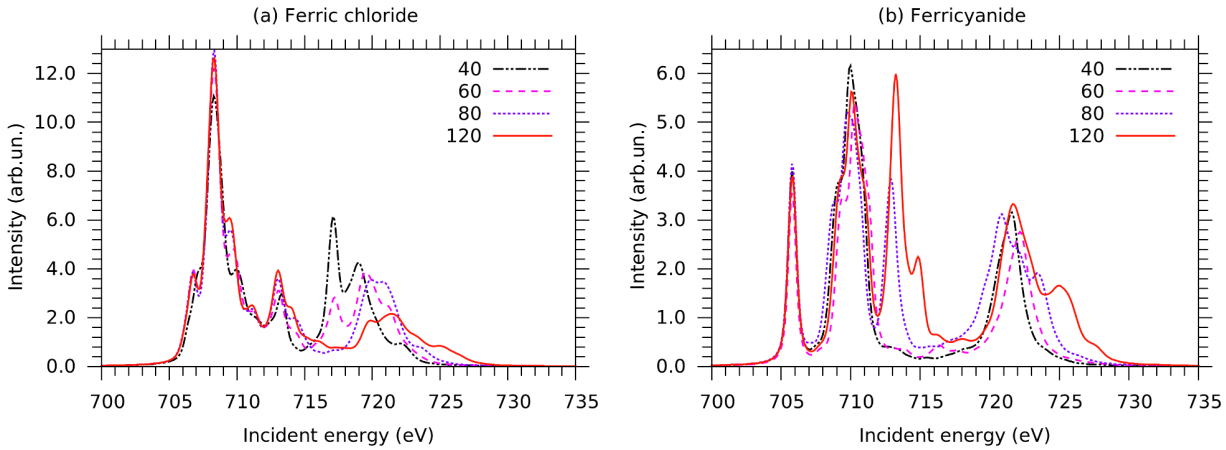


Figure 2: L-edge XAS spectra of (a) $[FeCl_6]^{3-}$ and (b)$[Fe(CN)_6]^{3-}$ calculated using RASPT2-cc/ANO-RCC-VTZP (15,1,0;3,10,0) with different numbers of states per symmetry and spin multiplicity.

increases by more than 2.0 eV when more states are included. Compared to experiment, increasing the number of states gives a slightly worse result for the L_3 edge while significantly improving the L_2 edge.

The effects are clearer for $[Fe(CN)_6]^{3-}$, see Figure 2. The spectra calculated with up to 60 states describe all the features of the t_{2g} and e_g peaks but lack the π^* peak in both edges. Increasing the number of states to 80 recovers this peak, albeit with relatively low intensity. Finally, with 120 states the intensity of the π^* peak increases and a high-energy shoulder appears.

The number of states can affect the calculations in several ways. In a SA-RASSCF calculation, the orbital optimization depends on the states that are included. To illustrate this, the $[Fe(CN)_6]^{3-}$ spectrum is calculated with the same number of final states (80) but using the orbitals optimized for 20, 40, 60 and 80 states respectively, see Figure 3a. In the ground state, the most stable RASSCF solution does not include the π^* orbitals, but rather a set of orbitals with metal 4d character that correlate with the 3d t_{2g} orbitals. For the π^* peak to appear, the RASSCF optimization must converge to a solution with the π^* orbitals in the active space instead. This requires that states with significant π^* character are included in the orbital optimization. Unfortunately, there are many d-d excitations with

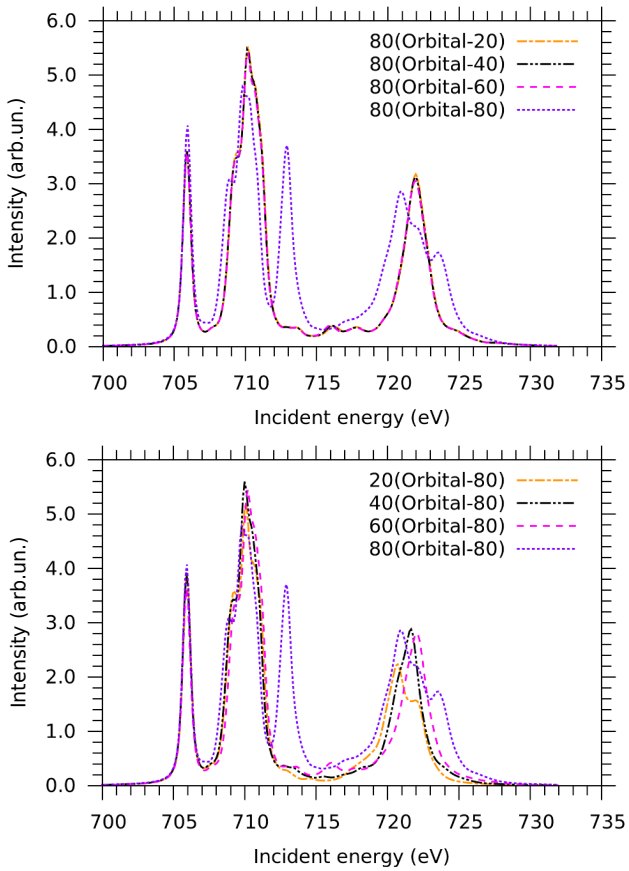


Figure 3: L-edge XAS spectra of $[FeCN_6]^{3-}$ calculated with RASPT2-cc/ANO-RCC-VTZP (15,1,0:3,10,0) with a) varying number of final states using orbitals optimized for 80 states, and b) 80 final states using orbitals optimized for a varying number of states. The number of states are reported per symmetry and spin state.

lower energy than the first π^* excitation, and only by including 80 final states the RASSCF converges to a solution that includes the π^* orbitals, see Figure 3. The large change in the active space affects other parts of the spectra as well and explains why both t_{2g}-e_g and the L_3-L_2 splittings are not smooth functions of the number of states.

Another important factor is the necessity to include all final states with significant transition intensity. To illustrate this, the $[Fe(CN)_6]^{3-}$ spectrum is calculated with different number of final states while using exactly the same orbitals (*viz.* those optimized for 80 final states), see Figure 3b. As expected from the orbital optimization, the π^* peak only appears when including 80 states. There are also large effects in the shape of the L_2 edge. The same

effect is seen in the $[FeCl_6]^{3-}$ simulations where the orbitals converge around 40 states but the shape of the L_2 edge does not converge even with 80 states, see Figure SI 4.

Compared to the number of active-space orbitals, where more orbitals should give more accurate calculations, it is less straightforward to determine the right number of final states. It is clear that a large number of states is required to describe all possible excitations. At the same time, including too many states can lead to suboptimal active spaces. An example is when these additional states include excitations to ligand orbitals that neither correlate with the metal orbitals nor have any significant transition intensity. Then orbitals that correlate with the 3d orbitals may be pushed out of the active space, which leads to a worse description of the 2p→3d excitations.

Selecting a fixed number of states for all symmetry-spin combinations does not lead to a formally correct calculation of the SOC for all states. As an example, to get correct SOC for all sextet states would, for combinatorial reasons, require a larger number of quartet than sextet states. Unfortunately, it is not practically feasible to perfectly match all states of different symmetry. The reason the current procedure still gives reasonable results is that the correct coupling is first achieved for the low-energy states, and these are the ones that mainly determine the spectral shape. However, in the L_2 edge the number of states of different spin multiplicity might not be enough to correctly describe the SOC, which contribute to the problems of converging that region.

Due to the large sensitivity of the results on the number of states, it is recommended to test the sensitivity of the results with respect to the number of states and to carefully monitor the shape of the active orbitals in the state-average optimization.

Basis set

The size and quality of the basis set is an an important factor in most electronic structure calculations. The basis set dependence of the L-edge XAS spectrum of $[FeCl_6]^{3-}$ was evaluated using ANO-RCC basis sets from double-ζ (VDZP) to quadruple-ζ on all atoms (VQZP), see Figure 4. Diffuse functions are included in the definition of these basis sets and no separate analysis have been made of the role of these functions. In the basis-set

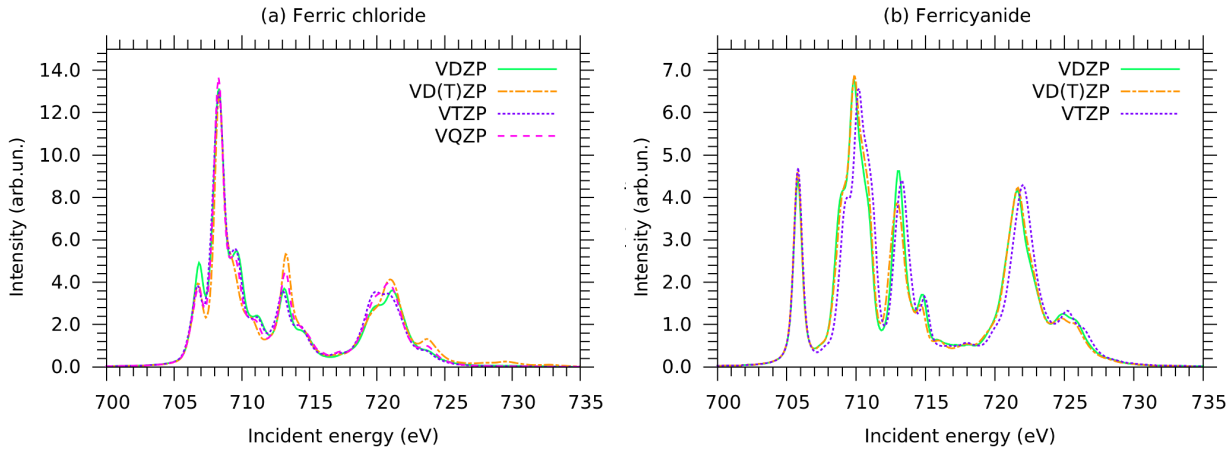


Figure 4: L-edge XAS spectra of (a) $[FeCl_6]^{3-}$ calculated with RASPT2-cc (15,1,0:3,10,0) and 80 final states (b) $[Fe(CN)_6]^{3-}$ calculated with RASPT2-cc (15,1,3:3,5,5) and 120 final states using different ANO-RCC basis sets, double-ζ (VDZP), triple-ζ on iron and double-ζ on all other atoms (VD(T)ZP), triple-ζ (VTZP), and quadruple-ζ (VQZP).

test, the relative intensities of the minor peaks change by up to 30%, but without any major changes in the relative energy. There is no clear convergence as the basis set increases up to VQZP, which is probably due to changes in the ordering of different states that in turn affects the orbital optimization procedure. In $[Fe(CN)_6]^{3-}$, going from a VDZP to a VTZP basis increases the t_{2g}-e_g splitting by 0.3 eV but there are no major changes in the relative intensity of the different peaks.

The observation that the results are not very sensitive to the choice of basis set might seem surprising considering the well-known basis-set sensitivity of correlated wavefunction methods, and the reason is the relatively low target accuracy. In studies of valence excitations, a change in relative energy of 0.3 eV between different states would be significant, while in this context the change is less significant.

Despite this relative stability, apparently large basis set effects appear in some cases. One example is a simulation of the ferricyanide spectrum with the smaller (15,1,3:3,5,5) active space. If only 80 states are included, the double ζ basis set gives a very weak π^* peak compared to the VTZP and VQZP results, see Figure SI 5.

In this case, it is the ordering of high-energy states of different character that is basis set

dependent. However, the effect on the spectrum would disappear with a sufficient number of states.

Core RAS procedure

The simulation results can also be affected by details of the RAS procedure. In the current calculations, the 2p orbitals are not re-optimized in order to ensure that the hole is confined to these orbitals. An alternative is to include a second step where instead the other occupied orbitals are frozen while the 2p orbitals are reoptimized together with the virtual orbitals. For $[FeCl_6]^{3-}$ this leads to increase in the intensity of the minor peak at 710 eV but no major changes in other parts of the spectrum, see Figure SI 6. The small effect could possibly due to a core orbital basis set that is not flexible enough to properly describe core relaxation.

An important decision is whether to perform the time-consuming RASPT2 step, that adds dynamical correlation, or stay at the RASSCF level. Spectra calculated with RASSCF and RASPT2 show clear differences, especially for $[Fe(CN)_6]^{3-}$, see Figure 5. With dynamical correlation, the energy difference between t_{2g} and e_g peaks decreases by >1 eV and the π^* peak is shifted to lower energy by >3 eV. For ferric chloride, the most obvious difference is a decrease of the transition energy of the shake-up transition by 1.5 eV. The relative intensities of different peaks are also affected by 25%.

The effects of dynamical correlation are equal to, or larger than, the differences between simulation and experiment. For $[Fe(CN)_6]^{3-}$ the RASPT2 step also leads to a significantly better agreement with experiment. As as an example, deviations in the relative position of the π^* peak, compared to experiment, is reduced from 4 eV in the RASSCF spectrum to 1 eV in the RASPT2 spectrum.

For simulations performed at the RASPT2 level the important task is to minimize the cost of the RASPT2 step without significant loss in accuracy. The RASPT2 calculations can be performed either as state-specific (SS) or MS-RASPT2. The latter is preferred in cases where several states of the same symmetry interact strongly.[30] However, with a very large number of states the SS-RASPT2 procedure has a lower computational cost. The calculations of individual states can also be performed separately, making this step trivially parallelizable.

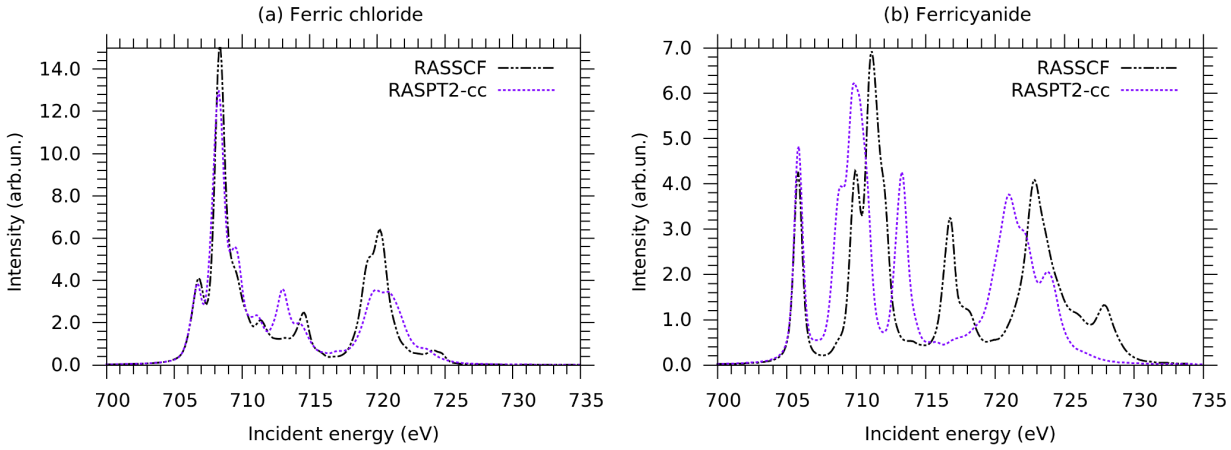


Figure 5: L-edge XAS spectra of $[FeCl_6]^{3-}$ (left) and $[Fe(CN)_6]^{3-}$ (right) at RASSCF and RASPT2-cc levels with a (15,1,0:3,10,0) RAS space and 80 final states.

Comparisons of SS and MS-RASPT2 spectra for both ferric chloride and ferricyanide are shown in Figure 6. In both complexes, the spectral differences are small. For ferric chloride, the intensity of the shake-up transition changes by 10% and in ferricyanide the position of the π^* peak differs by 0.2 eV. These relatively small changes suggest that the state-specific method can be used to decrease the time of heavy RASPT2 calculations.

In standard RASPT2 calculations the core orbitals are frozen. As the occupation of the Fe 2p orbitals changes during the X-ray absorption process, the current L-edge simulations have instead been performed with correlated core orbitals. A direct comparison between the two alternatives shows that the differences are small, both for ferric chloride and ferricyanide, with energy shifts of up to 0.2 eV, see Figure SI 7.

One common problem in RASPT2 calculations is the presence of intruder states that are artificially stabilized in the perturbation treatment. These states can be detected by a low weight of the reference RASSCF state in the RASPT2 calculation. A common method to reduce the problem of intruder states, in both ground and excited states, is to introduce an imaginary shift.[47] In the previous calculations an imaginary shift of 0.3 hartree was used for all spectra. Varying this shift between 0.15 and 0.50 hartree has small effects on the spectra, except for the energy of the π^* peak in ferricyanide that increases by 0.5 eV, see Figure SI 8. This suggests that problems with intruder states are limited in most regions of the current

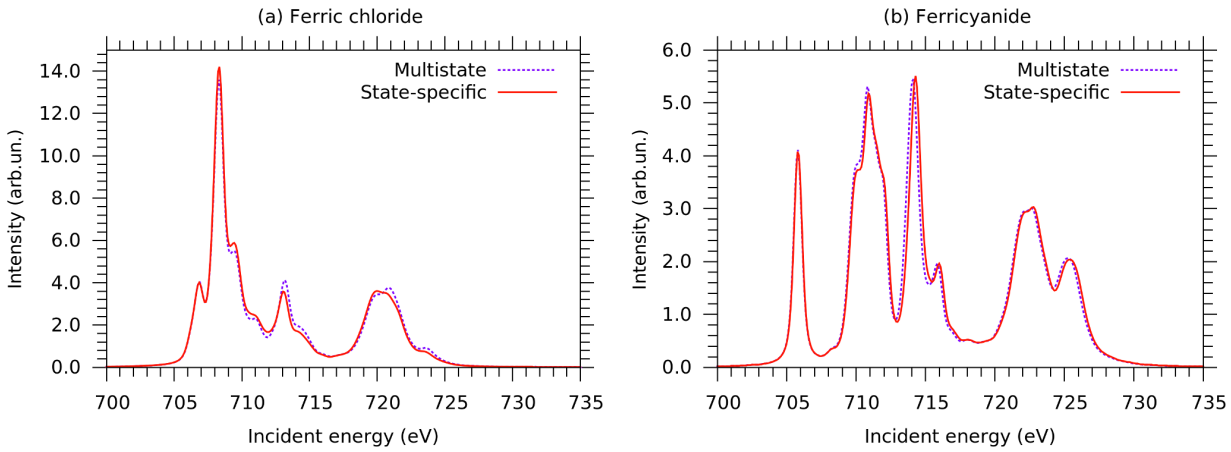


Figure 6: L-edge XAS spectra of $[FeCl_6]^{3-}$ (left) and $[Fe(CN)_6]^{3-}$ (right) calculated with state-specific (SS) and multi-state (MS) RASPT2 using (15,1,0:3,10,0) and (15,1,3:3,5,5) RAS spaces, respectively, and 80 final states.

spectra.

The RASPT2 algorithm has one global empirical parameter, the IPEA shift. This is a shift in the zeroth-order Hamiltonian that corrects for a systematic overestimation of the correlation energy in open shell systems.[48] A higher value of the IPEA shift stabilizes low-spin states relative to high-spin states and as the L-edge XAS spectrum is a mixture of at least two different spin states, the value of the IPEA shift could potentially affect the simulations.

The effects of increasing the IPEA shift from 0.00, through the default value of 0.25, up to 0.60 eV is shown in Figure 7. The effects are largest for ferricyanide where the energies of the e_g and π^* peaks increase by 0.5 eV and 1.5 eV respectively. A smaller IPEA shift would give a better agreement with experiment, although the intensity of the peak would remain too high. The reason for the large effect on the relative energies is that the t_{2g} transition gives a closed-shell t_{2g}^6 valence state while the e_g and π^* peaks have open-shell contributions that are destabilized by the IPEA shift.

It is not recommended to use the IPEA shift as an adjustable parameter to improve the agreement with experiment, but it is good practice to test the sensitivity of the results to this parameter. Large changes in the calculated spectra indicate that the results should be

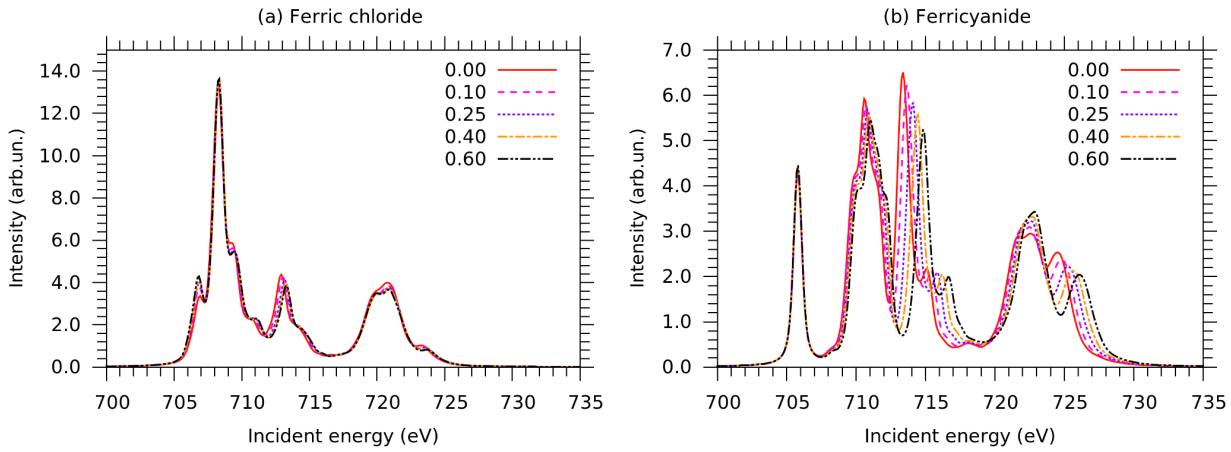


Figure 7: L-edge XAS spectra of $[FeCl_6]^{3-}$ (left) and $[Fe(CN)_6]^{3-}$ (right) with different IPEA shifts in the RASPT2 calculations using (15,1,0:3,10,0) and (15,1,3:3,5,5) RAS spaces, respectively, and 80 final states. The default IPEA shift is 0.25 eV.

treated with a healthy dose of scepticism.

Geometry

The core RAS results also depend on the geometries used in the simulations. Geometries can either be obtained from experiment, e.g., X-ray diffraction data, or, when reliable data are not available, predicted by theoretical calculations. For most transition-metal systems, geometries would be obtained from optimizations using density-functional theory. For $[Fe(CN)_6]^{3-}$ the X-ray structures give Fe−C distances of 1.941 Å on average.[35] Numerical optimization using CASPT2/ANO-RCC-VTZP (9,0,0;0,10,0) gives four Fe−C distances of 1.939 Å and two others of 1.916 Å, in good agreement with experiment, while B3LYP/ANO-RCC-VTZP gives bond distances that are ∼0.06 Å longer. To check whether differences in starting geometries of this magnitude have any effect on the final results, spectra were calculated with a range of Fe−ligand distances, see Figure 8.

For $[FeCl_6]^{3-}$ changes in Fe−Cl distance of ±0.03 Å mainly affects the intensity of the first minor peak, which changes by 30%, see Figure 8. For $[Fe(CN)_6]^{3-}$, distance changes of up to +0.06 Å have been included, matching the difference between CASPT2 and B3LYP optimized geometries. Larger iron-ligand distances are expected to show a decrease in the

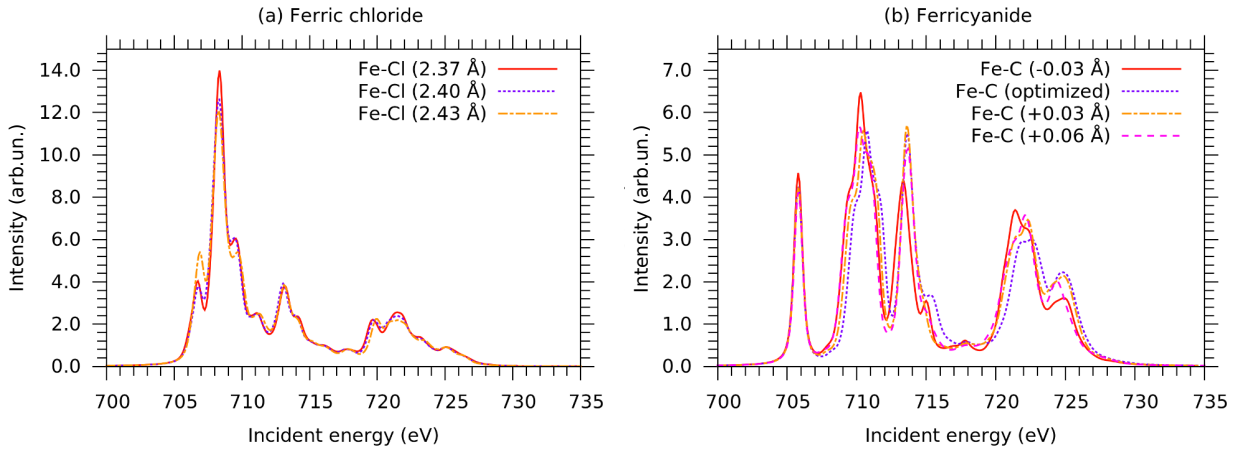


Figure 8: L-edge XAS spectra of $[FeCl_6]^{3-}$ (left) and $[Fe(CN)_6]^{3-}$ (right) calculated with RASPT2-cc/ANO-RCC-VTZP and 80 states using different Fe−L distances; the active space for these complexes are (15,1,0:3,10,0) and (15,1,3:3,5,5), respectively.

ligand-field, i.e., the t_{2g}-e_g splitting. This trend is observed for all but one distance in the series (−0.03 Å). The reason for this deviation is not fully understood, but as the relative position of the e_g peak was previously shown to depend on the number of states, see Figure 2, it is possible that this is related to a change in the ordering of the final states. The position of the π^* peak does not change, but the relative intensity is affected by 20 %. Compared to the metal-ligand distances, the effects of varying the ligand C−N distances are smaller, see Figure SI 9.

The results show that uncertainties in the geometry can affect the accuracy of the simulations. On the positive side, the observed spectral changes suggest that it can be possible to model the variations in the X-ray spectra during the dynamics of a bond-breaking process using the core RAS method.

Computational Timing

A drawback of the RAS method is the relatively high computational cost, at least compared to the semi-empirical CTM[6,7] or the density-functional theory restricted open-shell configuration interaction with singles (DFT-ROCIS) method.[11] To be able to apply the core RAS method to medium-sized systems, it is important to balance the choice of active space,

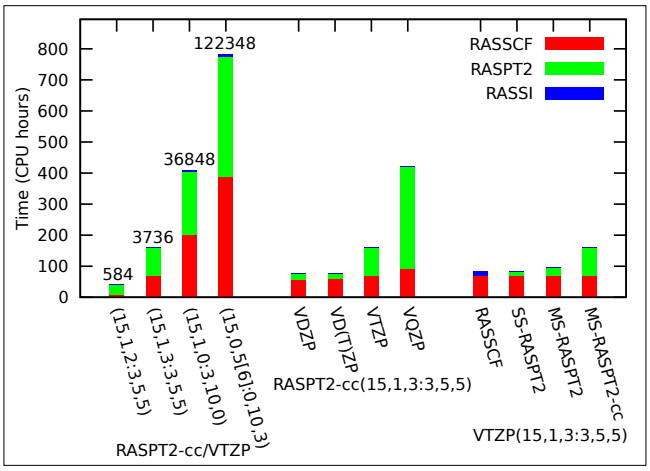


Figure 9: Analysis of the total computational cost for the various steps in the calculation of L-edge XAS spectra of $[Fe(CN)_6]^{3-}$ as function of active space, basis set, and computational procedure. The timings refer to an Intel Xeon E5-2660 "Sandy Bridge" processor at 2.2 GHz using 8 GB RAM. The number of doublet configuration state functions in B_{1g} symmetry is also shown.

number of states, basis set, and computational procedure to get accurate results with limited computational cost.

The reference calculation in the cost comparison is a simulation of the L-edge XAS spectrum of $[Fe(CN)_6]^{3-}$ using RASPT2-cc/ANO-RCC-VTZP (15,1,3:3,5,5) and 80 states. In this calculation, the 2p orbitals are in RAS1, while the e_g and π^* orbitals are in RAS3 allowing a maximum of three electrons. The total computational cost is 161 CPU hours, with comparable timings for the RASSCF and RASPT2 steps, see Figure 9. All calculations are performed in serial mode.

For this system, which has relatively few atoms, the size of the active space is the most important factor. To give an estimate of the size of the problem the number of doublet configurations in B_{1g} symmetry is given, and the numbers are similar for other symmetries. The total computational cost can be decreased to $\sim 25\%$ by limiting the number of RAS3 excitations from three to two, but this again leads to unacceptably large changes in the spectrum.[26] Allowing all possible excitations among the active valence orbitals, RAS(15,1,0:3,10,0), increases the number of configurations significantly and more than doubles the computational

cost for both the RASSCF and the RASPT2 steps.

As outlined in the methodology section, for systems that lack centrosymmetry it might be required to put the 2p orbitals in the RAS3 space and allowing a maximum of five electrons in the excited state, RAS(15,0,5[6]:0,10,3). The number of configuration state functions in the CI expansion increases compared to the RAS1 method, from 36 848 to 122 348, because high-energy configurations with up to six holes in the 2p orbitals are also included. These new configurations have very high energies compared to those with a single core excitation, and do not significantly affect the relative energy of the relevant states, but the computational cost increases by a factor of two.

Increasing the size of the basis set mainly affects the cost of the RASPT2 calculation. Calculations with the smaller VDZP and the mixed VD(T)ZP basis set are dominated by the RASSCF step. The savings in computational time compared to the VTZP calculation are therefore only 50%, even though the cost of the RASPT2 step itself decreases by 80%. The gains of using a smaller basis set will be more significant with a larger number of atoms or smaller active spaces. A recommendation is to start the simulations using a VDZP basis set and then use these orbitals as a starting point when expanding the basis to VTZP at a later stage.

Neglecting core correlation reduces the cost of the RASPT2 step significantly from 90 to 26 CPU hours. As the effects on the spectrum are relatively small, core correlation should not be included by default. The singe-state calculations takes even less time (15 CPU hours), again without any significant loss in accuracy compared to the multi-state approach. Without core correlation, the RASPT2 step is less time-consuming than the RASSCF part, at least for these relatively small systems. Thus, it is not recommended to calculate the spectrum only at the RASSCF level, especially as the effects of dynamical correlation can be large. The time for the RASPT2 calculations could also decrease further with the implementation of parallel RASPT2 in Molcas.[51]

CONCLUSIONS

The RAS approach can be used to simulate metal L-edge XAS spectra with reasonable accuracy without the use of any system-specific fitting parameters. These attributes make RAS a powerful approach to extract information from X-ray spectra, especially in cases where electronic and/or geometric structure is not known.

For ferric chloride and ferricyanide the errors in peak positions are $\leqslant$1 eV and the errors in the relative intensity are $\leqslant$30% for the major peaks and $\sim$50% for some of the minor peaks in ferric chloride. This accuracy is enough to assign individual peaks and give a good fingerprint for the electronic structure. Relative to these errors, the factors that significantly affect the simulated spectra are; the active space, the number of states, inclusion of dynamical correlation and the IPEA shift. They all have effects that are larger than 1 eV in peak positions and/or affects the relative intensities by more than 50%. These observations are obtained for systems with well-defined ground states, and systems with low-lying valence excited states can be more sensitive.

There is no simple rule when it comes to the optimal number of states in the state-average optimization. As the effects are large, it is recommended to systematically increase the number of states, carefully monitoring the shape of the active orbitals and the changes in the spectrum. For ferricyanide up to 80 states per spin multiplicity in four irreducible representations are required to get the π^* orbitals into the active space and the shape of the π^* peak changes up to at least 120 states. The IPEA shift should not be used as an adjustable parameter to improve the agreement with experiment, but a large change in the calculated spectra is an indication that the results must be interpreted with care.

Due to the large effect of dynamical correlation it is recommended to perform the simulations at the RASPT2 level. To save computational time, the calculations can be performed using the state-specific formalism without correlating the core orbitals. This leads to time savings of 80% compared to multi-state calculations involving correlated core orbitals, and leads to small changes in the calculated spectra, less than 0.2 eV in peak position and 10% in intensity. The imaginary shift also has a relatively small effect on the final spectra, which suggests that there are no major problems with intruder states for the current systems.

The spectra simulated with a sufficient number of states are not very sensitive to the quality of the basis set, with peak energies affected by 0.5 eV. In the present systems even a double-ζ basis gives good results. However, with a small basis set the bottleneck can instead be the size of the active space rather than the number of basis functions, and using triple-ζ does not lead to any huge increases in computational cost.

Finally, uncertainties in metal-ligand distances of $\pm$ 0.06 Å lead to changes in peak energies of 0.5 eV. The use of geometries from density-functional theory optimizations should in most cases result in reasonable spectra. At the same time, the sensitivity to ligand distances suggests that it is possible to study changes in the L-edge XAS spectra during a bond-breaking process.

The RAS approach can become a very important tool in the analysis of electronic structures and reaction dynamics probed by different X-ray spectroscopies. The current cost and sensitivity analysis should help users at all levels to design more stable models and to find good compromises between accuracy and timing. With optimized settings, the core RAS approach can be applied to e.g., heme systems, where the iron centers are very challenging to by standard UV/Vis methods due to the strong absorption by the porphyrine ligand. For these systems, the combination of X-ray spectroscopy with a RAS analysis could lead to new insights into the electonic structure of the iron center.

ACKNOWLEDGMENTS

Financial support was provided by the Carl Trygger foundation (RVP and MO), the Marcus and Amalia Wallenberg Foundation (ML) and the Swedish Research Council (MO and ML). The computations were performed on resources provided by SNIC through the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under project s00112-267 and National Supercomputer Centre (NSC) at Linköping University under project liu-2012-00060-46.

References

[1] R. Laskowski, P. Blaha, *Phys. Rev. B* **2010**, *82*, 205104.

[2] J. Vinson, J. Rehr, J. Kas, E. Shirley, *Phys. Rev. B* **2011**, *83*, 115106.

[3] P. Krüger, C. R. Natoli, *Phys. Rev. B* **2004**, *70*, 245120.

[4] P. Krüger, *Phys. Rev. B* **2010**, *81*, 125121.

[5] F. M. F. de Groot, J. C. Fuggle, B. T. Thole, G. A. Sawatzky, *Phys. Rev. B* **1990**, *42*, 5459.

[6] A. Tanaka, T. Jo, *J. Phys. Soc. Jpn.* **1994**, *63*, 2788–2807.

[7] F. M. F. de Groot, *Coord. Chem. Rev.* **2005**, *249*, 31–63.

[8] K. Ogasawara, T. Iwata, Y. Koyama, T. Ishii, I. Tanaka, H. Adachi, *Phys. Rev. B* **2001**, *64*, 115413.

[9] H. Ikeno, T. Mizoguchi, Y. Koyama, Y. Kumagai, I. Tanaka, *Ultramicroscopy* **2006**, *106*, 970–975.

[10] M. W. Haverkort, M. Zwierzycki, O. K. Andersen, *Phys. Rev. B* **2012**, *85*, 165113.

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

[12] D. Maganas, M. Roemelt, T. Weyhermüller, R. Blume, M. Hävecker, A. Knop-Gericke, S. DeBeer, R. Schlögl, F. Neese, *Phys. Chem. Chem. Phys.* **2014**, *16*, 264–276.

[13] E. Otero, N. Kosugi, S. G. Urquhart, *J. Chem. Phys.* **2009**, *131*, 114313.

[14] I. Josefsson, K. Kunnus, S. Schreck, A. Föhlisch, F. M. F. de Groot, Ph. Wernet, M. Odelius, *J. Phys. Chem. Lett.* **2012**, *3*, 3565–3570.

[15] B. O. Roos in *The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations*, John Wiley & Sons, Inc., **2007**, pp. 399–445.

[16] J. Olsen, B. O. Roos, P. Jørgensen, H. J. A. Jensen, *J. Chem. Phys.* **1988**, *89*, 2185–2192.

[17] P.-Å. Malmqvist, A. Rendell, B. O. Roos, *J. Phys. Chem.* **1990**, *94*, 5477–5482.

[18] R. Klooster, R. Broer, M. Filatov, *Chem. Phys.* **2012**, *395*, 122–127.

[19] I. Josefsson, S. K. Eriksson, N. Ottosson, G. Öhrwall, H. Siegbahn, A. Hagfeldt, H. Rensmo, O. Björneholm, M. Odelius, *Phys. Chem. Chem. Phys.* **2013**, *15*, 20189–20196.

[20] Ph. Wernet, K. Kunnus, S. Schreck, W. Quevedo, R. Kurian, S. Techert, F. M. F. de Groot, M. Odelius, A. Föhlisch, *J. Phys. Chem. Lett.* **2012**, *3*, 3448–3453.

[21] S. I. Bokarev, M. Dantz, E. Suljoti, O. Kühn, E. F. Aziz, *Phys. Rev. lett.* **2013**, *111*, 083002.

[22] K. Atak, S. I. Bokarev, M. Gotz, R. Golnak, K. M. Lange, N. Engel, M. Dantz, E. Suljoti, O. Kühn, E. F. Aziz, *J. Phys. Chem. B* **2013**, *117*, 12613–12618.

[23] E. Suljoti, R. Garcia-Diez, S. I. Bokarev, K. M. Lange, R. Schoch, B. Dierker, M. Dantz, K. Yamamoto, N. Engel, K. Atak, O. Kühn, M. Bauer, J.-E. Rubensson, E. F. Aziz, *Angew. Chem. Int. Edit.* **2013**, *52*, 9841–9844.

[24] N. Engel, S. I. Bokarev, E. Suljoti, R. Garcia-Diez, K. M. Lange, K. Atak, R. Golnak, A. Kothe, M. Dantz, O. Kühn, E. F. Aziz, *J. Phys. Chem. B* **2014**, *118*, 1555–1563.

[25] K. Kunnus, I. Josefsson, S. Schreck, W. Quevedo, P. S. Miedema, S. Techert, F. M. F. de Groot, M. Odelius, Ph. Wernet, A. Föhlisch, *J. Phys. Chem. B* **2013**, *117*, 16512–16521.

[26] R. V. Pinjari, M. G. Delcey, M. Guo, M. Odelius, M. Lundberg, *J. Chem. Phys.* **2014**, *141*, 124116.

[27] R. V. Pinjari, M. G. Delcey, M. Guo, M. Odelius, M. Lundberg, *J. Chem. Phys.* **2015**, *142*, 069901.

[28] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, K. Wolinski, *J. Phys. Chem.* **1990**, *94*, 5483–5488.

[29] K. Andersson, P.-Å. Malmqvist, B. O. Roos, *J. Chem. Phys.* **1992**, *96*, 1218–1226.

[30] J. Finley, P.-Å. Malmqvist, B. O. Roos, L. Serrano-Andrs, *Chem. Phys. Lett.* **1998**, *288*, 299 – 306.

[31] P.-Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, L. Gagliardi, *J. Chem. Phys.* **2008**, *128*, 204109.

[32] B. O. Roos, P.-Å. Malmqvist, *Phys. Chem. Chem. Phys.* **2004**, *6*, 2919–2927.

[33] Ph. Wernet, K. Kunnus, I. Josefsson, I. Rajkovic, W. Quevedo, M. Beye, S. Schreck, S. Grubel, M. Scholz, D. Nordlund, W. Zhang, R. W. Hartsock, W. F. Schlotter, J. J. Turner, B. Kennedy, F. Hennies, F. M. F. de Groot, K. J. Gaffney, S. Techert, M. Odelius, A. Föhlisch, *Nature* **2015**, *520*, 78–81.

[34] E. C. Wasinger, F. M. F. de Groot, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **2003**, *125*, 12894–12906.

[35] R. K. Hocking, E. C. Wasinger, F. M. F. de Groot, K. O. Hodgson, B. Hedman, E. I. Solomon, *J. Am. Chem. Soc.* **2006**, *128*, 10442–10451.

[36] F. Aquilante, L. De Vico, N. Ferré, G. Ghigo, P.-Å. Malmqvist, P. Neogrády, T. B. Pedersen, M. Pitoňák, M. Reiher, B. O. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov, R. Lindh, *J. Comput. Chem.* **2010**, *31*, 224–247.

[37] M. Douglas, N. M. Kroll, *Ann. Phys. New York* **1974**, *82*, 89–155.

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

[39] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, P.-O. Widmark, *J. Phys. Chem. A* **2005**, *109*, 6575–6579.

[40] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, P.-O. Widmark, *J. Phys. Chem. A* **2004**, *108*, 2851–2858.

[41] F. Aquilante and T. B. Pedersen and R. Lindh, *J. Chem. Phys.* **2007**, *126*, 194106.

[42] F. Aquilante, P.-Å. Malmqvist, T. B. Pedersen, A. Ghosh, B. O. Roos, *J. Chem. Theory Comput.* **2008**, *4*, 694–702.

[43] J. Boström, M. G. Delcey, F. Aquilante, L. Serrano-Andrés, T. B. Pedersen, R. Lindh, *J. Chem. Theory Comput.* **2010**, *6*, 747–754.

[44] B. A. Hess, C. M. Marian, U. Wahlgren, O. Gropen, *Chem. Phys. Lett.* **1996**, *251*, 365 – 371.

[45] P.-Å. Malmqvist, B. O. Roos, *Chem. Phys. Lett.* **1989**, *155*, 189 – 194.

[46] P.-Å. Malmqvist, B. O. Roos, B. Schimmelpfennig, *Chem. Phys. Lett.* **2002**, *357*, 230 – 240.

[47] N. Forsberg, P.-Å. Malmqvist, *Chem. Phys. Lett.* **1997**, *274*, 196–204.

[48] G. Ghigo, B. O. Roos, P.-Å. Malmqvist, *Chem. Phys. Lett.* **2004**, *396*, 142 – 149.

[49] M. O. Krause, J. H. Oliver, *J. Phys. Chem. Ref. Data* **1979**, *8*, 329–338.

[50] M. Ohno, G. A. van Riessen, *J. Electron Spectrosc.* **2003**, *128*, 1 – 31.

[51] S. Vancoillie, M. G. Delcey, R. Lindh, V. Vysotskiy, P.-Å. Malmqvist, V. Veryazov, *J. Comput. Chem.* **2013**, *34*, 1937–1948.