

Blueprint XAS: a Matlab-based toolbox for the fitting and analysis of XAS spectra

Mario Ulises Delgado-Jaime, Craig Philip Mewis and Pierre Kennepohl*

Received 14 March 2009
Accepted 4 November 2009The University of British Columbia, Department of Chemistry, 2036 Main Mall, Vancouver, Canada BC V6T 1Z1.
E-mail: pierre@chem.ubc.ca

Blueprint XAS is a new *Matlab*-based program developed to fit and analyse X-ray absorption spectroscopy (XAS) data, most specifically in the near-edge region of the spectrum. The program is based on a methodology that introduces a novel background model into the complete fit model and that is capable of generating any number of independent fits with minimal introduction of user bias [Delgado-Jaime & Kennepohl (2010), *J. Synchrotron Rad.* **17**, 119–128]. The functions and settings on the five panels of its graphical user interface are designed to suit the needs of near-edge XAS data analyzers. A batch function allows for the setting of multiple jobs to be run with *Matlab* in the background. A unique statistics panel allows the user to analyse a family of independent fits, to evaluate fit models and to draw statistically supported conclusions. The version introduced here (v0.2) is currently a toolbox for *Matlab*. Future stand-alone versions of the program will also incorporate several other new features to create a full package of tools for XAS data processing.

© 2010 International Union of Crystallography
Printed in Singapore – all rights reserved**Keywords:** X-ray absorption spectroscopy data analysis; optimization tools; *Matlab*.

1. Introduction

Analysis of near-edge XAS data typically involves the following steps applied sequentially: (i) scan averaging, (ii) energy calibration, (iii) background subtraction, (iv) normalization and (v) peak fitting. This general approach is built into a number of data analysis packages, dealing with each step independently (George & Pickering, 1993; Ressler, 1998; Webb, 2005). A particular concern with this approach resides in the fact that background subtraction may, in principle, have a significant impact on peak fitting and analysis. Practitioners therefore tend to perform several different fits starting from different background subtractions to determine whether a particular fit is robust. However, this approach still assumes that performing these steps independently does not introduce additional uncertainty, nor does it address the potential influence of user bias in each step. We have recently investigated (Delgado-Jaime *et al.*, 2006) and developed (Delgado-Jaime & Kennepohl, 2010) a new methodology that incorporates background subtraction into the fitting procedure and that involves the simultaneous merging of all of the relevant parent functions (*i.e.* functions governing the background of different regions of the spectrum) within a single complete fit model. Simultaneous fitting of the background also allows for judicious use of parameters to ultimately decrease the total number of fitting parameters (Delgado-Jaime & Kennepohl, 2010).

In addition to the above, optimization of multiple-variable problems presents an additional challenge: the very real possibility that a unique best fit solution is unattainable. To solve this issue, current approaches generally rely on the user generating a number of independent fits. Unfortunately, user bias in the selection of parameter starting values tends to limit the effectiveness of a manual

approach and makes it difficult to determine whether a particular solution is reliable and/or unique. To overcome this problem, we have proposed in a companion manuscript (Delgado-Jaime & Kennepohl, 2010) a methodology that uses a Monte Carlo approach to the simultaneous fitting of background and edge features in XAS spectra.

Herein, we introduce *Blueprint XAS*, a *Matlab*-based program that incorporates this methodology through a graphical user interface (GUI). Within this program, the user can interactively set up an overall fit model that incorporates functions and parameters related to the edges, background and/or peaks. Furthermore, a statistical module focuses on data analysis tools that allow users to evaluate the validity of their fit model and assist in determining the reliability of obtained solutions. This manuscript introduces the GUI generally and then focuses on specific features of the interface, as well as the system requirements and general information regarding the technical aspects of the program.

2. System requirements

Blueprint XAS is coded in *Matlab* R2007a and makes use of several ‘add-on’ optimization and statistics functions belonging to the curve-fitting and statistics toolboxes, respectively. A brief description of some of the specific curve-fitting functions is given in §S1 of the supplementary information¹. Additional details can be consulted at the Mathworks website (Mathworks, 2009; Moler, 1999). Therefore, to run *Blueprint XAS*, *Matlab*, version 7.4.0 R2007a (or a more recent

¹ Supplementary data for this paper are available from the IUCr electronic archives (Reference: OT5603). Services for accessing these data are described at the back of the journal.

version), equipped with the respective curve-fitting and statistics toolboxes, is required. In terms of visualization of the features in the GUI, *Blueprint XAS* is best viewed using wide-screen monitors with WSXGA resolution (1440 × 900 pixels) or better; the minimum recommended resolution is XGA (1024 × 768 pixels). Future versions, which will include a full toolbox for XAS data processing, will be introduced as stand-alone executables, but, given that the current suite focuses exclusively on the fitting and analysis modules, it has been left as a *Matlab*-based toolbox.

The major components of the fitting toolbox of *Blueprint XAS* are encoded in a single .m file (*BlueprintXAS.m*); a supporting .mat file (*BlueprintXAS.mat*) is also required. This latter file contains the current values for all of the objects in the GUI, as well as the data and parameters introduced by the user during the last time the toolbox was in use. Other secondary components are optional although necessary to accomplish certain tasks (see §5 and §6). The basic software and supporting documentation is currently available through SourceForge.net via an open source licensing agreement (see <http://www.sourceforge.net/projects/blueprintxas> for details). Further developments will be distributed through SourceForge and additional contributions to improve and expand features are welcome.

3. Input/output and other basic features

Fig. 1 shows a snapshot of the fitting toolbox of *Blueprint XAS* as it is displayed after typing the command 'BlueprintXAS' into the command window of *Matlab*. To administer the space in the GUI without sacrificing quick access to the different tools and objects that are available in the program, a series of tab panels are available, each focusing on a different aspect of the data analysis process. The graphics panel is visible at all times, although its scaling controls are located in the 'General' panel.

Data can be loaded into the program in several ways from the 'General' tab: (i) 'Calibrated' is an internal data format that will be used to load energy-calibrated and averaged spectra from another pre-fitting module of *Blueprint XAS* (under development); (ii) 'Fit' allows for retrieval of previously fitted parameters and data as well as the GUI parameters that lead to the generation of these fits; (iii) 'Excel' allows users to automatically load the first two columns of a Microsoft Excel file (.xls); and (iv) 'Parameters' retrieves just the parameters and functions used in other sessions. This function is particularly useful if similar settings from a previous session need to

Table 1

Allowed identifiers for the setting up of functions and parameters in a fit model using *Blueprint XAS*.

GUI panel	Allowed identifiers	Definition (according to default function parameters)
Edges, peaks	<i>I</i>	Edge intensity or peak area
	<i>O</i>	Energy position at edge's inflection point or at maximum height of a peak
	<i>W</i>	HWHM of edge or peak
	<i>G</i>	Shape of edge or peak (in Gaussian fraction)
	<i>B</i>	Branching ratio (in multiple-edged problems) or any other ratio
Background	<i>m</i>	Slope in a parent linear polynomial function
	<i>YI</i>	<i>Y</i> intercept in the lowest-energy parent polynomial function
	<i>p</i>	Quadratic coefficient in a parent quadratic polynomial function
	<i>q</i>	Linear coefficient in a parent quadratic polynomial function
	<i>O</i>	Position at the inflection point of a 'switch-like' background (Delgado-Jaime & Kennepohl, 2010)
	<i>W</i>	HWHM of a 'switch-like' background (Delgado-Jaime & Kennepohl, 2010)
	<i>V</i>	Pivot energy of a 'handle-like' background (Delgado-Jaime <i>et al.</i> , 2006)

be applied to the current data loaded into the program. The 'General' tab also includes the tools to run and to create batches of queued jobs (see below).

The other tab panels ('Edges', 'Background' and 'Peaks') focus on the different aspects of the overall evaluation function and the tools that help define the parameters therein. Although the default functions used in each of these tab panels are very specific to near-edge XAS data analysis, the fitting methodology is not. For example, adding an edge feature using the 'Add Edge' function (see Fig. 2) creates a new component to the fit model that corresponds to a suggested edge function. However, the fit component is completely user-editable, so that the parameters and even the functional form itself can be modified with simplicity.

For consistency, all variables within the program are labelled using mostly five identifiers: peak areas or edge intensities are labelled *I* (intensity); peak or edge energies are labelled *O* (position); half width at half-maximum (HWHM) parameters are labelled *W* (width); the shape of a Voigt peak or edge is labelled *G* (Gaussian fraction); and branching ratios (*i.e.* for multi-edge spectra) are given the label *B*. Table 1 shows the complete list of identifiers supported by *Blueprint XAS*. When developing the fit model for a data set, the program

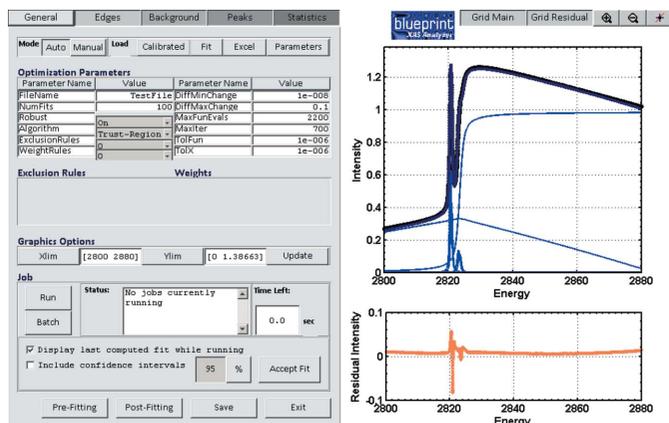


Figure 1
Snapshot of the fitting toolbox of *Blueprint XAS* showing the tools and options in the 'General' panel.

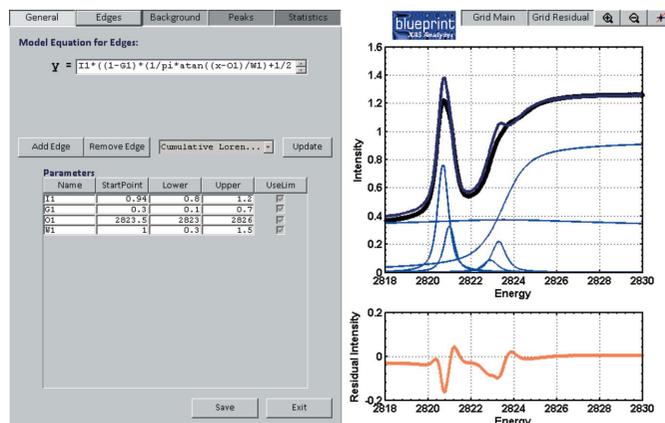


Figure 2
Snapshot of the fitting toolbox of *Blueprint XAS* showing the tools and options in the 'Edges' panel.

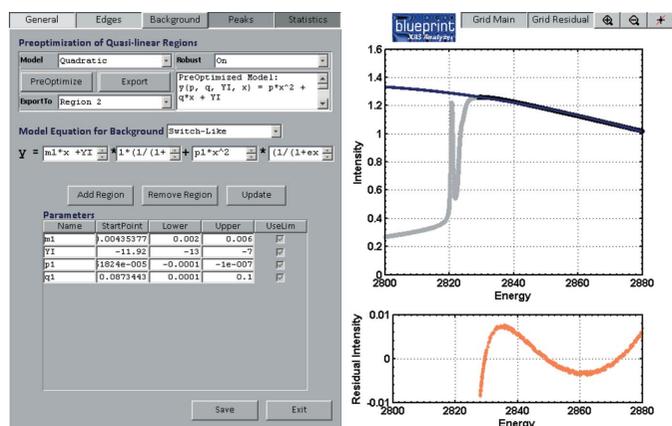


Figure 3 Snapshot of the fitting toolbox of *Blueprint XAS* showing the tools and options in the ‘Background’ panel. The graphics display the result of the optimization of the fit and residual for a quasi-linear region of the spectrum using a quadratic polynomial and a robust optimization method, which outweigh possible outlying points.

provides users with the ability to automatically add as many as four edges and as many as six peaks. This current limitation results from the GUI and can easily be circumvented by adding additional components directly into the evaluation function. In any case, it should be noted that data that require more than the above number of edges and peaks would be rather unusual.

The ‘Background’ panel offers specific functionality to assist in the building of a reasonable background. Initial guesses for polynomial functions fitting a particular background region are not as intuitive for the user as are the parameters in functions fitting an edge or a peak. Therefore, a pre-optimization tool, which uses linear least squares, is included to allow the user to rapidly define reasonable starting points for the parameters in these functions. Results from this pre-optimization can be easily exported as initial guesses to the parameters panel by using the ‘Export’ key (Fig. 3).

To set up the background, in addition to the models that can be created by the user, two built-in weighting models are currently supported in *Blueprint XAS*: the ‘handle-like’ (Delgado-Jaime *et al.*, 2006) and the ‘switch-like’ (Delgado-Jaime & Kennepohl, 2010) model functions. Unlike in the cases of ‘Edges’ and ‘Peaks’, a single profile for the whole background is displayed, regardless of the number of parent functions used (see Fig. 3).

An important feature is that the coefficient parameters can be easily shared among the different components of the evaluation function, thus allowing the user to minimize the number of fit parameters. A particularly useful example is the sharing of the energy position and width between an edge and the switch-like background model (Delgado-Jaime *et al.*, 2006). In situations where the background differs in the pre- and post-edge regions, this serves the dual purpose of minimizing parameters as well as providing a reasonable switch from the pre- to post-edge background functions. This feature helps the user to obtain a very general perspective of the problem, allowing for an insightful linking of the different parameters with relative simplicity. This is especially useful in situations where the fit model is rather complex.

All of the settings and completed fits can be saved to disk at any time using the ‘Save’ function, which is available in each of the five panels of the GUI. The ‘Exit’ function (also visible in all panels) not only closes the GUI but also saves the current settings into the ‘BlueprintXAS.mat’ file (*i.e.* the local workspace file).

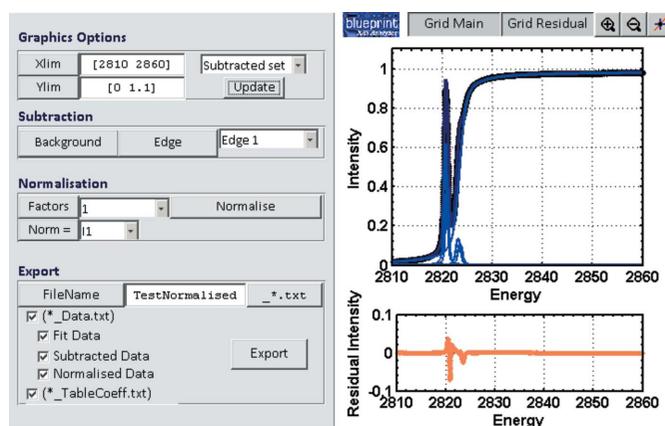


Figure 4 Post-fitting toolbox for *Blueprint XAS*.

4. Background subtraction and normalization

4.1. Post-fitting toolbox

As of the current version (v0.2), *Blueprint XAS* has a post-fitting toolbox (Fig. 4) in which background subtraction and normalization to the data is possible using the results of a fit job. By default, the average values of the coefficient parameters of the background from the independent fits computed within a fit job are used to retrieve the background. The ‘subtract background’ function on the post-fitting GUI uses the retrieved background and subtracts it from the original data set. In a similar way, the subtraction of a fitted edge jump is also possible. The normalization is accomplished by the use of a product that may include up to three factors in the norm. Each of these factors is set up by the user to be the average of any of the coefficient parameters obtained from the independent fits of a fit job. The normalized and subtracted data obtained this way, along with the fit, the original data set, the average and standard deviation on the coefficient parameters in the fits, can all be exported as formatted .txt files. These files can be easily read in Microsoft Excel or other spreadsheet or graphical software packages for further exploration, distribution and publication.

4.2. Internal normalization

The normalization of the data, based on the average of fits computed and performed throughout the use of the post-fitting toolbox of *Blueprint XAS*, is mostly useful to graphically compare the data obtained for several compounds. However, taking advantage of the highly interactive environment of *Blueprint XAS*, an ‘internal normalization’ of the peak features is possible. To accomplish this task, it is necessary to edit the appropriate fit-component field, within the ‘Peaks’ panel, so that the corresponding function on it is multiplied by the parameter chosen to be the intensity of the normalizing edge jump (I_1 in the example given in Fig. 5).

5. Job options

There are several options to run jobs in *Blueprint XAS*. If the operational mode is set to ‘manual’ (see Fig. 1), a single option is available. Under these circumstances the ‘Run’ function (in the General tab) will generate a single fit using the initial guess (‘Start-Point’) provided by the user. *Blueprint XAS* updates the graphics only after the full convergence is achieved. Hence, this operational

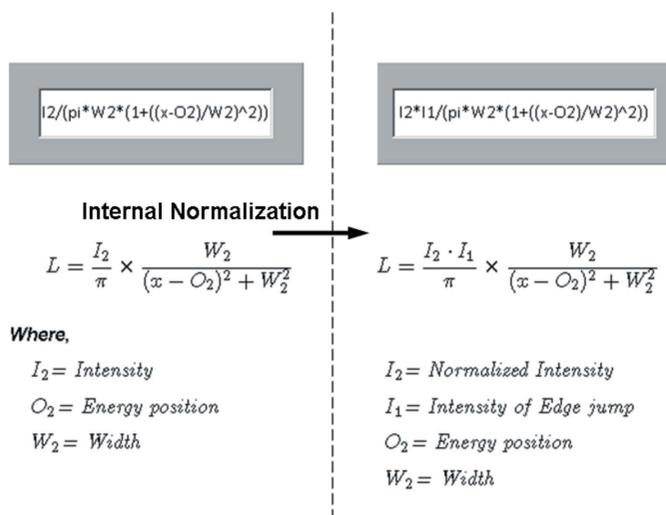


Figure 5 Internal normalization of a Lorentzian peak, using the intensity parameter I_1 of an edge jump.

mode is not significantly different from other conventional methods implemented in other existing software.

However, in ‘Auto’ mode the full methodological approach as described in the companion manuscript is applied (Delgado-Jaime & Kennepohl, 2010). With additional options available (see checkboxes in the Job panel, Fig. 1), the user has the choice of updating the graphics in the GUI every time a new fit is successfully generated or wait until completion of the fits before updating the graphics. The former option may be of interest to novice users or those seeking a visual cue as to the progress of the fits, but it can dramatically increase total calculation times and it is not recommended during typical data analysis.

Given that a typical fit cycle will include a large number of independent fits (we recommend 100 fits as a reasonable default setting) (Delgado-Jaime & Kennepohl, 2010), the overall fit time is longer than one would generally observe using traditional fitting software. For single edges with two to three peaks and less than 1000 data points we have observed average fit times of 1–2 h on a typical personal computer. For more complex fit models (multiple edges plus five to six peaks), fit times can increase to 5–8 h. Such fitting times make it useful to use off-line and/or batch job type operation. The ‘batch’ option opens an additional window (see Fig. 6) specifically designed to control a sequence of fitting jobs submitted through *Blueprint XAS*. Such fitting jobs run in the background with no interactive component. This approach is well suited for downloading complex fitting jobs to a high-throughput computing cluster if available.

6. Statistical features

The ‘Statistics’ panel of *Blueprint XAS* is a crucial component of the overall methodology and represents a unique aspect of this data analysis toolbox. This panel allows the user to visualize relevant statistical information and explore the resultant fits. After a job has been completed, each fit, along with its individual components and its corresponding residuals, can be easily visualized using the selector in the top corner of the panel (Fig. 7). Additionally, by selecting a fit, the summary for each fit, which includes the goodness-of-fit parameters and the value of every coefficient in the evaluation function, along

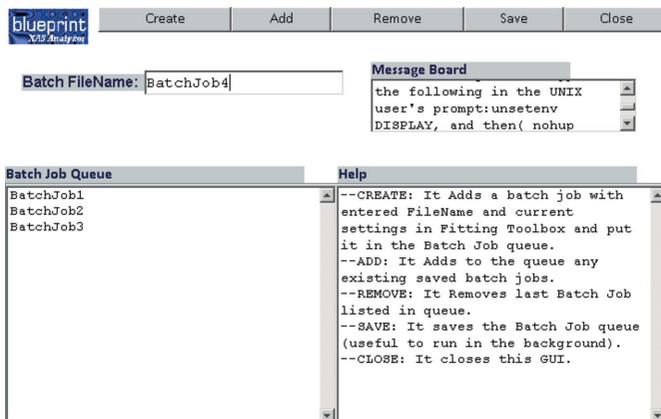


Figure 6 Batch job creator, accessed after clicking on ‘Batch’ in the ‘General’ panel (Fig. 1) of the fitting toolbox of *Blueprint XAS*.

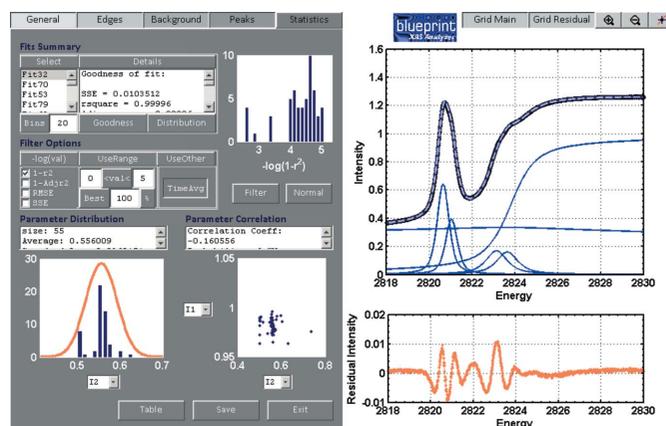


Figure 7 Snapshot of the fitting toolbox of *Blueprint XAS* showing the diverse features in the ‘Statistics’ panel.

with their confidence intervals, is displayed in the ‘Details’ display field.

Moreover, the user can visualize the frequency distribution of fits generated in a given job according to the value of the four goodness-of-fit indicators on a logarithmic scale. Undesired fits can be filtered out by selecting a range of acceptable fit indicator values. Additionally, the user can filter a percentage of the fits to visualize only the best-ranked fits (according to their goodness of fit). Filtering by time is another useful tool for whenever fits that did not converge need to be removed. In the lower left corner of the panel, the user can explore the frequency distribution (and the deviation from an ‘ideal’ Gaussian distribution) for all parameters in the fits (filtered or non-filtered). In the lower right of Fig. 7, correlation plots can be used to determine possible relationships between different parameters in the fit model.

The ‘Table’ button in the ‘Statistics’ panel (see Fig. 8) generates a table with the numerical summary of the fit parameters (in filtered or non-filtered sets). A plot displaying the time to generate each fit in the set is also displayed with the entire non-filtered population set (in thin blue lines) in contrast to the filtered (or non-filtered) set (in blue dots). We have observed that fits that require significantly more time to complete than the average of fits generally correlate with poor fits owing to non-convergence. A small number of such fits can easily be filtered out, but a fit set that includes a very large number of slowly

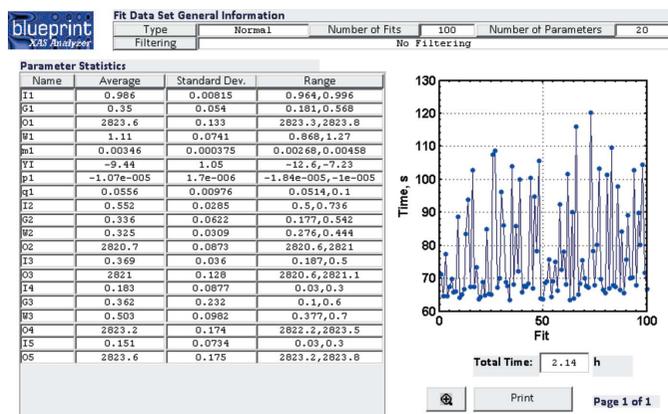


Figure 8 Display of the numerical summary generated from a non-filtered set of fits using the ‘Table’ button in the ‘Statistics’ panel (Fig. 7) of *Blueprint XAS*.

converging fits (as compared with the median fit time) should be considered as a warning flag of a poor fit set for which either the fit model is causing difficulties or some of the fitting parameters (e.g. MaxIter and MaxFunEvals) should be adjusted.

7. Simulation of calculated spectra using *Blueprint XAS*

Simulated spectroscopic data obtained from (for example) time-dependent density functional theory (TD-DFT) calculations or atomic multiplet simulations can also be ‘fit’ to a particular experimental data set using *Blueprint XAS* although the graphical interface for this feature is currently lacking

By using an auxiliary code, the user can directly read computed transition energies and oscillator strengths from an output file of the *Amsterdam Density Functional (ADF)* computational package, the ‘xy’ file from a *TD-multiplier* calculation or else the first two columns of a worksheet in Microsoft Excel. Each transition line is then modelled as a Gaussian, Lorentzian or pseudo-Voigt, but contrary to the regular functions used to fit peaks to the data the same single parameter is imposed to control the shift in the energy position of all the simulated lines. Under these settings the effect of changing the value of this parameter has the effect of shifting the entire set of lines as needed. In a similar way, a scaling factor parameter is used to proportionally amplify the intensity given by the computed oscillator strengths. To account for the shape and width of each line, either a single parameter or a modulating function is used.

The use of a modulating function (f_M ; Fig. 9) allows the user to impose a gradual increase/decrease on the shape and width of the lines, depending on the proximity of the function to the inflection point of an edge jump. Four parameters are used to define the widths and/or shapes of all lines according to this modulating function: W_{Min}/G_{Min} (the minimum width or shape); W_{Max}/G_{Max} (the maximum width or shape); LLW/LLG [the energy, relative to the inflection point of an edge jump (O), at which the width/shape starts to increase/decrease from its minimum/maximum value]; and ULW/ULG [the energy, relative to the inflection point of the edge jump (O), at which the maximum/minimum width/shaper should be reached].

In theory, a maximum of six peaks are allowed per evaluation function in *Blueprint XAS*. In practice, a very large number of them is permitted, provided the number of parameters uniquely defined in the ‘peaks’ panel does not exceed 36. Therefore, this auxiliary code distributes the transition-line peaks into six (or less) peak functions

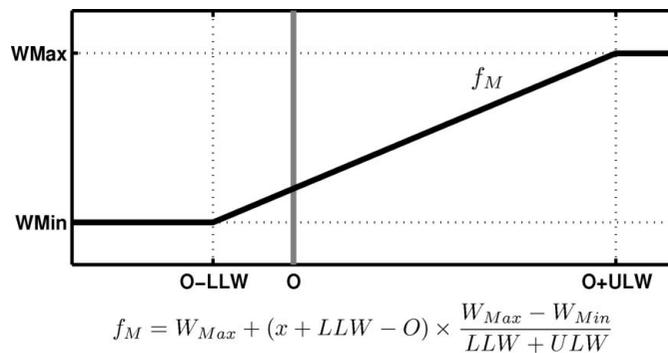


Figure 9 Modulation function (f_M) as used by the auxiliary code to define the width of each transition line.

according to the guidelines provided by the user. Each of the peak functions generated this way can be impressively long. In contrast, the number of total parameters, uniquely defined in the ‘peaks’ panel, could be quite small.

The results from four different TD-DFT calculations on tetragonal CuCl_4^{2-} were used to fit one experimental data set on the Cl *K*-edge XAS of $(\text{NEt}_4)_2\text{CuCl}_4$. These calculations were performed from the crystal structure of the CuCl_4^{2-} moiety of $(\text{NEt}_4)_2\text{CuCl}_4$ (Mahoui *et al.*, 1994) (see Table S2 in the supplementary information), using the *Amsterdam Density Functional 2007.01* program (ADF, 2007; te Velde *et al.*, 2001; Baerends *et al.*, 1973). An unrestricted triple- ζ basis set with double polarization and no core level was used, in conjunction with the Vosko–Wilk–Nusair local density approximation (Vosko *et al.*, 1980) and the Becke–Perdew gradient corrections (Becke, 1988; Perdew, 1986a,b). A total of 20 transition lines were computed in each of the four calculations (one for each chlorine atom). Tables S2 and S3 of the supplementary information list the energy and oscillator strengths for the 80 transition lines computed.

The obtained transition lines were modelled using pseudo-Voigt functions. A single parameter, O_2 , was used as an energy shift for these lines. Additionally, I_2 was set as a scaling factor and G_2 as a constant shape for all of the lines. A function with parameters W_2 (W_{Min}), W_3 (W_{Max}), B_1 (LLW) and B_2 (ULW) was used to model the width of each of the lines, about the inflection point of the edge jump, O_1 . The transition lines were nicely distributed in four peak functions (one for each of the transition-lines sets originated from each of the chlorine atoms).

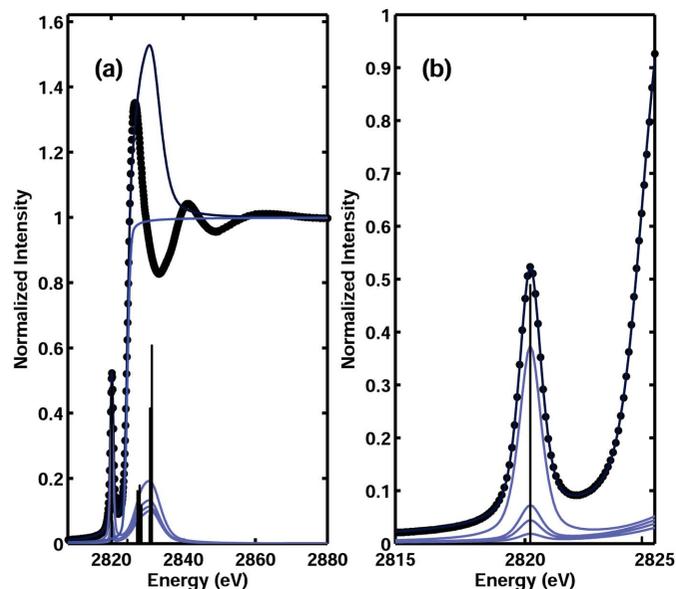
The evaluation function used in the fitting of the resulting model to the data included not only these four peak functions but also an edge jump, modelled as a cumulative pseudo-Voigt and a switch-like background (Delgado-Jaime & Kennepohl, 2010). Fig. S6 in the supplementary information shows the functional form of the evaluation function used in the optimization of this simulation. Since this particular problem is highly constrained by the computed transition lines, a single fit is obtained. Table S5 (in the supplementary information) lists the ‘fitoptions’ variables passed on to *Matlab* to undergo this fit. Moreover, Fig. 10 shows the obtained fit and Table 2 lists the results for the obtained optimized parameters.

From Fig. 10, it is evident that the simulated TD-DFT spectrum does not fit the experimental data particularly well. The relative energy position of the transition lines for the near-edge features is overestimated by the DFT calculation, when comparing these with the energy position of the pre-edge transition lines. Furthermore, assuming that self-absorption effects are not significant in the experimental Cl *K*-edge XAS data set used, either the intensity of the lines comprising the near-edge feature is overestimated or the

Table 2

Results for the fitting of a Cl *K*-edge XAS data set of (NEt₄)₂CuCl₄, based on 80 TD-DFT transition lines computed as detailed in the text.

Coefficient parameter	Bounds		Results	
	Lower	Upper	Average	95% confidence interval
Edge				
Intensity, I_1	1	3	1.541	0.029
Inflection point, O_1	2824	2830	2824.5	0.1
Shape, G_1	0	1	0.733	0.163
HWHM width, W_1	0.1	3	0.782	0.049
Background				
m_1	5.00×10^{-5}	0.002	1.338×10^{-4}	1.095×10^{-4}
YI	-5	0	-0.043	0.306
p_1	0	1000	2.155×10^{-5}	1.238×10^{-5}
q_1	-1000	0	-1.242×10^{-1}	7.12×10^{-2}
Transition lines				
Scaling factor, I_2	100	1000	252.2	9.3
Energy shift, O_2	84	88	86.5	<0.1
Shape, G_2	0	1	0.364	0.064
Wmin, W_2	0.2	1	0.502	0.013
Wmax, W_3	2	3	3.000	0.914
LLW, B_1	1	4	4	-
ULW, B_2	8	15	8	-

**Figure 10**

Optimized simulation fit for the Cl *K*-edge XAS data set 7 of (NEt₄)₂CuCl₄ from the output of TD-DFT calculations displayed in (a) the near-edge region and (b) the pre-edge region.

intensity of the ones accounting for the pre-edge are underestimated by the DFT calculation under the selected settings.

Nevertheless, the applicability of this tool goes beyond the fitting of computed transition lines for one compound. In this case, each peak function was built using the transition lines of one of the chlorine atoms in the compound. This implies that for cases in which the experimental data involve several species, each peak function can be built using a complete set of lines corresponding to each of these species in order to find, for instance, relative concentrations.

8. Summary

Blueprint XAS is a *Matlab*-based toolbox for *Matlab* still under development, but with a fully operational basic functionality. It

incorporates a powerful methodology that deals with the fitting of XAS data using holistic models that include the background, edges and peaks altogether and that is capable of generating any number of independent fits by means of a Monte Carlo approach that seeks for good start points. Since *Blueprint XAS* make use of the curve fitting and statistics toolboxes developed by The Mathworks, at least the 7.4.0 version of *Matlab* is required.

The versatile graphical user interface of *Blueprint XAS* introduced here allows the XAS user to easily set functions, parameters and lower and upper limits of the evaluation function, while at the same time offering an accessible way for other non-XAS users to customize these settings to their own needs. Parameters can easily be shared and combined amongst the different components in the evaluation function, allowing for a more general perspective on the visualization of complex problems. Moreover, its unique statistics panel allows for the fast analysis of any family of independent fits generated, the evaluation of fit models and the exploration of the behaviour of parameters and their possible correlation with others. An auxiliary code is useful to build models based on the calculation of transition lines by means of DFT, *TT-multiplet* or any other computational package whose output is a set of line strengths.

Overall, we believe that the use of *Blueprint XAS* will speed the processing of raw XAS data and that will increase the accuracy and robustness of conclusions and results obtained. We have made the code available through SourceForge to encourage others to participate on the development of *Blueprint XAS* to expand and improve the implementation described herein.

This research is funded by NSERC (the Natural Science and Engineering Research Council of Canada); infrastructure support provided by UBC. Special thanks from one of the authors (MUDJ) whose graduate fellowship is supported with funds from CONACYT (Consejo Nacional de Ciencia y Tecnología, México). *Matlab* was operated on infrastructure funded by CFI and BCKDF through the Centre for Higher Order Structure Elucidation (CHORSE).

References

- ADF (2007). *The Amsterdam Density Functional program system*. Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands.
- Baerdens, E. J., Ellis, D. E. & Ros, P. (1973). *Chem. Phys.* **2**, 41–51.
- Becke, A. D. (1988). *Phys. Rev. A*, **38**, 3098–3100.
- Delgado-Jaime, M. U., Conrad, J. C., Fogg, D. E. & Kennepohl, P. (2006). *Inorg. Chim. Acta*, **359**, 3042–3047.
- Delgado-Jaime, M. U. & Kennepohl, P. (2010). *J. Synchrotron Rad.* **17**, 119–128.
- George, G. N. & Pickering, I. J. (1993). *EXAFSPAK: a suite of computer programs for analysis of X-ray absorption spectra*, <http://ssrl.slac.stanford.edu/exafspak.html>.
- Mahoui, A., Lapasset, J. & Moret, J. (1994). *Acta Cryst.* **C50**, 358–362.
- Mathworks (2009). *fitoptions (curve-fitting toolbox)*, <http://www.mathworks.com/access/helpdesk/help/toolbox/curvefit/index.html?access/helpdesk/help/toolbox/curvefit/f2-17602.html>.
- Moler, C. (1999). *Optimally Speaking: Optimization Toolbox features new methods for large-scale problems*, http://www.mathworks.com/company/newsletters/news_notes/clevescorner/sm99cleve.html.
- Perdew, J. P. (1986a). *Phys. Rev. B*, **34**, 7406.
- Perdew, J. P. (1986b). *Phys. Rev. B*, **33**, 8822–8824.
- Ressler, T. (1998). *J. Synchrotron Rad.* **5**, 118–122.
- Velde, G. te, Bickelhaupt, F. M., Baerends, E. J., Fonseca Guerra, C., van Gisbergen, S. J. A., Snijders, J. G. & Ziegler, T. (2001). *J. Comput. Chem.* **22**, 931–967.
- Vosko, S. H., Wilk, L. & Nusair, M. (1980). *Can. J. Phys.* **58**, 1200–1211.
- Webb, S. M. (2005). *Phys. Scr.* **T115**, 1011–1014.