Multiplatform Applications for XAFS (MAX)

A. Michalowicz and K. Provost*

ICMPE, UMR7182, UPEC–CNRS, 2–8 Rue Henri Dunant, 94320 Thiais, France. *Correspondence e-mail: provost@u-pec.fr

Multiplatform Applications for XAFS (MAX) is a multiplatform (Windows, MacOS and Linux) set of standalone applications dedicated to most aspects of XAFS experiment preparation and spectral data analysis. ABSORBIX deals with absorbance calculations at any X-ray energy and fluorescence self-absorption correction. CHEROKEE allows EXAFS extraction and XANES normalization from raw data. CRYSTALFFREV links any crystallographic structure to FEFF input files. EXAFS nonlinear fitting is performed by ROUNDMIDNIGHT and chemometric aspects are treated by STRAIGHTNOCHASER. Normalization of X-ray magnetic circular-dichroism spectra is performed with SOS_XMCD.

1. General features

Multiplatform Applications for XAFS (MAX) is a set of six standalone applications covering almost all aspects of XAFS experiment preparation and data analysis (Michalowicz et al., 2009, 2013). ABSORBIX deals with X-ray absorbance and fluorescence tables, absorbance calculation at any X-ray energy and fluorescence self-absorption correction. CHEROKEE allows EXAFS extraction, Fourier analysis and XANES normalization from raw data. CRYSTALFFREV links crystallographic structures to FEFF input files. EXAFS nonlinear fits are performed by ROUNDMIDNIGHT and linear chemometric aspects are treated by STRAIGHTNOCHASER. Normalization of X-ray magnetic circular-dichroism (XMCD) spectra is performed with SOS_XMCD.

1.1. Language

Most of the codes, including numerical calculations and the user interface (commands, graphical and file facilities), are written in the multiplatform script language Livecode (https://livecode.com/), which allows the simultaneous creation of standalone applications for Windows, MacOS and Linux from a unique code on a unique platform. The user interfaces are automatically adapted to each operating system, while the functions are strictly identical. One exception is ROUNDMIDNIGHT, which is composed of two applications, a user interface written in Livecode and a fitting procedure based on the CERN minimization code MINUIT (James & Roos, 1975), which is kept in its original Fortran language and compiled separately on each platform.

1.2. Availability

MAX is freely downloadable from the Institute of Chemistry and Materials Paris East (ICMPE) website (Michalowicz, 2017) and is easy to install: just download, unzip and run. Its main advantages are (i) it is totally standalone and (ii) the user’s file management and the graphical and printing facilities

© 2021 International Union of Crystallography

https://doi.org/10.1107/S1574870720003419 1 of 5
follow the standards of each computer operating system. CHEROKEE is provided with a series of example files. Printable Help windows, including XAFS data-analysis principles, mathematical explanations and user manuals, are included in the code.

2. ABSORBIX

X-ray absorption edges, fluorescence lines and mass attenuation coefficients are based on McMaster’s tables (McMaster et al., 1970) as compiled in a Livecode version of Mucal (P. Bandypadhyay & C. U. Segre; http://csrri.iit.edu/mucal.html). ABSORBIX provides the following.
   (i) A search of X-ray absorption edges and fluorescence lines for any energy range.
   (ii) X-ray absorbance values and plots, sample concentrations and thickness optimization for X-ray absorption experiments. Sample chemical composition is entered via an interactive X-ray periodic table. Total absorbance is evaluated versus the absorption jump, the sample mass/surface, the thickness or the sample concentration for solid and liquid dilute samples, taking the solvent characteristics into account.
   (iii) Normalized fluorescence XANES or EXAFS spectra self-absorption correction using the nonlinear equations of Haskel (1999) and Booth & Bridges (2005), respectively, solved using Brent’s numerical equation root solver algorithm (Brent, 1973), without any absorption-length linear approximation.

3. CHEROKEE

This application contains all of the procedures for a complete treatment of X-ray absorption spectra to obtain EXAFS and XANES spectra for further analysis.
   (i) Energy recalibration based on monochromator characteristics.
   (ii) Edge-jump determination by extrapolation of the pre-edge and atomic backgrounds.
   (iii) Atomic background determination by three different algorithms: polynomial least-squares fitting, cubic spline least-squares fitting and spline smoothing (de Boor, 1978). We underline the advantages of manual access to different background-removal options, associated with a quantitative estimation of the systematic amplitude errors by Fourier analysis, in order to allow an easy switch to the most efficient algorithm in difficult cases where an unique automatic algorithm may be insufficient and lead to side effects.
   (iv) Energy-dependent EXAFS normalization using either Lengeler–Eisenberger expansion (Lengeler & Eisenberger, 1980) or McMaster’s tables (McMaster et al., 1970).
   (v) Fourier transform and filtering, Fourier smoothing, signal-to-noise evaluation and residual background subtraction. Scattering phase and amplitude extraction by Fourier filtering from any EXAFS spectrum (experimental or theoretical).
   (vi) Total absorbance or XANES normalization and calculation of derivatives.
CHEROKEE also provides many other functions, such as monochromator calibration, energy and k-space deglitching, and switchable automatic or manual step-by-step data treatments.

CHEROKEE is organized into two sections for the treatment of individual spectra and of series of spectra simultaneously.

4. CRYSTALFFREV

CRYSTALFFREV is a crystallographic tool for the conversion of crystal structures into FEFF (Rehr & Albers, 2000; Kas et al., 2020) or FDMNES (Joly, 2001; Bunău et al., 2021) input files.

The structure can be created either by manually introducing the cell parameters, space group and asymmetric unit content, or input from a CIF or PDB file, and saved in CRYSTALFFREV, CIF or PDF format. All International Tables for Crystallography space groups (Hahn, 2002) are available in a database and can be selected by name or number, imported, edited and checked. All International Tables for Crystallography origin choices and trigonal space-group descriptions (rhombohedral or hexagonal axes) can be used. It is also possible to enter nonconventional crystallographic symmetry operations.

A cluster is created by selecting the central absorbing atom and defining any cluster radius. It is possible to rotate the local cluster in order to choose the z axis (which is useful for axial ligand analysis of macrocycle metalloenzyme active sites). Disorder tools are available to deal with partial substitutions or vacancies through random generators, or to introduce random radial distribution functions for selected atoms (Gaussian distribution).

Finally, FEFF7, FEFF8/9 and FDMNES input files, including most control cards, can be directly edited in CRYSTALFFREV.

5. ROUNDMIDNIGHT

ROUNDMIDNIGHT is our EXAFS fitting program. The fit of theoretical versus experimental EXAFS is based on the standard multiple-scattering and spherical waves EXAFS formula as documented in the FEFF ab initio code (Rehr & Albers, 2000; Kas et al., 2020) and the general function minimization software MINUIT from the CERN computer library (James & Roos, 1975).

ROUNDMIDNIGHT is totally compatible with FEFF, both in the fitting module and in the multiple-scattering paths analysis module. However, any other sources of scattering phase and amplitude functions can also be admitted. The empirical photoelectron mean free path curve can be fitted on model standards and used in place of, or in combination with, the FEFF reduction factor and calculated photoelectron mean free path in order to improve the EXAFS fit in the low-k (<3 Å⁻¹) domain. In order to prevent correlations (Provost et al., 2013), it is possible to fit individual or grouped scattering paths, and various options are proposed to constrain the
coefficient numbers (any linear combination is admitted), effective Debye–Waller factors (multiplier option), distances (distances differing by a fixed distance shift to deal with linear ligands) and \( E_0 \) correction. A ‘batch’ module is available in order to edit, run and analyse the results of multiple fits of input files associated with a constraints editor.

The fitting procedure follows the IXS Standard and Criteria Subcommittee definitions and recommendations (Sayers, 2000) on statistical error estimation (either Student’s statistics on several EXAFS recordings or Fourier noise-filtering methods), reduced \( \Delta \chi^2 \) (quality factor QF) calculated either in \( k \)- or \( R \)-space, parameter correlations and fitting convergence criteria, and F-test calculator (Michalowicz et al., 1999).

6. STRAIGHTNOCHASER

STRAIGHTNOCHASER is dedicated to properly normalized XAFS spectral analysis of mixtures by statistical and linear algebra methods. Linear combination analysis (LCA) by linear least-squares fit (LSF) is used to determine the concentration profile of experimental spectra linearly composed of a known number of known standards (Bevington, 1969; Press et al., 1997). If the number of standards included in the experimental series is a priori unknown, principal component analysis (PCA) allows its evaluation, while target transformation (TT) can evaluate whether a standard is likely to be a component of the series (Klementiev, 2013; Malinowski, 2002; Manceau et al., 2014; Ravel, 2016; Ravel & Newville, 2005; Ressler et al., 2000; Wasserman et al., 1999; Webb, 2005). Multivariate curve resolution–alternating least squares (MCR-ALS) deals with time series of spectra where the number of standard components is known but the nature of these standards is partially unknown (Cassinelli et al., 2014; de Juan et al., 2014; Faro et al., 2011; Jaumot et al., 2005). It is particularly useful to recover the spectra of pure unknown intermediates mixed with other components in the series. Mathematically, the three methods are based on matrix singular values decomposition (SVD) and its applications, matrix pseudo-inversion, rank reduction and vector orthogonal projection (Press et al., 1997). The help manual included in the software contains the user manual, and most of the mathematical definitions and more developed texts with mathematical developments and demonstrations are provided in the package.

6.1. Linear combination analysis (LCA)

The LCA module provides LSF solutions, standard spectra variance–covariance and correlation matrices, statistical errors on coefficients and quality of fit (QF or reduced \( \Delta \chi^2 \)) evaluations (Press et al., 1997). We systematically use the SVD pseudo-inversion since it has been proved to be equivalent to the classical LSF solution but to be much more efficient when standard spectra are correlated. The LCA module also produces the standard-components concentration profile of a series of spectra. The fits can be unconstrained or positive concentration constrained. We propose a search of the more likely standards representing series of experimental mixture spectra among a set of standard spectra by combinatorial LCA with an original sorting ‘progressive’ algorithm (Michalowicz et al., 2013).

6.2. Principal component analysis (PCA)

The PCA of an XAFS data matrix is used to determine the number of independent components and to separate the physical signal from noise (Klementiev, 2013; Malinowski, 2002; Manceau et al., 2014; Ravel, 2016; Ravel & Newville, 2005; Ressler et al., 2000; Wasserman et al., 1999; Webb, 2005). STRAIGHTNOCHASER provides the standard Malinowski’s criteria (Malinowski, 2002). We also propose a new original test: the reconstruction of the data matrix by LSF on the basis of the PCA abstract components and a QF(\( M_d^* \)) plot versus the data indexes, where \( M_d^* \) is the number of abstract components used in each plot (Fig. 1). When the data matrix represents a smooth time evolution of the mixture, as in chemical reactions, the LSF of the data matrix can also be used in score plots of the components of LSF coefficients versus time. The time evolution of the coefficients corresponding to the principal components reflects the smooth evolution of the spectra, while the noisy nature of the components of lower singular values is clearly demonstrated (Fig. 2). The PCA module also provides a tool for determination of the data noise level. The target-transformation (TT) method (Malinowski, 2002) helps to test whether the spectrum of a standard is likely to be represented in the data matrix of experimental mixtures. In addition to TT,

![Figure 1](https://example.com/figure1.png)

**Figure 1**

PCA QF plot of a data matrix (example of a time series of 20 spectra of 400 points each) fitted with one (\( M_d^* = 1 \)), two and three or more abstract components. In this example, to reach good quality factors for the whole set of spectra, it is necessary to introduce at least three components. Adding a fourth component does not improve the fits: components 4–20 represent the noise.
we also provide least-squares fitting of the standards on the principal components basis, which provides equivalent results more quickly. The goodness of the TT and LSF tests are evaluated by the statistical QF and the SPOIL factor.

6.3. Multivariate curve resolution–alternating least squares (MCR-ALS)

*MCR-ALS* (de Juan *et al.*, 2014) is used to recover the spectra of *a priori* unknown intermediate compounds in a time-evolving mixture, even if these intermediates only appear as mixtures during the reaction and cannot be isolated directly in the data set. *MCR-ALS* provides both the calculated spectra of the pure components and the corresponding concentration profiles. While most previously published work was performed with graphical user interfaces linked to commercial mathematical software such as *Matlab* (Jaumot *et al.*, 2005), *STRAIGHTNOCHASER* in *MAX* is totally standalone and free. It provides all of the necessary tools to perform self-consistent cycles of *MCR-ALS*, alternating between LSF estimation of the concentration profile and the spectra basis. To reach convergence, after the choice of the number of pure real components and a first reasonable guess, the lack of spectral and concentration information is compensated by constraints: concentration non-negativity, closure (total concentration equal to 1), average edge jump equal to 1, null pre-edge, and beginning and end spectra optionally fixed. After convergence, the calculated spectra can be used as standards in other LCA studies or classically analysed for a XANES and EXAFS study of the local structures of the unknown intermediate species.

**Figure 2**

Scores plot for abstract components 1–4 of a time-series data matrix. The concentration profiles of abstract components 1–3 reflect the smooth time evolution of the mixture, while component 4 (dotted line) represents noise fluctuations.

7. **SOS XMCD**

*SOS XMCD* is dedicated to XMCD data analysis, providing treatment of both XAS and XMCD spectra in a unique tool. Any ASCII column-format data can be input, with energy entered as eV, keV or even pixels provided that the coefficients for calibration are known. *SOS XMCD* XANES normalization proceeds as in *CHEROKEE*. XMCD baseline definition is based on polynomial regression, with the possibility of defining exclusion zones in the spectra. XMCD normalization directly uses the normalization coefficients defined in the XANES analysis. Simple tools for XANES and XMCD joint analysis are provided.

**References**


