

The AXEAP2 program for $K\beta$ X-ray emission spectra analysis using artificial intelligence

In-Hui Hwang,^a Shelly D. Kelly,^a Maria K. Y. Chan,^{b*} Eli Stavitski,^c Steve M. Heald,^a Sang-Wook Han,^d Nicholas Schwarz^a and Cheng-Jun Sun^{a*}

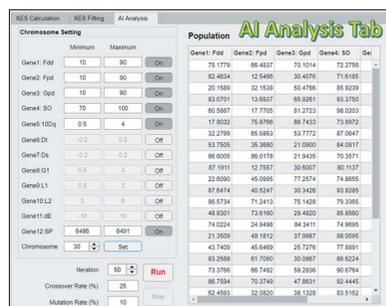
^aX-ray Science Division, Argonne National Laboratory, Lemont, IL 60439, USA, ^bCenter for Nanoscale Materials, Argonne National Laboratory, Lemont, IL 60439, USA, ^cNational Synchrotron Light Source II, Brookhaven National Laboratory, NY 11973, USA, and ^dDepartment of Physics Education and Institute of Fusion Science, Jeonbuk National University, Jeonju 54896, Republic of Korea. *Correspondence e-mail: mchan@anl.gov, cjsun@anl.gov

The processing and analysis of synchrotron data can be a complex task, requiring specialized expertise and knowledge. Our previous work addressed the challenge of X-ray emission spectrum (XES) data processing by developing a standalone application using unsupervised machine learning. However, the task of analyzing the processed spectra remains another challenge. Although the non-resonant $K\beta$ XES of $3d$ transition metals are known to provide electronic structure information such as oxidation and spin state, finding appropriate parameters to match experimental data is a time-consuming and labor-intensive process. Here, a new XES data analysis method based on the genetic algorithm is demonstrated, applying it to Mn, Co and Ni oxides. This approach is also implemented as a standalone application, *Argonne X-ray Emission Analysis 2 (AXEAP2)*, which finds a set of parameters that result in a high-quality fit of the experimental spectrum with minimal intervention. *AXEAP2* is able to find a set of parameters that reproduce the experimental spectrum, and provide insights into the $3d$ electron spin state, $3d-3p$ electron exchange force and $K\beta$ emission core-hole lifetime.

1. Introduction

The non-resonant $K\beta$ X-ray emission spectrum (XES) of $3d$ transition metals is a widely used tool in the investigation of the electronic structure of materials of interest in condensed matter physics, coordination chemistry and catalysis (Vankó *et al.*, 2006; Lafuerza *et al.*, 2020; Kucheryavy *et al.*, 2016). The XES is a result of an electric-dipole-allowed $3p \rightarrow 1s$ transition following an incident photon with sufficient energy to excite a $1s$ electron through the absorption process. The XES shape is known to highly depend on spin state and oxidation state. These characteristics are often collected as a complementary method to X-ray absorption spectroscopy (XAS) and are widely used to elucidate the unique structural and electronic information by comparing spectral shape with known standards samples (Bauer, 2014; Burkhardt *et al.*, 2017; Agote-Arán *et al.*, 2019; Lassalle-Kaiser *et al.*, 2017; Glatzel & Bergmann, 2005; Gretarsson *et al.*, 2013; Pellicieri *et al.*, 2017; Chin *et al.*, 2017).

Atomic multiplet theory and crystal field effects provide a solid theoretical foundation for explaining the XES spectra (Cowan, 1981; de Groot & Kotani, 2008). Atomic multiplet theory calculations are based on the interaction of two electrons and spin-orbital coupling in a many-body system. In $3d$ transition metals, the Coulomb and exchange force of $3d-3d$ and $3d-3p$ electrons are considered, which cause the $K\beta$ emission line to split into a $K\beta_{1,3}$ and a $K\beta'$ line. In the crystal



field effect, the ligands and positively charged metal cation break the degeneracies of the $3d$ electron orbital and separate them into energy levels such as t_{2g} and e_g , leading to differences in the number of pairing electrons and spin state. Usually, the high-spin (HS) spectrum has relatively prominent $K\beta'$ features compared with the low-spin (LS) spectrum, even if the number of $3d$ electrons is the same. Therefore, electron–electron interactions, the number of electrons and the spin state in the $3d$ orbital all play crucial roles in determining the spectral shape.

Our group recently developed a program, *Argonne X-ray Emission Analysis Package (AXEAP)* (Hwang *et al.*, 2022), that was able to dramatically increase the processing speed of raw data by using unsupervised machine learning. However, data analysis remained challenging due to the tedious and difficult process of finding parameters in multiplet spectral simulations that match experimental data, leading to a trial-and-error method for determining them. This problem occurs not only in XES but also in various fields trying to match theory and experimental data. To address this challenge artificial intelligence (AI), such as machine learning (ML) including deep learning, has been applied to data analysis (Zheng *et al.*, 2018; Timoshenko & Frenkel, 2019; Miyazato *et al.*, 2019; Benmore *et al.*, 2022). ML consists of numerical models that can predict output classes or values from input data. This method requires a lot of training data because model parameters must be trained to enable accurate prediction. However, there is insufficient experimental XES data with known ground truth to reliably train these ML models.

In this study, a methodology for analyzing $3d$ transition metal XES and a user-friendly program named *AXEAP2* are presented. The methodology includes three aspects. The first is the calculation of the XES, which is based on the *CTM4XAS* program (Stavitski & de Groot, 2010). The second is convolution after the XES calculation. The third part is parameter optimization driven by a genetic algorithm (GA), which is based upon an evolutionary metaphor that reflects the process of natural selection where the fittest individuals or parameters are selected repeatedly (Mitchell, 1996). This algorithm efficiently finds optimized values in multiple parameter spaces and is widely used to generate high-quality solutions for optimization problems (Terry *et al.*, 2021; Li *et al.*, 2017; Pettersson & Takahashi, 2021). In addition, we developed a viewer function to aid in the interpretation of results. The performance of the methodology was evaluated by analyzing the emission spectra of Mn, Co and Ni oxides.

2. XES measurement and experimental broadening calculation

To test the performance of *AXEAP2*, the XES of MnO, MnCO₃, Mn₂O₃, CoO, LiCoO₂ and LiNiO₂ were measured at

Table 1
Sample information and their experimental energy resolution.

Std is the standard deviation of the HWHM.

Sample	Analyzer	HWHM (eV)	Std (eV)	Formal valence	$3d$ occupancy	Spin state	Symmetry
MnO	Ge (044)	0.393	0.009	2+	$3d^5$	HS	O_h
MnCO ₃	Ge (044)	0.425	0.016	2+	$3d^5$	HS	O_h
Mn ₂ O ₃	Ge (044)	0.414	0.007	3+	$3d^4$	HS	O_h
CoO	Ge (444)	0.470	0.007	2+	$3d^7$	HS	O_h
LiCoO ₂	Ge (444)	0.492	0.034	3+	$3d^6$	LS	O_h
LiNiO ₂	Si (444)	0.552	0.013	2+ or 3+	$3d^8$ or $3d^7$	HS or LS	O_h

the 20-ID-C beamline of the Advanced Photon Source at Argonne National Laboratory, USA. The chemical characteristics of each sample are listed in Table 1 with reference to the literature (Fubini & Stone, 1983; Qiao *et al.*, 2013; Gamblin & Urch, 2001; Siculo *et al.*, 2020; Montoro *et al.*, 1999). These samples were obtained from Aldrich without further purification and selected to compare the differences in the spectrum according to the local environment around the transition metal ion as follows.

- (1) The same element but different number of electrons (MnO and Mn₂O₃).
- (2) The same element and the same number of electrons (MnO and MnCO₃).
- (3) The same element but different number of electrons and different spin state (CoO and LiCoO₂).
- (4) A sample whose valency is not yet clear (LiNiO₂).
- (5) The same number of electrons but different spin state when the LiNiO₂ is trivalent (CoO and LiNiO₂).

A miniature X-ray emission spectrometer (miniXES) was employed to obtain high-resolution spectra within ~ 1 eV (Solovyev *et al.*, 2021; Mattern *et al.*, 2012). The miniXES requires a micro-focused incident X-ray beam with a short working distance, flat analyzer crystals arranged on von Hamos geometry, and two-dimensional position-sensitive detector (2D-PSD) to collect energy-resolved fluorescence signal from the sample. The energies of monochromatic X-rays diffracted from a Si(111) double-crystal monochromator were calibrated to the first derivative peaks of the K -edges of Mn, Co and Ni foils. To focus the beam, Kirkpatrick–Baez mirrors were used, and the focused beam diameter at the sample position was about 50 μm . In order to diffract fluorescence emitted from the sample, eight analyzers were arranged side by side on a von-Hamos circle. Here, Ge(044), Ge(444) and Si(444) satisfying Bragg angles for the $K\beta$ emission range of Mn, Co and Ni, respectively, were used as the analyzers. For the 2D-PSD, a DECTRIS 500K was used and placed at a specific distance according to the optical design. More details about the experimental setup, data acquisition and data processing method are given elsewhere (Hwang *et al.*, 2022).

The broadening of the spectrum, half width at half-maximum (HWHM), due to the 2D-PSD pixel resolution was determined by fitting elastic scattering lines with a Gaussian function. The resolution depends on the pixel size, optical geometry and analyzer (Solovyev *et al.*, 2021). The elastic

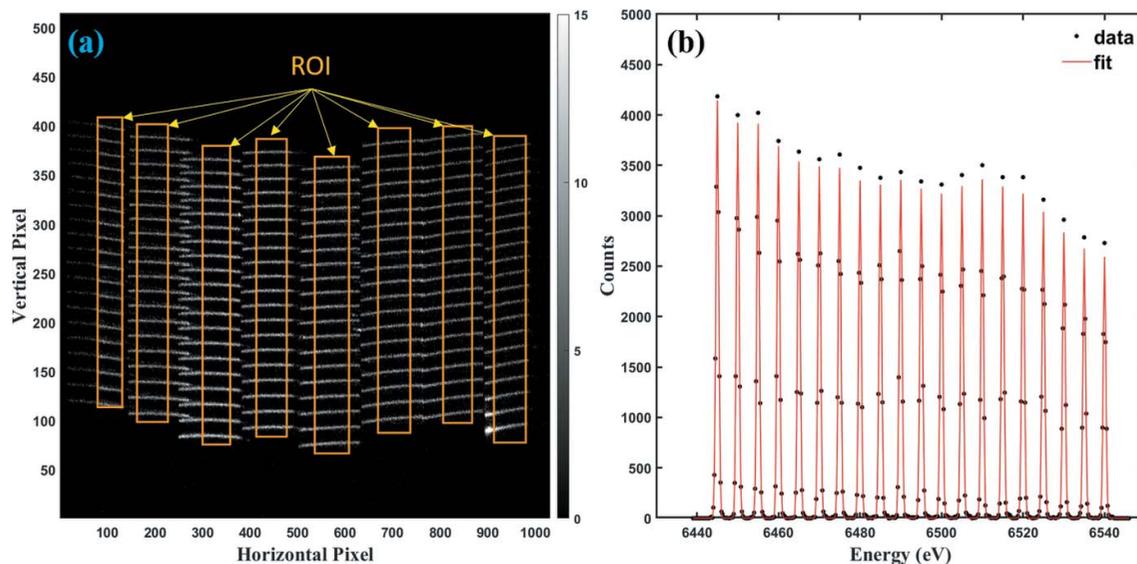


Figure 1

(a) An image of elastically scattered X-rays from MnO on the 2D-PSD for an incident energy range 6445–6540 eV at 5 eV intervals. The horizontal lines with high count rate (white) correspond to each energy and a ROI is set for each of the 20 lines arranged vertically. (b) An example of estimating the energy resolution elastic scattering line used for (a) calibration image with accompanying Gaussian fit.

scatterings were collected during the calibration. *AXEAP* provides the HWHM estimation and Fig. 1 shows an example of the calibration image and its line processed by *AXEAP*. The image shown in Fig. 1(a) is a merged image from 20 series scans corresponding to a Mn $K\beta$ emission range of 6445 to 6540 eV for detector energy calibration. Each horizontal line in the eight regions of interests (ROIs) represents the location of the photons collected using monochromatic X-rays and eight Ge(044) analyzers for 60 s. The image shows that some horizontal lines overlap horizontally with adjacent horizontal lines, which are treated as unusable data occurring in the experimental setup. Using unsupervised ML, *AXEAP* efficiently identifies the overlapping region and recognizes it as unusable data. This allows it to exclude these areas from the ROI. Additionally, *AXEAP* is intelligently designed to exclude a small area near the boundary of the ROI, thereby preventing any potential noise data from being included in the analysis. When energy values are assigned to all pixels in the ROI, the calibration image can be converted to a spectral form as a function of energy as shown in Fig. 1(b). Here, the y-axis represents the total photons counted within eight ROIs. The 20 peaks in Fig. 1(b) correspond to the collected monochromatic X-ray photons, and the HWHM of each peak. The average of the 20 HWHMs is defined as the instrumental broadening originating from experimental errors in this study, and the average HWHMs and standard deviation of all samples were estimated as listed in Table 1, which were used in the convolution process.

3. AXEAP2 features

Our methodology and *AXEAP2* can be divided into three parts: XES calculation, XES fitting and parameter optimization. These three parts are provided as modules within the

main window of *AXEAP2*, accessible via a series of interface windows in the control panel, as shown in Fig. 2.

The XES calculation module provides the ability to perform theoretical XES calculations with input parameters. This module was created by extracting and adapting the source codes of *CTM4XAS* related to XES calculation, ensuring familiarity and ease-of-use for those experienced with *CTM4XAS*, as the interface closely resembles that of *CTM4XAS*.

The XES fitting module is designed to fine-tune the theoretically calculated emission energy and its probability to match the shape of experimental spectra through the use of non-linear least square (NLS) fitting. This module supports Gauss, Lorentz and pseudo-Voigt functions.

Finally, the AI analysis module provides the ability to optimize the parameters used in XES calculation and XES fitting by specifying the optimization range and defining the required values. Upon completion of the optimization process, the results of the calculation and data analysis can be reviewed and confirmed through the GA viewer.

3.1. XES calculation

The theoretical approach in *CTM4XAS* is described thoroughly elsewhere (Cowan, 1981; de Groot & Kotani, 2008; Stavitski & de Groot, 2010). In this paper, a brief overview of the parameters is given.

The starting point of the XES calculation is to consider the electron–electron interaction and spin–orbital coupling for all electrons in an atom. The electron–electron interaction is described using Slater integrals. The Slater integrals are calculated within the Hartree–Fock limit and controlled by the first reduction factor. In this study, the reduction factors

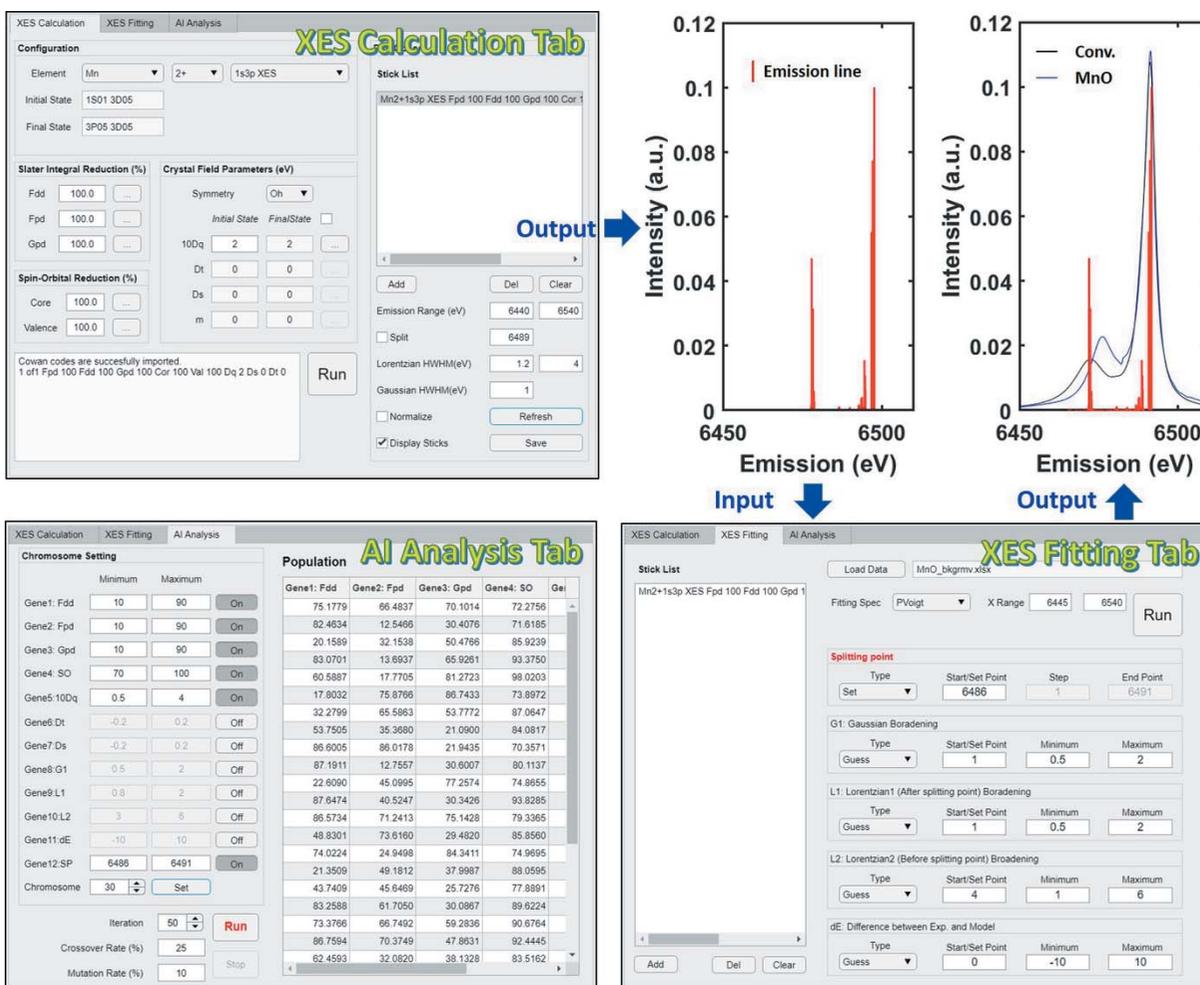


Figure 2 Screen shots of the three modules in *AXEAP2* (XES calculation, XES fitting and AI analysis tabs) and two example output plots. The red vertical lines on the two input plots represent emission probabilities. The blue line and black line are the MnO $K\beta$ XES spectrum and the spectrum convoluted by the XES fitting tab using vertical lines, respectively.

are expressed as F_{dd} ($3d-3d$ Coulomb interaction), F_{pd} ($3d-3p$ Coulomb interaction) and G_{pd} ($3d-3p$ exchange interaction).

The next step is to consider the spin-orbital term including core level and valence level. *CTM4XAS* defines the spin-orbital energy reduction factor as SO in an equivalent manner as the Slater integral reduction factors. We only consider the part of the valence SO optimization parameter because it is assumed that samples studied here have no screening effect on the core level.

The final step is to add a crystal field potential effect. This requires three input variables, which *CTM4XAS* describes as $10Dq$, Dt and Ds (Solomon & Lever, 1999). Varying $10Dq$, Dt and Ds in various symmetries such as octahedral (O_h), tetragonal (D_{4h}) and fourfold (C_4) can yield three spin states: a high spin (HS), an intermediate spin (IS) and a low spin (LS). In the *CTM4XAS* calculation, spin state is determined by $10Dq$ and the pairing energy (or exchange splitting). The pairing is present for every two parallel electrons and can be approximated by a linear combination of the Slater integrals related with F_{dd} (de Groot, 2005). The spin state is a matter

of great attention in $K\beta$ XES because HS tends to have a prominent $K\beta'$ feature relative to LS or IS. The crystal field parameters, therefore, were treated as optimization targets in the same way as atomic parameters.

3.2. XES fitting

In general, when the atomic multiplet theory, spin-orbital effect and crystal field effect are considered in the $K\beta$ main-line calculation for $3d$ transition metals, the *CTM4XAS* program generates up to 100 or more discrete emission probabilities. Convolution can be understood as applying core-hole lifetime and instrumental broadening by applying Lorentzian (L) and Gaussian (G) functions to emission probabilities. To reflect both L and G in the spectrum, the pseudo-Voigt function (V_p) is widely used (Ida *et al.*, 2000). V_p is given by

$$V_p(x, f_v) = \eta L(x, f_v) + (1 - \eta) G(x, f_v) \quad (1)$$

with $0 < \eta < 1$, where η is a function of the full width at half-maximum (FWHM) parameter given by

$$\eta = 1.36603 \left(\frac{f_L}{f_v} \right) - 0.47719 \left(\frac{f_L}{f_v} \right)^2 + 0.11116 \left(\frac{f_L}{f_v} \right)^3, \quad (2)$$

where f_L and f_G are the FWHM of L and G , respectively (Ida *et al.*, 2000). f_v is defined as the FWHM of V_p and is described by

$$f_v(f_L, f_G) = 0.5346 f_L + (0.2166 f_L^2 + f_G^2)^{1/2}. \quad (3)$$

As previously noted, $K\beta_{1,3}$ and $K\beta'$ splitting arises from core-to-core or core-to-valence electron interactions in a transition metal element, and they have different final states. In other words, different Lorentzians should be applied to each emission probability. However, since it is difficult to apply different broadening to 100 or more probabilities as well as theoretical core-hole life calculations, an alternative method of applying different widths from a point called the splitting point (SP) is used (Kucheryavy *et al.*, 2016; Peng *et al.*, 1994). For example, a broadening value L_1 is applied to the emission probability above SP, and L_2 is applied to the emission probability below SP. This is described by

$$f_L = \begin{cases} L_1, & \text{if } SP \geq E_i, \\ L_2, & \text{if } SP < E_i, \end{cases} \quad (4)$$

where E_i is an i th emission energy.

The analysis of $K\beta$ mainlines does not reflect the absolute intensity – only relative intensities and the energy range between $K\beta_{1,3}$ and $K\beta'$ are needed (Vankó *et al.*, 2006). To allow all generated emission probabilities to be adjusted equally according to the normalized spectral intensity, a parameter I_0 is defined. In this study the optimized results of I_0 are not presented. The dE parameter was adopted to correct the difference that occurs in the calibration process or in the calculation.

In summary, to apply the convolution, six parameters must be defined: Gaussian HWHM (G_1), splitting point (SP), Lorentzian HWHM after SP (L_1), Lorentzian HWHM before SP (L_2), energy difference between experiment spectrum and theoretical spectrum (dE), and overall emission intensity adjustment factor (I_0). To ensure that the emission spectrum calculated in theory is generated in close proximity to the experimental spectrum, the NLS fitting method was introduced.

3.3. AI analysis

Fig. 3 shows the workflow for XES parameter optimization. A parameter

to be optimized in the GA is defined as a gene. The atomic parameters and the crystal field parameters are genes, while the convolution parameters can optionally be either genes or fitting parameters. A set of genes is referred to as a chromosome and a set of chromosomes is defined as a population. \hat{y} can be defined as a spectrum function S resulting from the emission probabilities after the convolution through the NLS fitting,

$$\hat{y} = S(g_1, g_2, \dots, g_n), \quad (5)$$

where g_1, g_2, \dots, g_n are genes in a chromosome. The identification strategy is to find the best chromosome to minimize the difference between experimental XES spectra and \hat{y} . The root-mean-square error (RMSE) is the most appropriate indicator for this and is represented by

$$e_{ij} = \left[\frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{T} \right]^{1/2}, \quad (6)$$

where e_{ij} is the RMSE of the j th chromosome on the i th population, and y_t is an experimental spectrum. Selection probability (F_{ij}) helps to create a next generation based on randomly selected chromosomes from the previous generation, which is called fitness (Mitchell, 1996). Then the fitness is defined and estimated by

$$F_{ij} = \frac{(1/e_{ij})}{\sum_{l=1}^n (1/e_{il})}. \quad (7)$$

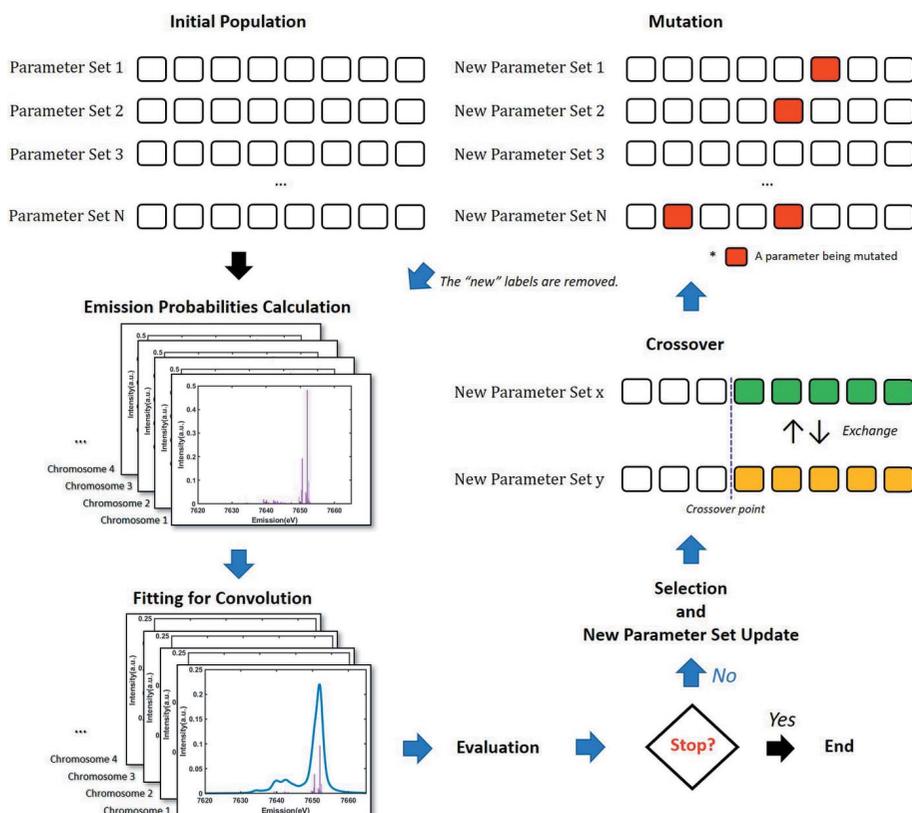


Figure 3 Workflow of the GA-based XES analysis.

The defined fitness increases the probability that a chromosome is selected when the fitness is high.

Selecting chromosomes adopted the roulette wheel method. The area of each chromosome on the roulette board directly reflects the fitness score. The number of chromosomes in all generations is the same because roulette always operates as many chromosomes in a population. Chromosomes can be chosen in duplicate, so if there is a chromosome with an overwhelmingly high fitness score it becomes dominant.

The next step is the crossover. In the previous step, some chromosomes are randomly selected within a population. The selected chromosomes are new chromosomes for the next generation, but they may also have a new combination of genes by exchanging genes (Mitchell, 1996). The exchange of genes between new chromosomes is called crossover. The use of crossover, therefore, can enable GA to search the entire parameter space in parallel, helping it avoid getting trapped in local maxima. The genes to be exchanged are randomly selected within a range that does not exceed the total number of genes in a chromosome.

There is another way to avoid getting trapped in local maxima, namely mutation. The crossover step produces a variety of gene arrangements. Since the selection roulette allows for duplicate selection, the gene pool would tend to become more and more homogeneous as one gene begins to dominate over generations (Mitchell, 1996). Mutation remedies the homogeneity by introducing genetic diversity, *i.e.* replacing genes with new values. This swap of genes also enables the GA to explore a large portion of the parameter space beyond that spanned by the initial population. With each iteration (or population) of the algorithm, some genes in a chromosome will be mutated within a gene's limitation.

3.4. GA viewer

During optimization, *AXEAP2* files all relevant information such as emission probabilities, fitting results and parameters for each spectrum. The GA viewer facilitates the visualization of this information, providing a comprehensive understanding of the results.

The uncertainties of parameters used in NLS fitting can be directly estimated, but the uncertainties of genes cannot be directly obtained. This is because the fitness of a chromosome can only be confirmed after convolution with the NLS fitting process. Instead, the GA viewer supports scatter plots for the parameters generated during GA iteration. GA is based on selection rules such as biological evolution, and, the higher the fitness, the more likely the chromosome is to be inherited by the next generation. The scatter plot is designed to help scientists see the distribution of genes and make inferences about them. An example of using this function is presented in the *Results* section.

Table 2

List of parameters used in GA and NLS fitting.

Parameter	Lower bound	Upper bound	Description
F_{dd}	10%	90%	Reduction factor of $3d-3d$ electrons Coulomb force
F_{pd}	10%	90%	Reduction factor of $3p-3d$ electrons Coulomb force
G_{pd}	10%	90%	Reduction factor of $3p-3d$ electrons exchange force
SO	70%	100%	Spin-orbital coupling reduction factor
$10Dq$	0.5 eV	4.0 eV	Crystal field energy
D_t		–	Crystal field energy
D_s		–	Crystal field energy
SP	$K\beta$ mainline – 5	$K\beta$ mainline + 1	Splitting point to apply different broadening
G_1	Instrumental broadening (Section 2)		Gaussian half width at half-maximum (HWHM)
L_1	1.0 eV	3.0 eV	Lorentzian HWHM after SP
L_2	2.0 eV	8.0 eV	Lorentzian HWHM before SP
dE	–10 eV	+10 eV	Energy difference between experiment and theory
I_0	0.0	1.0	Overall emission intensity adjustment factor

3.5. Programming environment

AXEAP2 is written in MATLAB2021b and works only on Windows OS. This program is freely available, by email. In order to run the program, MATLAB Runtime 2021b is required – a standalone set of shared libraries that enables the execution of compiled *MATLAB* applications or component without license. MATLAB Runtime can be accessed via the MathWorks website (<https://www.mathworks.com/products/compiler/mcr/index.html>).

4. Application of the GA to experimental data

Prior to optimization, all experimental spectra were normalized to have an integral area of 1 within the effective energy range. The orbital configuration of samples during $3p \rightarrow 1s$ transition was determined by the number of electrons as listed in Table 1. For example, Mn^{2+} ($1s3d^5 \rightarrow 3p^53d^5$) and Mn^{3+} ($1s3d^4 \rightarrow 3p^53d^4$) configurations are used for MnO and Mn_2O_3 , respectively. For $LiNiO_2$, since the electronic structure of the Ni ions is still under debate and it is unclear whether it is HS divalent or LS trivalent, two configurations were applied for comparison (Sicolo *et al.*, 2020; Montoro *et al.*, 1999). The parameters F_{dd} , F_{pd} , G_{pd} , $10Dq$, SO and SP were set as genes, and O_h symmetry was used.

To convolute the XES calculation results, we chose the pseudo-Voigt function, which can reflect instrumental and theoretical broadening as described in Section 3.2. The value of the experimental broadening parameter G_1 for each sample was set to the HWHM listed in Table 1, and the upper and lower limits of each parameter are shown in Table 2.

The crossover and mutation rate should be adjusted so that appropriate crossover and mutations can occur in one generation. Excessive crossover and mutations make optimization difficult due to frequent variable changes before becoming dominant, while, in the opposite case, genes are trapped in local optimization and cannot escape easily. Appropriate values for population size, crossover and mutation rates were found and applied through several trials. The use of 20% to 30% for crossover and 10% to 15% for mutation showed stable optimization, and the default values for crossover and mutation rates in *AXEAP2* were set to 25% and

Table 3

Pairing energies calculated by *AXEAP2* and parameter values from the lowest RMSE gene.

	Pairing energy (eV)	F_{dd} (%)	$10Dq$ (eV)	Spin state	G_{pd} (%)	F_{pd} (%)†	SO (%)†	SP (eV)†	L_1 (eV)	L_2 (eV)	dE (eV)
MnO	3.14	73.7	2.24	HS	77.9	29.1	74.0	6486.2	1.64 ± 0.01	3.70 ± 0.06	-4.69 ± 0.01
MnCO ₃	3.14	81.5	2.54	HS	84.4	83.3	89.6	6489.9	1.32 ± 0.01	3.05 ± 0.07	-4.49 ± 0.01
Mn ₂ O ₃	3.25	84.1	2.24	HS	76.6	22.4	84.6	6489.0	1.46 ± 0.03	2.02 ± 0.19	-4.29 ± 0.01
CoO	2.80	32.3	0.82	HS	70.6	86.1	94.6	7648.0	1.37 ± 0.02	4.61 ± 0.11	-3.67 ± 0.01
LiCoO ₂	2.46	28.0	0.86	LS	81.0	62.0	80.4	7644.4	1.51 ± 0.01	3.29 ± 0.25	-3.40 ± 0.01
LiNiO ₂	2.92	56.2	2.40	LS	58.2	52.5	98.5	8264.4	1.84 ± 0.01	4.89 ± 0.16	-4.04 ± 0.01

† No dependence on the optimization process.

10%, respectively. In this study, we utilized the default values set in *AXEAP2* (Hermawanto, 2013). In addition, the number of chromosomes in a population and iterations were set to 30 and 50, respectively, to generate enough data for optimization. It took about 1 h for a desktop computer with a single core of Intel 10th Generation i7 CPU to optimize a spectrum by the above description.

5. Results

In this study, the criterion for an acceptable optimization result was defined as achieving lower than the mean RMSE among populations. The MnO, MnCO₃, Mn₂O₃, CoO, LiCoO₂ and LiNiO₂ spectra and lowest RMSE model are presented in Fig. 4. The parameters for the model with lowest RMSE are listed in Table 3.

It is important to note that our method does not provide uncertainty for the gene type parameter. Instead, an alternative method was used to analyze the relationship between

the RMSE and all parameters generated during the optimization process by comparing them. An example of this is shown in Fig. 5, which displays scatter plots of MnO genes. Points represent all genes generated across all populations, with red markers indicating values with lower than the mean RMSE among all optimized fitting models, which are considered acceptable results of the fitting optimization process.

As shown in Figs. 5(a) and 5(b), the red F_{dd} points are primarily located above the 60% F_{dd} value, while the red $10Dq$ points are concentrated below 2.20 eV. The pairing energy of divalent Mn without reduction is about 3.14 eV, as calculated using *AXEAP2*. When the lowest RMSE F_{dd} value of MnO is applied, the pairing energy is about 2.31 eV (3.14×0.737), which is larger than $10Dq$, so all five electrons of MnO are aligned in spin-up, resulting in HS. In the O_h symmetry, the theoretical XES spectrum becomes HS when the condition (pairing energy) \times ($F_{dd}/100$) $>$ $10Dq$ is satisfied. Fig. 6(a) shows the effect of the spin state by applying the optimized value of MnO listed in Table 3 and changing the $10Dq$ values. If the $10Dq$ value is less than 2.31 eV, which is the electron pairing energy, all become HS. The HS spectra exhibit relatively better fits with RMSE values less than the mean when compared with LS, thus creating empty spaces in the F_{dd} and $10Dq$ scatter plots corresponding to LS.

The overall distribution of G_{pd} points in Fig. 5(c) is V-shaped and converged to 77.9% of the G_{pd} value, meaning that G_{pd} has a strong effect on spectral shape and is well defined by *AXEAP2*. In fact, Fig. 6(b) shows the effect of the G_{pd} value. As the G_{pd} value increases, the $K\beta_{1,3}$ and $K\beta'$ peaks become farther apart. Other parameters SO , F_{pd} and SP show no significant relationship to RMSE as shown in Figs. 5(d), 5(e) and 5(f), respectively, and it was difficult to find a clear role in the optimization results of all samples. Therefore, these values are not determined by *AXEAP2*.

The MnCO₃ spectrum shows a clear separation of $K\beta_{1,3}$ and $K\beta'$, similar to the separation in MnO, and the energy difference between the $K\beta_{1,3}$ and $K\beta'$ of MnCO₃ is about 1.2 eV larger than that of MnO. In addition, when the intensities of the $K\beta_{1,3}$ and $K\beta'$ peaks are compared by normalizing the two spectra, the peak intensity of MnCO₃ is higher than that of MnO. Since MnCO₃ has the same number of electrons as MnO, the pairing energy is the same as that of MnO when there is no reduction. The lowest RMSE reduction factor of the F_{dd} value is 81.5%, and, when this is applied, the final pairing energy is about 2.56 eV (3.14×0.815), which is higher

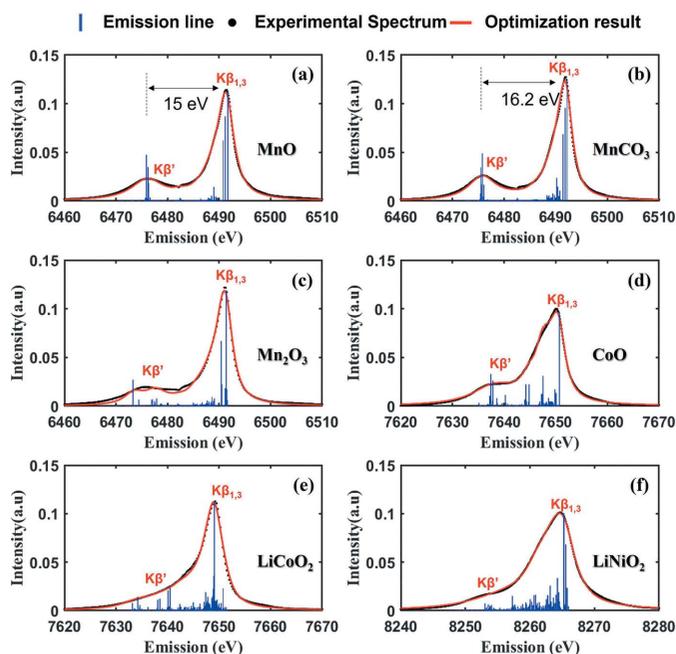


Figure 4
The best results of GA-based optimization for (a) MnO, (b) MnCO₃, (c) Mn₂O₃, (d) CoO, (e) LiCoO₂ and (f) LiNiO₂. The dotted black line, the solid red line and blue bar represent the experiment data, optimized spectrum and emission probability, respectively.

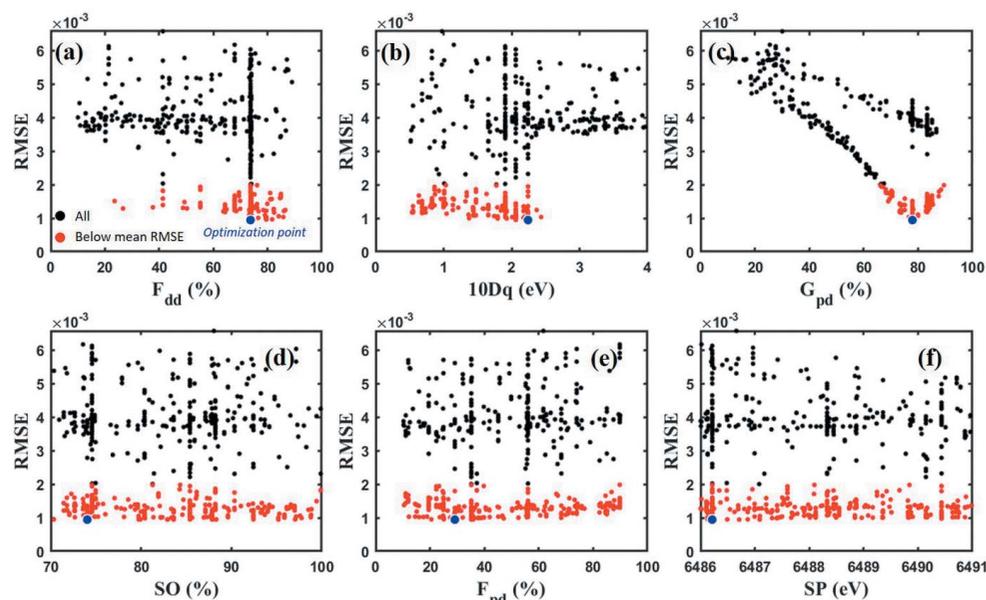


Figure 5 Scatter plot of (a) F_{dd} , (b) $10Dq$, (c) G_{pd} , (d) SO , (e) F_{pd} and (f) SP genes of MnO. The red dots represent genes below the mean RMSE and the blue dot represents the optimized value found by AXEAP2.

than the $10Dq$ value of 2.54 eV for $MnCO_3$. This means that $MnCO_3$ has HS with the same spin alignment as MnO.

The experimental results of Mn_2O_3 show that the $K\beta'$ intensity is weaker than that of MnO, and a mismatch is found between $K\beta'$ experimental and optimization results. Also, L_2 is significantly lower than other optimization results. The

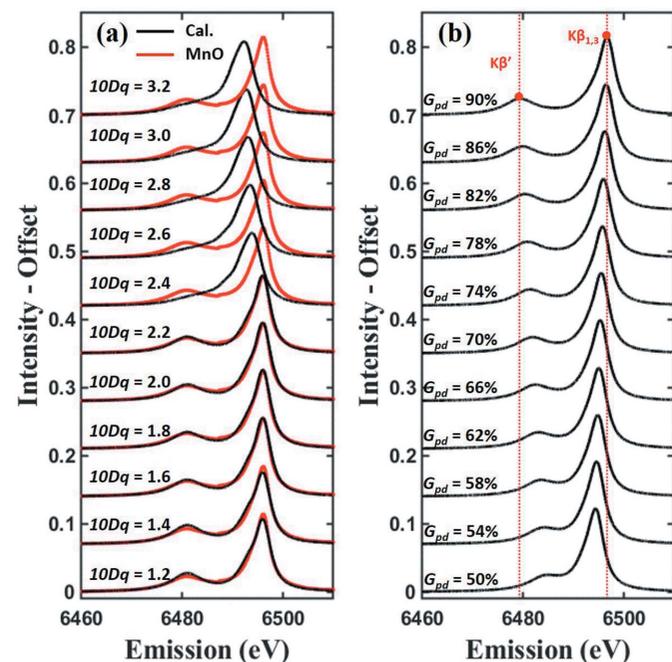


Figure 6 Calculated spectra of Mn^{2+} showing the effect of (a) $10Dq$ and (b) G_{pd} values. Other parameters are set to optimized values of MnO as listed in Table 3. The red lines in (a) are the same MnO spectrum shown in Fig. 4(a) and the black lines in (a) and (b) are calculated spectra. The red dotted vertical lines in (b) are a guide to the eye.

pairing energy of trivalent Mn was calculated using the same method as for MnO and was found to be approximately 2.73 eV (3.25×0.841) in Mn_2O_3 . This value is higher than the $10Dq$ value of 2.24 eV for Mn_2O_3 , resulting in HS.

The experimental CoO spectrum has a higher-intensity $K\beta'$ peak than $LiCoO_2$. The Co ions in CoO and $LiCoO_2$ are known to have seven and six electrons in the $3d$ orbital, respectively. Interestingly, CoO and $LiCoO_2$ had low $10Dq$ values, but AXEAP2 predicted CoO as HS and $LiCoO_2$ as LS. So, four electrons in CoO are paired and three electrons are aligned in spin-up, whereas the six electrons in $LiCoO_2$ are all paired.

The $K\beta$ peak width of $LiNiO_2$ is wider than other measured spectra and L_1 and L_2 also are the highest. As described previously, for this material it is not clear whether the Ni ion is divalent or trivalent. When the divalent and trivalent Ni configurations were applied to $LiNiO_2$, the LS trivalent model was found to be closer. This means that six electrons are paired and one electron is aligned in spin-up. As with $LiCoO_2$, an F_{dd} of 32.5% is small when multiplied by the pairing energy to give a value lower than the low $10Dq$ value for LS.

The uncertainty of L_1 , L_2 and dE can be directly estimated by NLS fitting, indicating a high level of precision and confidence in the results with small uncertainties. The parameter dE , which represents the difference between the optimized spectrum and the experimental spectrum, shows a similar difference when the element species is the same. Negative-optimized values as listed in Table 3 indicate that the energy of the experimental spectrum is shifted to higher energy, compared with the spectrum calculated by AXEAP2.

6. Discussion

AXEAP2 was tested using six spectra to evaluate its performance and capabilities. This program does not provide uncertainty on the parameters used as a gene, so an alternative method was used to understand the results using a comparison method based on the relationship between RMSE and parameter. Theoretical spectra were generated as close as possible to the experimental spectra within the physical/chemical boundaries. A range of values were found with acceptable RMSE values as shown in the scatter plots of Fig. 5 for F_{dd} , $10Dq$ and G_{pd} . The values for SO , F_{pd} and SP did not have a significant effect on the data and were not determined.

The theoretical spin state was determined by comparing the spin pairing energy with the optimized F_{dd} and $10Dq$ values

based on the atomic multiplet theory and crystal field effect. Some of our samples have lowest RMSE $10Dq$ values that are similar to values calculated or used in previous literature, but there are also $10Dq$ values that differ significantly. The *AXEAP2* algorithm reports the lowest RMSE value, although the scatter plots show a range with similar RMSE values that need to be considered. For example, $10Dq$ values for MnO and MnCO₃ have been previously reported to be in the range 1.0–1.4 eV (Qiao *et al.*, 2013; de Groot *et al.*, 1989; Glatzel *et al.*, 2004). *AXEAP2* reports the lowest RMSE value of 2.2 to 2.5 eV (blue dot in Fig. 5) and a range of acceptable values from 0.5 to 2.5 (red dots in Fig. 5). The scatter plots show the lowest RMSE value at one end of this range but with little differentiation between the RMSE values throughout the range. As shown in Fig. 5, points satisfying (pairing energy) $\times (F_{dd} / 100) > 10Dq$ for HS include both the literature reported values and the lowest RMSE values. Figs. 5(a) and 5(b) can be interpreted as a strong trend of making HS for the MnO spectrum by lowering the $10Dq$ value and increasing the F_{dd} value. The spin states of the six samples were predicted and the predicted results were consistent with the references, except for LiNiO₂ where the valency is not yet clear. A clue that the G_{pd} value played a decisive role in a spectral shape was found in a comparison of the scatter plot and the spectral shapes with different G_{pd} values, as shown in Figs. 5 and 6(b). The low uncertainty of L_1 and L_2 indicates that the effect on the core-hole lifetime can be estimated numerically, if the misfit of the spectrum is not large.

6.1. MnO and MnCO₃

Previous studies have shown that both materials have the same number of electrons, valency and spin state. The results of spin states predicted by *AXEAP2* are also consistent with previous studies. What is noteworthy is that the energy difference between $K\beta_{1,3}$ and $K\beta'$ in the two spectra appears differently despite having the same electron configuration. Pollock *et al.* (2014) investigated the characteristics of the $K\beta_{1,3}$ and $K\beta'$ emission spectra. They utilized the restricted active space configuration interaction and crystal field multiplet calculations to determine the decisive parameters in shaping the appropriate broadening of the spectra. They found that the $3p$ – $3d$ exchange interaction parameters have major impact on the spectral shape when the spin state is preserved, while other parameters have minimal influence. Their work provides a reasonable explanation for why the G_{pd} value is the most sensitive compared with other parameters in our findings, suggesting that the $3p$ – $3d$ exchange force plays a key role in contributing to the energy difference between $K\beta_{1,3}$ and $K\beta'$ even in MnO and MnCO₃ spectra with the same electron number and configuration. In fact, the G_{pd} value of MnCO₃ is higher than that of MnO as listed in Table 3.

6.2. MnO (or MnCO₃) and Mn₂O₃

The number of electrons in trivalent Mn is 4, which is one less than for divalent Mn such as MnO. As the number of unpaired electrons increases, the intensity of the $K\beta'$ line is

known to be relatively low even though it has the same spin state. This fact is well reflected on experimental spectra of MnO and Mn₂O₃ and the optimization spectra also show that the overall emission probabilities of trivalent Mn in the $K\beta'$ region is lower than that of the divalent Mn model. However, a discrepancy around the $K\beta'$ region was found between experiment and GA optimization. The discrepancy can be understood in two ways: either the optimization process has not reached the best values during the limited iteration, or the calculation model has not reached optimal state. The first hypothesis was ruled out through ten additional optimizations, all of which produced similar results. Thus, it is likely that the model was not appropriate for the experimental data. For this reason, the low L_2 value of Mn₂O₃, which represents the core-hole broadening effect of the $K\beta'$ line, can be interpreted as an optimized result that attempts to fit the spectral line as closely as possible.

6.3. CoO and LiCoO₂

The comparison of the two spectra represents conditions in which a different spin state is introduced in addition to comparison conditions of MnO and Mn₂O₃. The predicted spin states are consistent with the references, leading to more dramatic changes in $K\beta'$ due to the spin state difference. The $10Dq$ value of 0.7–1.1 eV for CoO can be found in the previous works, but the $10Dq$ value of 0.86 eV for LiCoO₂ is very low compared with the value (~ 2.4 eV) obtained by experiments and calculation (van Elp *et al.*, 1991; Enslin *et al.*, 2014; Iida *et al.*, 2021). Furthermore, F_{dd} for LiCoO₂ was found to be a very low value, which is a result of constructing a lower pairing energy than the $10Dq$ value. This means that F_{dd} and $10Dq$ are more focused on the spin state during optimization and may differ significantly from the actual value, similar to the MnO and MnCO₃ cases, as described above.

6.4. LiNiO₂ and CoO

AXEAP2 concludes that LiNiO₂ is close to trivalent LS. The L_1 and L_2 values of LiNiO₂ are larger than for other spectra, which reflects the relatively wide widths of the $K\beta$ mainlines. Several studies have investigated core-hole broadening in the $K\alpha$ and $K\beta$ lines of the $3d$ transition metals and reported that the width of the spectrum widens as the atomic number increases (Krause & Oliver, 1979; Hölzer *et al.*, 1997; Ito *et al.*, 2018). These differences affect the shape of the two spectra, and the area around $K\beta'$ in particular shows the difference between HS and LS, even though both spectra have the same number of electrons.

7. Summary and conclusion

In this study, we have proposed a method for optimizing the XES parameters using a GA-based algorithm and matching the spectrum between the experimental data and the theoretical model. We set six genes, 30 chromosomes in a population, 25% crossover rate, 10% mutation and 50 iterations per experiment data set. The XES calculation part in the

CTM4XAS program was extracted and re-coded into a friendly package that is easy to use. To convolute the result of the *CTM4XAS* calculation, a pseudo-Voigt function and NLS fitting were introduced. Although a discrepancy was found in the $K\beta'$ of Mn_2O_3 , the overall spectral appearance of the optimized result is very similar to the experimental data within the range of physical and chemical limitations. Also, *AXEAP2* has merit in that it can numerically estimate the spin state, $3d-3p$ electron exchange force, and the broadening effect that arises from the core-hole lifetime, which is difficult to obtain information. Therefore, *AXEAP2* will be of great help in finding an acceptable range of parameters that most strongly determine the spin state and G_{pd} when the experimental environment changes, *i.e.* experimental results with pressure or temperature changes, have not been reported before. Although a GA-based algorithm requires high computational costs, it will be very helpful from the perspective of reducing trial-and-error approaches.

Funding information

This research used resources of the Advanced Photon Source, an Office of Science User Facility operated for the US Department of Energy (DOE) Office of Science by Argonne National Laboratory, and was supported by the US DOE under Contract No. DE-AC02-06CH11357, and the Canadian Light Source and its funding partners. This work is also funded by Argonne National Lab, Lab Directed Research and Development (LDRD-Project: PRJ1007376), and DOE project KC040602 for FWP#35909, 'Integrated Platform for multimodal data capture, exploration and Discovery Driven by AI Tool'. Use of the Center for Nanoscale Materials, an Office of Science user facility, was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences (BES), under Contract No. DE-AC02-06CH11357. MKYC acknowledges the support from the BES Scientific User Facilities (SUF) Division Early Career award.

References

Agote-Arán, M., Lezcano-González, I., Greenaway, A. G., Hayama, S., Díaz-Moreno, S., Kroner, A. B. & Beale, A. M. (2019). *Appl. Catal. Gen.* **570**, 283–291.

Bauer, M. (2014). *Phys. Chem. Chem. Phys.* **16**, 13827–13837.

Benmore, C., Bicer, T., Chan, M. K. Y., Di, Z., Gürsoy, D., Hwang, I.-H., Kuklev, N., Lin, D., Liu, Z., Lobach, I., Qiao, Z., Rebuffi, L., Sharma, H., Shi, X., Sun, C.-J., Yao, Y., Zhou, T., Sandy, A., Miceli, A., Sun, Y., Schwarz, N. & Cherukara, M. J. (2022). *Synchrotron Radiat. News*, **35**(4), 28–35.

Burkhardt, L., Holzwarth, M., Plietker, B. & Bauer, M. (2017). *Inorg. Chem.* **56**, 13300–13310.

Chin, Y.-Y., Lin, H.-J., Hu, Z., Kuo, C.-Y., Mikhailova, D., Lee, J.-M., Haw, S.-C., Chen, S.-A., Schnelle, W., Ishii, H., Hiraoka, N., Liao, Y.-F., Tsuei, K.-D., Tanaka, A., Hao Tjeng, L., Chen, C.-T. & Chen, J.-M. (2017). *Sci. Rep.* **7**, 3656.

Cowan, D. R. (1981). *The Theory of Atomic Structure and Spectra*. Berkeley: University of California Press.

Elp, J. van, Wieland, J. L., Eskes, H., Kuiper, P., Sawatzky, G. A., de Groot, F. M. F. & Turner, T. S. (1991). *Phys. Rev. B*, **44**, 6090–6103.

Enslin, D., Cherkashinin, G., Schmid, S., Bhuvaneshwari, S., Thissen, A. & Jaegermann, W. (2014). *Chem. Mater.* **26**, 3948–3956.

Fubini, B. & Stone, F. K. (1983). *J. Chem. Soc. Faraday Trans. 1*, **79**, 1215–1227.

Gamblin, S. D. & Urch, D. S. (2001). *J. Electron Spectrosc. Relat. Phenom.* **113**, 179–192.

Glatzel, P. & Bergmann, U. (2005). *Coord. Chem. Rev.* **249**, 65–95.

Glatzel, P., Bergmann, U., Yano, J., Visser, H., Robblee, J. H., Gu, W., de Groot, F. M. F., Christou, G., Pecoraro, V. L., Cramer, S. P. & Yachandra, V. K. (2004). *J. Am. Chem. Soc.* **126**, 9946–9959.

Gretarsson, H., Saha, S. R., Drye, T., Paglione, J., Kim, J., Casa, D., Gog, T., Wu, W., Julian, S. R. & Kim, Y. (2013). *Phys. Rev. Lett.* **110**, 047003.

Groot, F. de (2005). *Coord. Chem. Rev.* **249**, 31–63.

Groot, F. M. F. de, Grioni, M., Fuggle, J. C., Ghijsen, J., Sawatzky, G. A. & Petersen, H. (1989). *Phys. Rev. B*, **40**, 5715–5723.

Groot, F. M. F. de & Kotani, A. (2008). *Core Level Spectroscopy of Solids*. CRC Press.

Hermawanto, D. (2013). *arXiv:1308.4675*.

Hölzer, G., Fritsch, M., Deutsch, M., Härtwig, J. & Förster, E. (1997). *Phys. Rev. A*, **56**, 4554–4568.

Hwang, I.-H., Solovyev, M. A., Han, S.-W., Chan, M. K. Y., Hammonds, J. P., Heald, S. M., Kelly, S. D., Schwarz, N., Zhang, X. & Sun, C.-J. (2022). *J. Synchrotron Rad.* **29**, 1309–1317.

Ida, T., Ando, M. & Toraya, H. (2000). *J. Appl. Cryst.* **33**, 1311–1316.

Iida, S.-I., Terashima, M., Mamiya, K., Chang, H.-Y., Sasaki, S., Ono, A., Kimoto, T. & Miyayama, T. (2021). *J. Vac. Sci. Technol. B*, **39**, 044001.

Ito, Y., Tochio, T., Yamashita, M., Fukushima, S., Vlaicu, A. M., Syrocki, Ł., Ślabkowska, K., Weder, E., Polasik, M., Sawicka, K., Indelicato, P., Marques, J. P., Sampaio, J. M., Guerra, M., Santos, J. P. & Parente, F. (2018). *Phys. Rev. A*, **97**, 052505.

Krause, M. O. & Oliver, J. H. (1979). *J. Phys. Chem. Ref. Data*, **8**, 329–338.

Kucheryavy, P., Lahanas, N., Velasco, E., Sun, C.-S. & Lockard, J. V. (2016). *J. Phys. Chem. Lett.* **7**, 1109–1115.

Lafuerza, S., Carlantuono, A., Retegan, M. & Glatzel, P. (2020). *Inorg. Chem.* **59**, 12518–12535.

Lassalle-Kaiser, B., Gul, S., Kern, J., Yachandra, V. K. & Yano, J. (2017). *J. Electron Spectrosc. Relat. Phenom.* **221**, 18–27.

Li, Y., Li, H., Pickard, F. C., IV, Narayanan, B., Sen, F. G., Chan, M. K. Y., Sankaranarayanan, S. K. R. S., Brooks, B. R. & Roux, B. (2017). *J. Chem. Theory Comput.* **13**, 4492–4503.

Mattern, B. A., Seidler, G. T., Haave, M., Pacold, J. I., Gordon, A., Planillo, J., Quintana, J. & Rusthoven, B. (2012). *Rev. Sci. Instrum.* **83**, 023901.

Mitchell, M. (1996). *An Introduction to Genetic Algorithms*. Cambridge: MIT Press.

Miyazato, I., Takahashi, L. & Takahashi, K. (2019). *Mol. Syst. Des. Eng.* **4**, 1014–1018.

Montoro, L. A., Abbate, M., Almeida, E. C. & Rosolen, J. M. (1999). *Chem. Phys. Lett.* **309**, 14–18.

Pellicciari, J., Huang, Y., Ishii, K., Zhang, C., Dai, P., Chen, G. F., Xing, L., Wang, X., Jin, C., Ding, H., Werner, P. & Schmitt, T. (2017). *Sci. Rep.* **7**, 8003.

Peng, G., deGroot, F. M. F., Haemaelaenen, K., Moore, J. A., Wang, X., Grush, M. M., Hastings, J. B., Siddons, D. P. & Armstrong, W. H. (1994). *J. Am. Chem. Soc.* **116**, 2914–2920.

Petterson, L. G. M. & Takahashi, O. (2021). *Theor. Chem. Acc.* **140**, 162.

Pollock, C. J., Delgado-Jaime, M. U., Atanasov, M., Neese, F. & DeBeer, S. (2014). *J. Am. Chem. Soc.* **136**, 9453–9463.

Qiao, R., Chin, T., Harris, S. J., Yan, S. & Yang, W. (2013). *Curr. Appl. Phys.* **13**, 544–548.

Sicolo, S., Mock, M., Bianchini, M. & Albe, K. (2020). *Chem. Mater.* **32**, 10096–10103.

Solomon, E. I. & Lever, A. B. (1999). *Inorganic Electronic Structure and Spectroscopy*. New York: Wiley.

- Solovyev, M. A., Lockard, J. V., Huang, X., Heald, S. M. & Sun, C.-J. (2021). *Rev. Sci. Instrum.* **92**, 073105.
- Stavitski, E. & de Groot, F. M. F. (2010). *Micron*, **41**, 687–694.
- Terry, J., Lau, M. L., Sun, J., Xu, C., Hendricks, B., Kise, J., Lnu, M., Bagade, S., Shah, S., Makhijani, P., Karantha, A., Boltz, T., Oellien, M., Adas, M., Argamon, S., Long, M. & Guillen, D. P. (2021). *Appl. Surf. Sci.* **547**, 149059.
- Timoshenko, J. & Frenkel, A. I. (2019). *ACS Catal.* **9**, 10192–10211.
- Vankó, G., Rueff, J., Mattila, A., Németh, Z. & Shukla, A. (2006). *Phys. Rev. B*, **73**, 024424.
- Zheng, C., Mathew, K., Chen, C., Chen, Y., Tang, H., Dozier, A., Kas, J. J., Vila, F. D., Rehr, J. J., Piper, L. F. J. & Persson, K. A. (2018). *NPJ Comput. Mater.* **4**, 12.