Surface and Interface Analysis

WILEY

SURFACE and INTERFACE ANALYSIS

RESEARCH ARTICLE OPEN ACCESS

KherveFitting: An Open Source Software for Fitting X-Ray Photoelectron Spectroscopy Data

Gwilherm Kerherve[1] | William S. J. Skinner[1] | Julian A. Hochhaus[2] | Arthur Graf[3,4] | David J. Morgan[3,4] | Mark A. Isaacs[4,5] | Benjamen P. Reed[6] | Hideki Nakajima[7] | David J. Payne[1]

[1]Department of Materials, Imperial College London, London, UK | [2]Department of Physics/DELTA, TU Dortmund University, Dortmund, Germany | [3]School of Chemistry, Cardiff University, Cardiff, UK | [4]HarwellXPS, Research Complex at Harwell, Didcot, UK | [5]Department of Chemistry, University College London, London, UK | [6]National Physical Laboratory, Hampton Rd, Teddington, UK | [7]Synchrotron Light Research Institute, Nakhon, Ratchasima, Thailand

Correspondence: Gwilherm Kerherve (g.kerherve@ic.ac.uk)

Received: 27 September 2025 | **Revised:** 4 November 2025 | **Accepted:** 12 November 2025

Keywords: curve fitting | data analysis | ESCA | peak fitting | quantification | XPS | X-ray photoelectron spectroscopy

ABSTRACT

KherveFitting is a software written in Python and designed for fitting X-ray photoelectron spectroscopy (XPS) data. It provides a user-friendly graphical interface for researchers in materials science and surface analysis, offering features that include multiple background subtraction methods, various peak fitting models, and automated multipeak fitting with customisable constraints. This software aims to fill a gap in the market for XPS analysis tools, providing a freely available solution for users. This paper describes the software architecture and its theoretical background, including fitting algorithms and peak models. Its performance is validated by directly comparing quantification accuracy and peak fitting capabilities against established commercial software packages, demonstrating robust and reliable results across diverse test cases including ionic liquids, conducting oxides, and rare earth materials.

Introduction

X-ray photoelectron spectroscopy (XPS) is a widely used surface analysis technique that provides crucial information on the elemental composition and chemical state of materials. Peak fitting of XPS spectra is essential to extract both quantitative and qualitative data, but the process is often complex and time-consuming. At present, available XPS fitting software is largely Windows-based and proprietary, restricting both accessibility and customization. The most widely known XPS data processing tools include Thermo Avantage [1], SDP [2], UniFit [3], XPSPEAK41 [4], AAnalyzer [5], and CasaXPS [6], with CasaXPS being the most commonly used software in the XPS community, offering a breadth of functionality for data analysis, making it a powerful tool for experts in the field. However, while its wide-ranging capabilities are powerful for advanced users, it can be overwhelming for a novice or nonexpert user. Occasional XPS users may find themselves navigating a vast array of features, even when their needs are limited to basic peak fitting without extensive data analysis.

To address these challenges, we have developed KherveFitting or LG4X-V3, an open-source software written entirely in Python. KherveFitting is an evolution of the LG4X-V1 [7]. and LG4X-V2/lmfitxps [8, 9]. projects, which feature a user-friendly graphical user interface (GUI) designed to simplify the XPS peak fitting process by (1) enabling efficient data input and output through Microsoft Excel integration, (2) providing a range of background correction and peak shape models, (3) offering flexible peak constraints, and (4) incorporating numerous shortcuts to improve

Abbreviations: ESCA, electron spectroscopy for chemical analysis; XPS, X-ray photoelectron spectroscopy.

workflow. Together, these features significantly reduce the time required for analysis and enhance the user experience, making it as intuitive and straightforward as possible for quick, easy, and efficient peak fitting of data and subsequent preparation of figures for presentations and publications.

Software Characteristics

Architecture and Front-End

KherveFitting is implemented in Python, leveraging several powerful libraries to provide a robust and user-friendly experience. The software architecture is modular, consisting of the following main components:

- Graphical User Interface (GUI): Implemented using wxPython [10], providing a cross-platform (Windows, macOS, Linux) interface for user interaction.
- Data Visualization: Matplotlib [11]. is utilized to plot and visualize XPS data and fitted results, giving high customisability.
- Numerical Computation: NumPy [12]. is employed for efficient numerical operations, along with lmfit [13]. and lmfitxps [9]. for curve fitting algorithms.
- Data Handling: Utilizes Pandas [14]. and openpyxl for manipulating Excel files, enabling seamless input and output of data.
- Core Functionality: Contains modules for background subtraction, peak fitting, binding energy correction, survey peak identification, and labeling.
- Other Functionality: Copy, crop, merge, differentiate, integrate, shift, add, subtract, multiply data.

The software follows an event-driven programming model, where user interactions trigger appropriate responses in the data processing and visualization components. A central data model maintains the state of the XPS data and fitted parameters, ensuring consistency across different views and operations.

Figure 1 presents an annotated image of the software's front end. The interface is divided into three sections: the plot (left side), the peak fitting parameter grid (top right), and the peak results grid (bottom right). Two toolbars are also available. The horizontal toolbar includes tools to open or save files, select core levels, toggle various plot components, create an area under a shape, generate a peak model, calculate the noise level, identify elements within a survey spectrum (wide scan), and export the peak fitting data to the results grid. The vertical toolbar allows users to adjust the X and Y plot limits. Various keyboard shortcuts provide quick access to most of these tools, thereby accelerating the fitting process.

Process Workflow

The workflow is depicted in Figure 2. Users begin by opening a KherveFitting-formatted Excel file (see manual for specific formatting requirements), either by dragging and dropping it onto the plot section of the GUI or by clicking the open button on the horizontal toolbar. The software supports importing Vamas (.vms), Thermo (.avg), Kratos (.kal), Phi (.spe). csv (.csv), Ascii (.asc or .txt), and Excel files (.xls, .xlsx) exported from Avantage, which are automatically converted

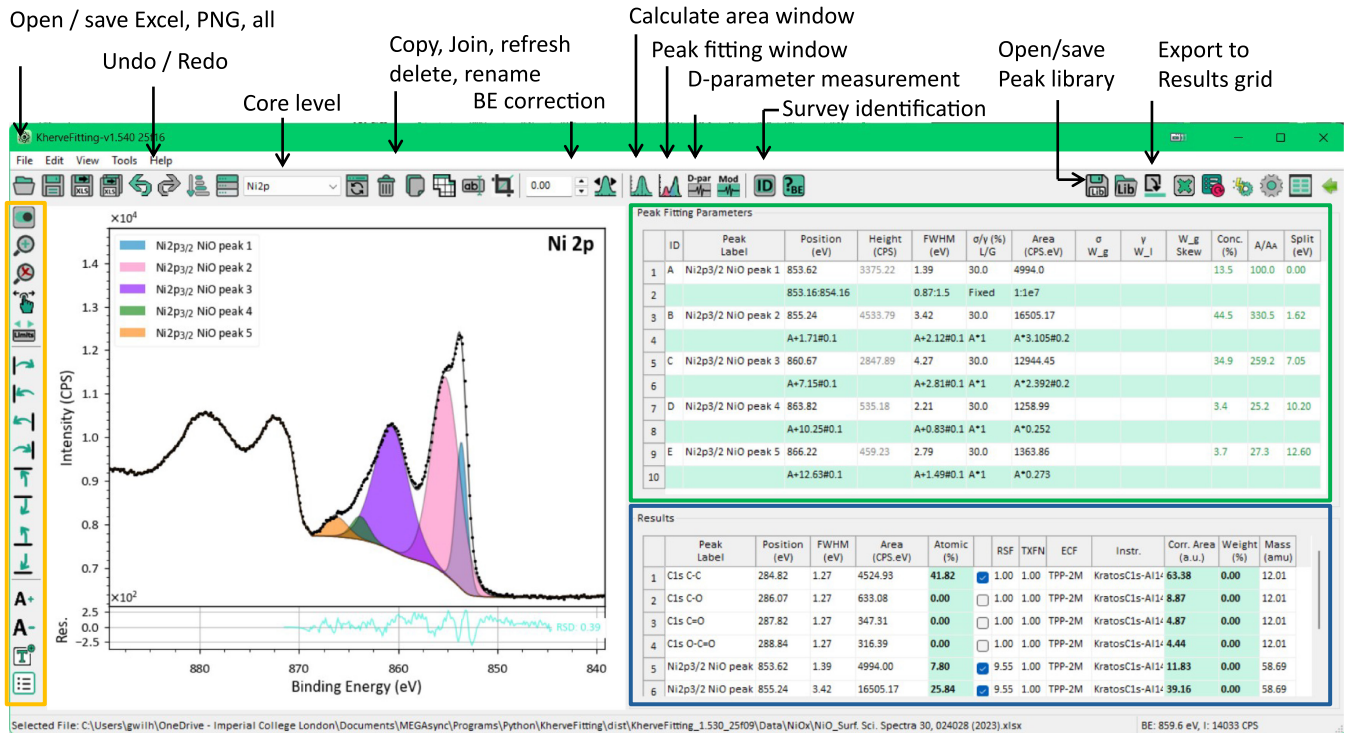

FIGURE 1 | GUI of KherveFitting. The green box area indicates the peak fitting grid in which all peaks applied to this core level are defined. The green rows represent the constraint rows. The blue box area indicates the results grid in which the atomic concentrations for each element are calculated. The orange box around the left vertical toolbar indicates the location of the plot controls: zoom, labels edit, and toggles.

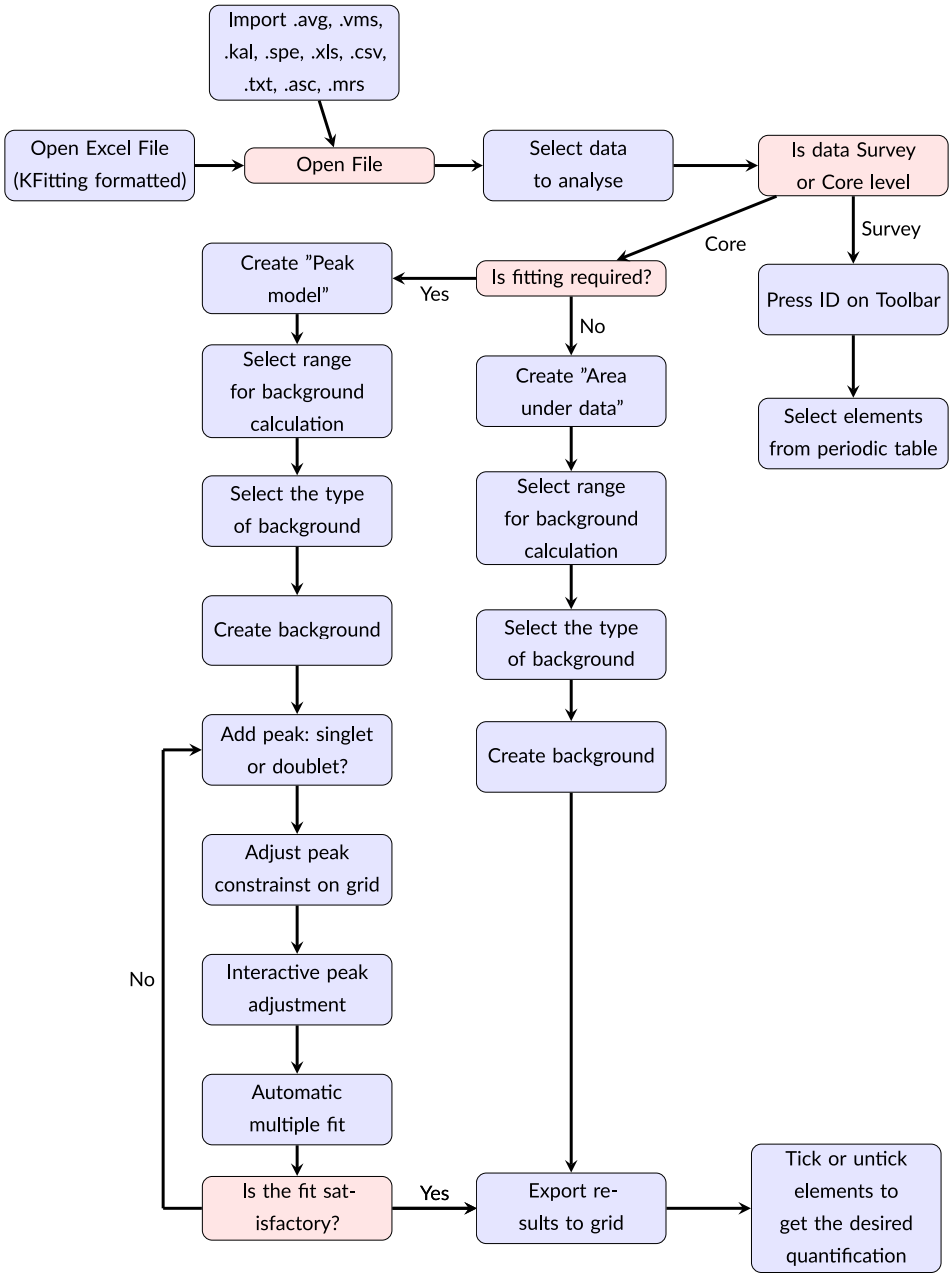

FIGURE 2 | Workflow to analyze XPS data using KherveFitting.

into an appropriately-formatted (Kfitting) Excel data file and a corresponding JSON configuration file. The JSON file contains the peak properties for all core levels present in the Excel data file.

Visualization and comparison of multiple samples is available through the sample manager window; see Figure 3. This manager offers various tools such as plotting multiple core levels together, normalizing to specific binding energy, copying, renaming, or summing core levels.

Once a dataset has been imported, the desired narrow/wide scan to be analyzed can be selected either by scrolling the mouse wheel or by choosing it from the drop-down menu, as illustrated in Figure 1. Depending on the data type, users can identify elements in the spectrum (surveys or wide scan spectra) or create a peak model to identify the different chemical species present

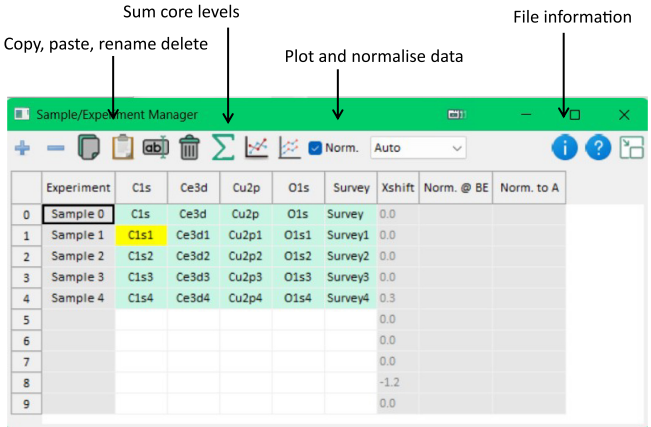

FIGURE 3 | Representation of the sample manager window. The samples loaded in the manager are Cu-doped CeO_2. The narrow scans are C 1*s*, O 1*s*, Cu 2*p* and Ce 3*d*, as well as a survey scan.

(core level spectra) and determine the composition of the sample using their relative areas.

For element identification in a survey, a periodic table is available via the "ID" button on the horizontal toolbar. Clicking on an element in the periodic table displays its binding energy along with the respective intensity ratio related to the main core line. An example survey scan of a graphite sample is presented in Figure 4.

When creating a peak model, the peak fitting window (Figure 5) first prompts the fitting of a background over a selected binding energy range, with models that currently include linear, Shirley [15], Smart, and Tougaard. Further details of each of these methods are included in the Background Models section.

Once the background is established, the user can create singlet or doublet peaks of different shapes. The peak shapes available in KherveFitting are based on combinations of Gaussian and Lorentzian functions, which are widely used in XPS peak fitting due to their ability to model both intrinsic line shapes and instrumental broadening effects. A number of different peak models are currently available in KherveFitting—details of the most commonly applied peak fitting models are included in the Peak Models section, and further details of all available peak models

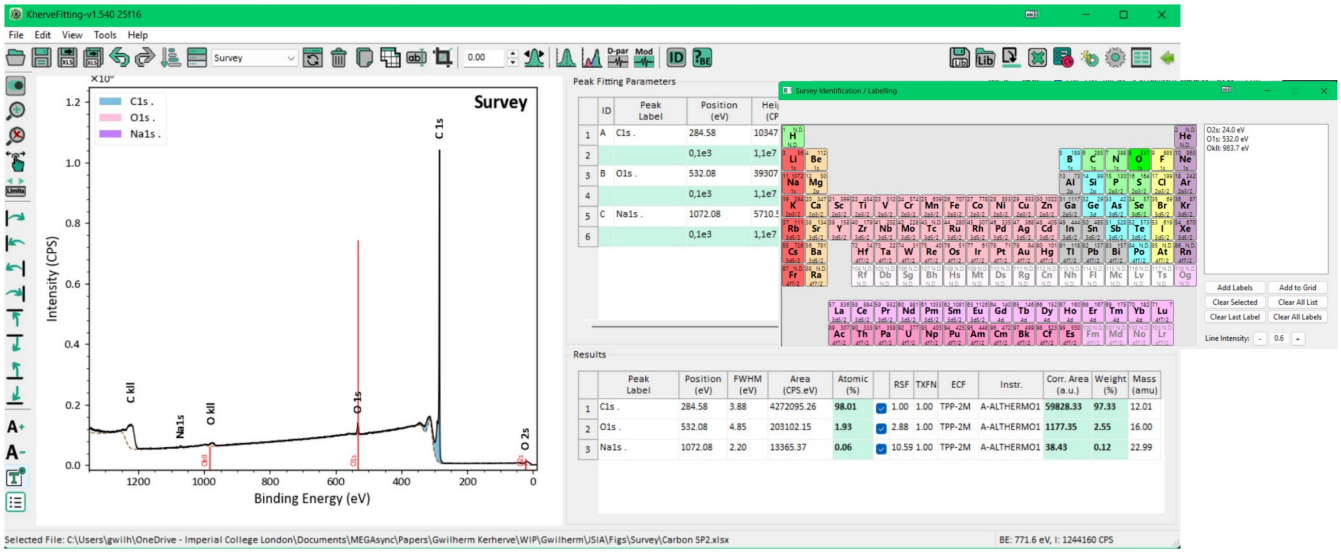

FIGURE 4 | Survey spectrum of a hard carbon sample showing the presence of O, C, and Na. KherveFitting allows the identification, labeling, and measurement of the area contribution of each elements.

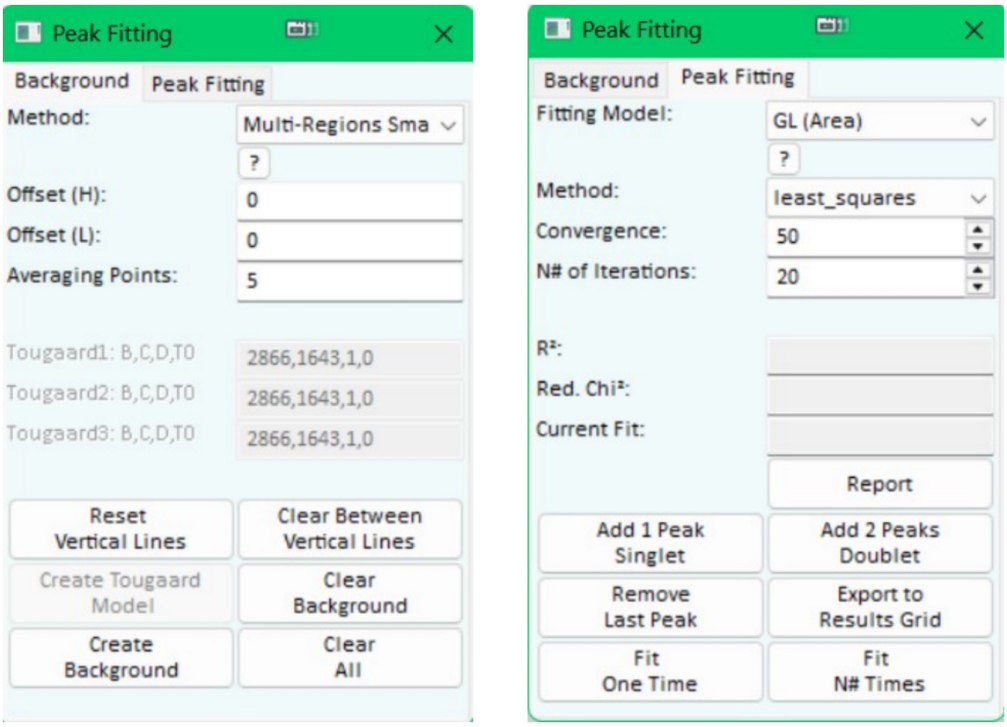

FIGURE 5 | Representation of the fitting screen window consisting of the background tab (left) and the fitting tab (right).

are included in the KherveFitting manual. Users can choose the appropriate peak shape and adjust parameters via the peak fitting parameter grid, with real-time visualization of how changes affect the overall fit.

The peak fitting parameter grid, as shown in Figure 1, allows users to modify the properties of each peak and to set constraints. Furthermore, peak positions, heights, and full-width-at-half-maximum (FWHM) can be interactively adjusted through cursor movement.

Once the peak model is created, the user can fit the data by clicking the fit button. By default, fitting is performed using the Levenberg-Marquardt algorithm [16, 17], as implemented in the LMFIT Python library [13]. The goodness of fit (R^2, reduced Chi2) is shown in the peak fitting parameter grid, while the residual standard deviation (RSD) is displayed in the residual plot. Users can further refine the fit by manually adjusting peak parameters or by adding and removing peaks as necessary. The undo/redo functionality makes it easy to experiment with different fitting approaches.

This process can be repeated for multiple core levels within the same file. The software maintains consistency across core levels, allowing for a thorough analysis of complex XPS data.

Once a satisfactory fit is achieved for all core levels, users can apply a binding energy correction if required. This is typically done by identifying a reference peak (e.g., C 1*s* C-C at 284.8 eV), which applies the calculated offset to all core levels in the same row of the grid. KherveFitting includes a binding energy correction function that searches for the label "C1*s* C-C" peak within the peak fitting grids and adjusts all core levels accordingly. The name and binding energy of the reference peak can be customized in the preference window to any name and value.

Finally, further data analysis, such as calculating atomic concentrations or comparing chemical states across different core levels, can be performed on fitted data by exporting to the results grid (see Figure 1). Atomic concentrations are calculated using the relative sensitivity factors for elements and other correction factors, which can be defined in the preference window. These correction factors will be discussed in further detail in the theoretical background section. The peak-fitted data, fitting parameters, and plots for each core level are saved in an Excel format, which is compatible with other plotting software, enabling seamless integration into existing workflows.

Theoretical Background

Fitting Algorithm

The fitting algorithms available in KherveFitting, provided through the LMFIT library [13], are outlined in Table 1.

By default, fitting is performed using the Levenberg-Marquardt algorithm [16, 17], as implemented in the LMFIT Python library [13]. For further details of the available fitting algorithms, please see the LMFIT documentation.

TABLE 1 | Optimization methods available in KherveFitting via the LMFIT library.

Method	Description
Levenberg–Marquardt (leastsq)	Efficient for least-squares problems
Least-squares	Provides multiple algorithmic options for flexibility
Powell	Derivative-free, ideal for complex problems
Nelder–Mead	Simplex algorithm, robust for noisy data
Cobyla	Handles constrained optimization without requiring derivatives
Trust-region	Optimizes nonlinear constraints efficiently

Background Models

KherveFitting offers a variety of background models, namely linear, Shirley, Smart, and Tougaard. The linear background is the simplest of these models, with the background fitted as a straight line between the range markers. This background model is typically only appropriate in the case where the background decreases with increasing binding energy. When the background increases with increasing binding energy, either a Shirley or Tougaard background should be fitted. The default option in KherveFitting is the Smart background, which applies a linear background if the background decreases with increasing binding energy or a Shirley background if the background increases.

The Shirley background [15]. is an approach that leverages information about the spectrum to construct a background sensitive to changes in the data and is defined as follows:

$$B_S(E) = I_{Low} + \kappa \frac{A_{Low}(E)}{A_{High}(E) + A_{Low}(E)}. \quad (1)$$

In this equation, I_{Low} and I_{High} are the intensity values at the low and high binding energy endpoints and define the boundary conditions for the background calculation. κ is the step in the background, typically equal to the difference $(I_{High} - I_{Low})$, which defines the total background intensity change across the spectral region. $A_{Low}(E)$ and $A_{High}(E)$: The integrated areas from energy E to the low and high binding energy endpoints, respectively, which are used to determine the background contribution at each energy point. The key aspect of the Shirley algorithm is its iterative nature for determining the background. Since the quantities $A_{High}(E)$ and $A_{Low}(E)$ require knowledge of the background $S(E)$, which is initially unknown, the Shirley background calculation becomes an iterative process. The integrated areas $A_{High}(E)$ and $A_{Low}(E)$ for each energy point must first be computed using an initial approximation of $S(E)$, after which the background $S(E)$ is updated based on these calculations, and the entire process is repeated until convergence is achieved.

The Tougaard U4 background was first introduced and developed by Sven Tougaard [18]. and is a widely used model in XPS to describe the inelastic scattering of electrons as they exit a sample. Unlike the simpler Shirley background, which assumes a monotonic buildup of signal due to electron scattering, the Tougaard background incorporates a more physically accurate model by accounting for energy losses that photoelectrons experience due to inelastic scattering processes. Firstly, the Tougaard inelastic scattering cross-section is expressed as follows:

$$U(E;B,C,D,T_0)=\begin{cases}\dfrac{B\cdot E}{(C-E)^2+D\cdot E^2} & \text{if } E\geq T_0,\\ 0 & \text{otherwise.}\end{cases} \tag{2}$$

In this equation, B is a scaling factor that adjusts the overall intensity of the background. The C-parameter is closely tied to the position of the maximum energy loss in the scattering tail, depending on material properties, and is associated with the excitation of plasmons. D controls the broadening or width of the scattering tail, reflecting the spread of inelastic scattering energy losses. T_0 is the energy threshold below which the Tougaard background is zero, corresponding to regions where no inelastic scattering is expected. The final Tougaard background B_T(E) at energy E can then be expressed as follows:

$$B_T(E)=I_{Low}+\int_E^{\infty}U(E'-E;B,C,D,T_0)\cdot I(E')\,dE' \tag{3}$$

where I_{Low} is the baseline intensity and I(E') is the measured intensity after baseline subtraction.

Peak Models

There are a variety of models available in KherveFitting. This section outlines the most commonly used models currently implemented; however, for further details of all available peak models, please see the latest version of the KherveFitting manual.

The Gaussian G(E) and Lorentzian L(E) line shape used by the different peak methods are defined as follows:

$$\begin{aligned}&G(E;E_c,F,H,lg)\\ &=\ H\times\exp\left[-4\cdot\ln 2\times(1-lg)\cdot\left(\frac{E-E_c}{F}\right)^2\right],\end{aligned} \tag{4}$$

and

$$L(E;E_c,F,H,lg)=\frac{H}{1+4\cdot lg\cdot\left(\frac{E-E_c}{F}\right)^2}, \tag{5}$$

where F is the full width at half maximum of the model and E_c is the center position (in binding energy) of the peak, H is the height of the peak, and lg is the ratio between Lorentzian and Gaussian. The term lg is used only in the GL and SGL model. A full Gaussian peak ($lg=0$) has a symmetric, bell-shaped profile, while a full Lorentzian peak ($lg=1$) is a little narrower at its apex and extends out further on its peak tails.

The simplest peak method used by KherveFitting are the product GL(E) and sum SGL(E) of Gaussian and Lorentzian:

$$\begin{aligned}GL(E;E_c,F,H,lg)\ &=H\times L(E;E_c,F,1,lg)\\ &\quad\times G(E;E_c,F,1,lg),\end{aligned} \tag{6}$$

and

$$\begin{aligned}SGL(E;E_c,F,H,lg)\ &=H\cdot(lg\cdot L(E;E_c,F,1,1)\\ &\quad+(1-lg)\cdot G(E;E_c,F,1,0)).\end{aligned} \tag{7}$$

where lg, F, E_c, and H are parameters already defined above.

The asymmetric Lorentzian LA(E) is a Lorentzian line shape raised by the power σ and γ. It is a pure Lorentzian model if both σ and γ are equal to 1 and is symmetric if σ is equal to γ. This model lacks physical significance but is comparable to the LA(α, β, m) model used in CasaXPS software when the Gaussian width parameter m is maintained at sufficiently low values to minimize its contribution to the lineshape [6]. The LA(E) is expressed as follows:

$$\mathrm{LA}(E;H,E_c,F,\sigma,\gamma,0)=\begin{cases}\dfrac{H}{\left[1+4\cdot\left(\frac{E-E_c}{F}\right)^2\right]^{\gamma}} & E\leq E_c\\ \dfrac{H}{\left[1+4\cdot\left(\frac{E-E_c}{F}\right)^2\right]^{\sigma}} & E>E_c,\end{cases} \tag{8}$$

where α and β are the asymmetry parameters on the higher and lower binding energy sides of the peak, respectively. The higher σ or γ is, the more Gaussian the peak is on the higher or lower binding energy side.

Another variation of the LA model called LA∗G(E) is the LA lineshape, which further convolutes it with a Gaussian lineshape of width W_g. This model corresponds to the CasaXPS LA(α, β, m) model, which is described as being a superset of Voigt functions, as this model can create asymmetric and symmetric Voigt-like lineshapes. The LA∗G(E) is expressed as follows:

$$\begin{aligned}&\mathrm{LA}*\mathrm{G}(E;H,E_c,F,\sigma,\gamma,W_g)\\ &=\mathrm{LA}(E;H,E_c,F,\sigma,\gamma,0)*G(E;E_c,W_g,H,0).\end{aligned} \tag{9}$$

The Voigt function is a combination of a Gaussian function convolved with a Lorentzian function and is provided directly from the LMFIT. The Voigt model can be expressed in simple terms as follows:

$$V(E;E_c,\sigma,\gamma)=\int_{-\infty}^{\infty}L(E';E_c,\gamma)\cdot G(E-E';0,\sigma)\,dE', \tag{10}$$

where the Lorentzian and Gaussian parts of the Voigt are expressed as follows:

$$L(E;E_c,\gamma)=\frac{\gamma/\pi}{(E-E_c)^2+\gamma^2}, \tag{11}$$

$$G(E;E_c,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(E-E_c)^2}{2\sigma^2}\right). \tag{12}$$

Note that in the convolution integral, the Gaussian is centered at 0 and shifted by the integration variable E', which is why it appears as $G(E-E';0,\sigma)$.

The Gaussian convoluted Doniach–Sunjic model G∗DS(E) is based on the Doniach–Sunjic model [19]. and is particularly useful for fitting asymmetric peaks. Similarly to the Voigt function, the FWHM (F) is calculated from the width of the Gaussian (σ) and the width of the Lorentzian (γ), whereas the S term controls the asymmetry. The function can be expressed as follows:

$$G*DS(E;A,E_c,\sigma,\gamma,S)=A\cdot G(E;E_c,\sigma)*DS(E;E_C,\gamma,S), \tag{13}$$

where $DS(E;E_C,\gamma,S)$ is the Doniach–Sunjic model, G(E, σ) is the Gaussian model, A is the amplitude/area, E_c is the center position, and S is the asymmetry parameter or skew. While the Gaussian model has been already expressed in Equation (12), the Doniac–Sunjic model $DS(E;E_C,\gamma,S)$ model can be expressed as follows:

$$DS(E;E_C,\gamma,S)=\frac{\cos\left(\frac{\pi S}{2}+(1-S)\cdot\arctan\left(\frac{E-E_c}{\gamma}\right)\right)}{((E-E_c)^2+\gamma^2)^{(1-S)/2}}, \tag{14}$$

where for S = 0, DS(E;Ec, γ, 0) is of the form of a Lorentzian:

$$\begin{aligned}DS(E;Ec,\gamma,0) &= \frac{1}{((E-E_c)^2+\gamma^2)^{1/2}}\cdot\cos\left(\tan^{-1}\left(\frac{E-E_c}{\gamma}\right)\right)\\ &=\frac{1}{\sqrt{((E-E_c)^2+\gamma^2)}}\cdot\frac{1}{\sqrt{1+\frac{(E-E_c)^2}{\gamma^2}}}\\ &=\frac{\gamma}{(E-E_c)^2+\gamma^2}.\end{aligned} \tag{15}$$

Table 2 shows the *lg* ratio of the different models required to fit a peak of Voigt lineshape with a width of 1 eV and varying *lg* ratio. The LA lineshape used had no Gaussian contribution attributed to the model.

Fit Quality Indicators

Coefficient of determination (R_2) measures goodness of fit over a range of 0 to 1, where 1 indicates a good fit.

$$R^2=1-\frac{\sum_i(y_i-f_i)^2}{\sum_i(y_i-\bar{y})^2}, \tag{16}$$

TABLE 2 | Conversion table between the *lg* ratio values of the Voigt model and the GL, SGL, and LA models.

Voigt *lg* (%)	GL *lg* (%)	SGL *lg* (%)	LA σ &γ
10	30.27	10.23	9.58
20	49.04	20.29	4.42
30	66.14	31.47	2.74
40	80.44	43.54	1.98
50	90.46	56.44	1.57

Note: The simulated Voigt peak has an area A of 100, a height H of 95, a width of 1 eV, and a varying *lg* ratio.

where y_i are the observed values, f_i are the predicted values and $\bar{y}$ is the mean of observed values.

Relative standard deviation (RSD) indicates data scatter and goodness of fit:

$$RSD=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i-f_i}{\sqrt{y_i}}\right)^2}, \tag{17}$$

where N is the number of data points, y_i are the observed values, and f_i are the predicted values.

Reduced chi-square statistic (Red. Chi_2) for fit quality:

$$\chi^2_{red}=\frac{1}{N-p}\sum_{i=1}^{N}\frac{(y_i-f_i)^2}{\sigma_i^2}, \tag{18}$$

where N is the number of data points, p is the number of fitted parameters, y_i are the observed values, f_i are the predicted values, and σ_i are the uncertainties in the measurements. Lower values of χ^2_{red} generally indicate better fits, with $\chi^2_{red}\approx 1$ being ideal when using proper statistical weighting ($\sigma_i=\sqrt{y_i}$). However, the target value depends strongly on the weighting scheme employed. When using equal weighting ($\sigma_i=1$) or other weighting approaches commonly used in XPS analysis, the optimal χ^2_{red} values may deviate significantly from unity.

Corrected Area

The corrected area A_{corr} of a peak of area A fitted using one of the models previously described can be described as follows:

$$A_{corr}=\frac{A_{measured}}{\mathrm{RSF}\times\mathrm{TXFN}\times\mathrm{ECF}\times\mathrm{ACF}} \tag{19}$$

where RSF is the relative sensitivity factor for the element, ECF is the energy compensation factor, and TXFN is the transmission function of the instrument for a specific pass energy (PE). In some instances, the ECF, ACF, and/or TXFN are incorporated within the RSFs provided by the supplier. For example, in the case of AMRSFs, the ECF and ACF are set to 1 as they are already taken into account. In the case of the Thermo K-Alpha, the TXFN is set to 1.0 as it is incorporated into the

data (for excel file only). For all other files imported using Vamas (.vms), Kratos (.Kal), or Ulvac-Phi (.spe), KherveFitting reads the transmission values and corrects the intensity values accordingly within the Excel file. The ACF is the angle correction factor, which accounts for the angular distribution of photoelectrons.

The factor for the ACF is given by the following:

$$\mathrm{ACF} = 1 - \frac{\beta}{4}(3\cos^2\theta - 1), \tag{20}$$

where β is the asymmetry parameter for the orbital of interest and θ is the take-off angle of the photoelectrons relative to the sample surface. For instruments with a magic angle of 54. 7°, the correction factor ACF = 1.

The ECF depends on the selected library and accounts for the attenuation of photoelectrons due to inelastic scattering. Three methods are available to calculate the ECF: the kinetic energy approximation, the TPP-2 M formula for calculating the inelastic mean free path (IMFP), and Seah's universal equations for calculating the IMFP and effective attenuation length (EAL).

The simplest approach is the kinetic energy approximation, where the ECF is taken as the IMFP calculated using an approximate relationship between the IMFP (λ) and kinetic energy (E_k), based on the general shape of the universal curve:

$$\lambda \propto E_k^{0.6}. \tag{21}$$

A more rigorous approach to calculating the IMFP is the TPP-2 M formula (short for Tanuma, Penn, and Powell, who developed this relation [20].), which provides an accurate empirical formula for calculating the IMFP of electrons in various materials. This method is based on the kinetic energy of electrons and incorporates material properties such as density, molecular weight, and the number of valence electrons. Calculation of the IMFP is vital in surface-sensitive techniques like XPS, where understanding the escape depth of electrons is crucial for quantification. The TPP-2 M formula for IMFP (λ) is given by the following:

$$\lambda = \frac{E_k}{E_{\mathrm{pl}}^2} \cdot \left[\beta \ln\left(\gamma \cdot E_k\right) - \frac{C}{E_k} + \frac{D}{E_k^2}\right], \tag{22}$$

where

$$E_{\mathrm{pl}} = 28.8 \cdot \sqrt{\frac{N_v \cdot \rho}{M}}, \tag{23}$$

$$\beta = -0.10 + \frac{0.944}{\sqrt{E_{\mathrm{pl}}^2 + E_g^2}} + 0.069 \cdot \rho^{0.1}, \tag{24}$$

$$\gamma = 0.191 \cdot \rho^{-0.5}, \tag{25}$$

$$C = 1.97 - 0.91 \cdot \left(\frac{N_v \cdot \rho}{M}\right), \tag{26}$$

$$D = 53.4 - 20.8 \cdot \left(\frac{N_v \cdot \rho}{M}\right), \tag{27}$$

where E is the electron kinetic energy (eV), ρ is the material density (g/cm3), M is the molecular weight (g/mol), N_v is the number of valence electrons per atom, E_{pl} is free electron plasmon energy (eV), and E_g is the bandgap energy (eV). The resulting IMFP is reported in nanometers (nm).

For a general approximation applicable to metals and inorganic compounds, average matrix parameters can be used. Based on reference data, the following average values are assumed:

$$\begin{aligned} N_v &= 4.684 \quad \text{(valence electrons/atom)} \\ \rho &= 6.767\ \mathrm{g/cm^3} \\ M &= 137.51\ \mathrm{g/mol} \\ E_g &= 0\ \mathrm{eV}. \end{aligned}$$

Substituting these values into the equations, we get the following:

$$\begin{aligned} E_{\mathrm{pl}} &= 28.8 \cdot \sqrt{\frac{4.684 \cdot 6.767}{137.51}} \approx 16.26\ \mathrm{eV}, \\ U &= \frac{N_v \cdot \rho}{M} \approx 0.230. \end{aligned}$$

To enable comparison of the corrected area obtained using the TPP-2 M method with the corrected area using the Thermo Avantage software, a scaling factor of 26.2 was also applied to the calculated IMFP.

$$\mathrm{ECF} = \lambda \times 26.2. \tag{28}$$

An alternative to the TPP-2 M approach is Martin Seah's universal equations [21]. for the IMFP (λ) and effective attenuation length (EAL). These equations are mainly used for calculating the IMFP and EAL of thin film samples and can be expressed as follows:

$$\lambda = \frac{0.73 + 0.0095 \cdot E_k^{0.872}}{Z^{0.3}}\ (\mathrm{nm}), \tag{29}$$

$$\mathrm{EAL} = \frac{0.65 + 0.007 \cdot E_k^{0.93}}{Z^{0.38}}\ (\mathrm{nm}), \tag{30}$$

where E_k is the photoelectron kinetic energy (in eV) and Z is the average atomic number of the material.

The IMFP describes the average distance electrons can travel before undergoing inelastic scattering, while the EAL accounts for attenuation effects that reduce the signal's effective sampling depth. For thin-film quantification in XPS, the EAL model is particularly relevant as it adapts the analysis depth based on material and kinetic energy parameters. This ensures accurate quantification of elements in homogeneous thin films and avoids overestimation of the signal contribution from deeper layers. As it accounts for the effects of both elastic and inelastic scattering of electrons on the detected signal, the EAL is used for escape depth correction when Seah's universal equations are chosen as the method for calculating the ECF.

Examples of Application

To validate its reliability and accuracy, KherveFitting has undergone extensive testing and comparison with some of the most widely used XPS software, CasaXPS and Thermo Avantage. Tests against three different use cases have been included: quantifying relative atomic percentages of different elements in a room temperature ionic liquid, peak fitting the asymmetric peaks observed in the Ru 3*d* core level spectrum of RuO_2, and applying a complex, multicomponent fitting model to the Ce 3*d* core level spectrum of CeO_2 to determine the relative proportions Ce (III) and Ce (IV).

Quantitative Comparison

To rigorously assess the capability of KherveFitting to accurately reproduce the quantification of atomic percent fractions obtained with established commercial software, a comparison was performed using reference data from an ideal homogeneous sample—a room-temperature ionic liquid. Room-temperature ionic liquids are organic salts that are liquid below 100°C. These materials exhibit a high degree of homogeneity throughout their volume and have well-defined stoichiometric ratios of light elements, and because of their high surface tension and low vapor pressure, they are stable under ultrahigh vacuum. These properties make them ideal materials for XPS analysis. 1-propyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)imide is an ionic liquid consisting of five detectable elements in XPS (F, O, N, C, and S). An XPS spectrum of this ionic liquid has previously been reported [22]. with an accurate quantification exhibiting an average relative deviation of 3.4% of the expected stoichiometric ratio. These data (Accession #01708-01) were reanalyzed with KherveFitting v1.521 using data analysis parameters similar to those reported in the published paper (Figure 6). The survey spectrum was intensity corrected using a transmission function determined with the NPL method [23], and the inelastic contributions were removed using the Tougaard U4 background, $U(E; 367, 497, 436, 0)$, over the binding energy range of 800 to –7 eV. The peak areas were corrected using the average matrix RSFs (AMRSFs) published by NPL [24]. This was achieved in KherveFitting by a manual overwrite of the RSF values in the results table. The ECF term is already included in the AMRSFs, so this was set to unity in the results table. The equivalent homogeneous atomic composition of the ionic liquid determined by KherveFitting is presented in Table 3. The average relative deviation of the KherveFitting composition from the composition reported by Reed et al. is 2.9 % and 5.4 % from the theoretical stoichiometry of 1-propyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)imide. Note that the expanded uncertainties of the XPS compositions of this ionic liquid (considering both Types A and B errors) vary between 3 at.% and 7.5 at.% [25].

Asymmetric Lineshapes

Figure 7 shows the Ru 3*d*/C 1*s*(top figure) and O 1*s* (bottom figure) of a anhydrous Ruthenium dioxide material. The data were acquired using a Kratos Axis Ultra-DLD photoelectron spectrometer, using monochromatic Al K*α*X-rays (1486.69 eV) operating at 10-mA emission and 12-kV anode voltage (120-W operating power). The data were acquired using a pass energy of

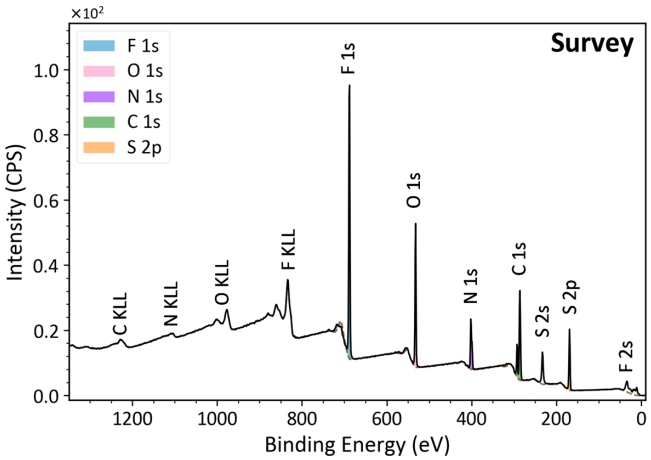

FIGURE 6 | Survey of a 1-propyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)imide obtained using data from Reed et al.

TABLE 3 | The average equivalent homogeneous atomic composition of 1-propyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)imide obtained using data from Reed *et al.* and analyzed using KherveFitting v1.521.

Element	Measured area (cps.eV)	AMRSF	At.%
F	267.65	3.26	23.7
O	141.11	2.46	16.5
N	66.83	1.66	11.6
C	135.32	1.00	39.0
S	57.11	1.80	9.2

40 eV. At this pass energy, the FWHM of the Ag $3d_{5/2}$ peak was 0.63 eV, and the valence band Fermi edge was determined to be 0.60 eV. Within the software, the Kratos F1*s* sensitivity factors were used, with no escape depth correction applied, as this is already included in the sensitivity factors.

Ruthenium dioxide (RuO_2) is a conducting oxide possessing an asymmetric core-level line shape, which results from screened states, and an energy window for the Ru 3*d* core-level which overlaps the C 1*s* region [26]. The peak model was constructed using a combination of LA asymmetric (for Ru 3*d* peaks) and GL lineshapes (for C 1*s* peaks) using constraints developed in CasaXPS, [27]. with the addition of small ($\pm$ 0.2 eV) movements in the binding energies. The C 1*s*/Ru 3*d* core-level spectra were fitted between 275.9 and 293.5 eV, while the O 1*s* core-level spectrum was fitted between 525.6 and 537.5 eV. Quantification using CasaXPS gave an O / Ru ratio of 1.95, while quantification in KherveFitting gave a O / Ru ratio of 2.2, which is in excellent agreement, given the uncertainties in XPS quantification, complexity of the peak model, and potential differences in the background positions and type (Shirley vs. Smart).

Complex, Multicomponent Fitting Model

CeO_2 XPS data were acquired using a Thermo NEXSA, using monochromated Al K*α*X-rays (1486.69 eV) at 6 mA emission

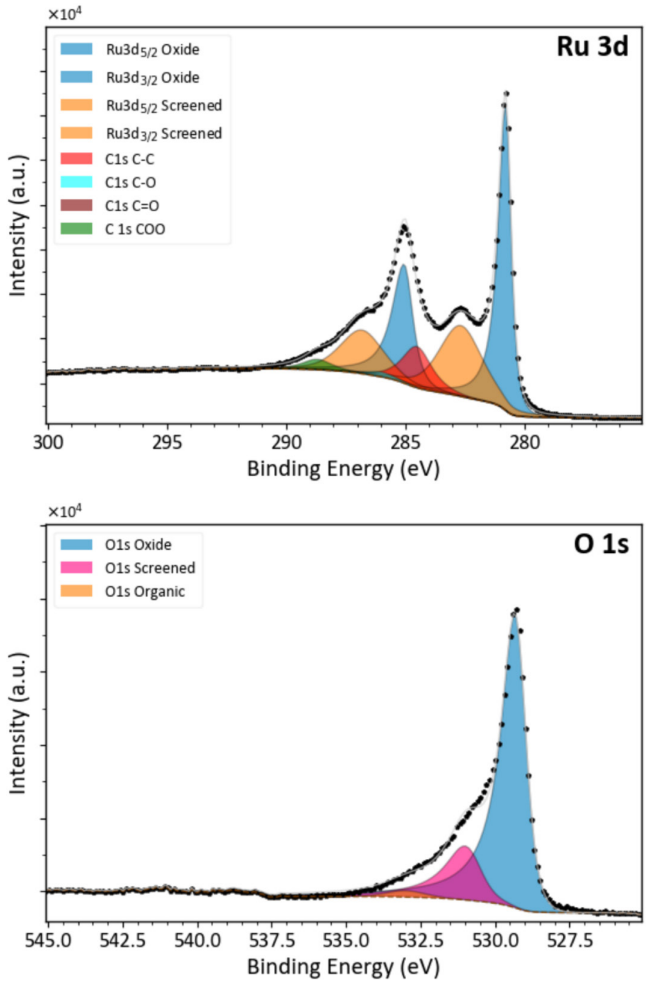

FIGURE 7 | Fitted Ru 3*d*/C 1*s* (top plot) and O 1*s* (bottom plot) core-levels for a reference anhydrous RuO_2 material.

and 12 kV HT (72 W), an elliptical spot size of 400 μm, and a 180° hemispherical analyzer in conjunction with a two-dimensional detector that integrates intensity across the entire angular distribution range. The instrument was calibrated to gold metal Au 4*f* (83.95 eV) and dispersion adjusted to give a BE of 932.6 eV for the Cu $2p_{3/2}$ line of metallic copper. Ag $3d_{5/2}$ line FWHM at 10 eV pass energy was 0.54 eV. The source resolution for monochromated Al K*α*X-rays is 0.3 eV, and the instrumental resolution was determined to be 0.36 eV at 10-eV pass energy using the Fermi edge of the valence band for metallic silver.

The Ce 3*d* spectral region presents particular challenges for peak fitting due to multiple final state effects and overlapping contributions from Ce(III) and Ce(IV) oxidation states [28, 29]. Additionally, experimental considerations during data acquisition can significantly influence the observed Ce(III)/Ce(IV) ratios, including X-ray induced reduction, vacuum exposure effects, and instrumental parameters such as X-ray flux density [30]. Users should be aware of these complexities when developing fitting models for cerium-containing materials and are encouraged to consult established literature on best practices for CeO_2 analysis to ensure robust and reproducible quantification.

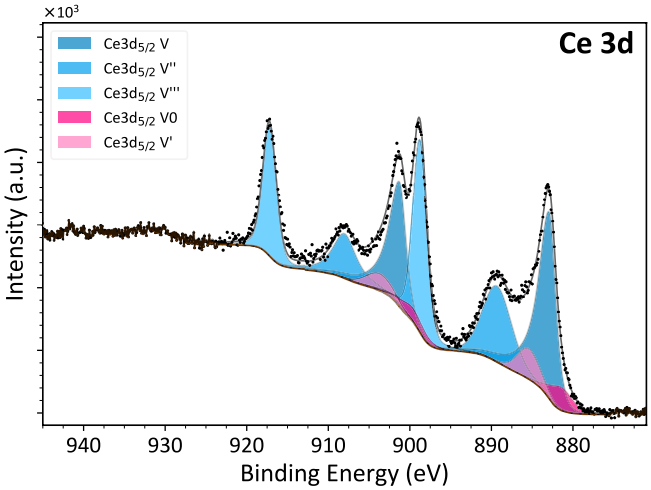

FIGURE 8 | Ce(III)/Ce(IV) fit of a CeO_2 reference material.

CeO_2 was fitted using a combination of a LA-type peak for the asymmetric Ce 3*d* v feature, and SGL lineshapes for the remaining nine peaks, according to a previously developed model for the same dataset (Figure 8) [30]. Three separate Shirley backgrounds were applied in the fitting range of the Ce 3*d*spectral region between 875.0 and 922.5 eV.

Conclusions

We have presented KherveFitting, a free and open-source XPS peak fitting software designed to provide an accessible, yet comprehensive solution for XPS data analysis. The software can import data in the majority of widely used XPS file formats, maintains compatibility with CasaXPS workflows, and features an intuitive sample manager for efficient multisample analysis. KherveFitting also offers a growing library of peak models, facilitating direct application of widely accepted peak models from the literature to new datasets, and incorporates a broad range of relative sensitivity factors and methods to calculate the energy compensation factors required for accurate quantification.

The software has been validated across diverse analytical scenarios, demonstrating reliable quantification and sophisticated peak fitting capabilities that are comparable to those of established commercial packages. With its rapid uptake from the XPS community and highly positive user feedback since its release, we anticipate that KherveFitting is on track to become a widely adopted software for XPS analysis. Our ultimate goal is to replicate the success of other widely used and respected open-source characterization software (e.g., Gwyddion [31]. for scanning probe microscopy, GSAS-II [32]. for X-ray diffraction, and ImageJ [33]. for scientific imaging), providing a similar service to the XPS community to make XPS more accessible and improve the quality of XPS peak fitting.

Future developments of the Khervefitting software will primarily focus on three key areas: extending its application towards synchrotron XPS measurements, incorporating additional, specialized peak models, and enhancing automation features based on community feedback and contributions.

Best Practices and Parameter Reporting

A significant challenge in the XPS community is incomplete reporting of peak fitting parameters, which hinders reproducibility of published results [34]. To address this, KherveFitting includes a Knowledge section in the help menu providing direct links to key literature and resources, including M. Biesinger's guidelines demonstrating best practice peak fitting examples [35]. and V. Crist's XPS library highlighting common pitfalls [2]. Users are strongly encouraged to report comprehensive fitting parameters including peak positions, FWHM values, peak shapes, constraints, and background parameters when publishing XPS results to ensure transparency and reproducibility.

Software Location

Archive: SourceForge
Persistent identifier: https://sourceforge.net/projects/khervefitting/
Licence: BSD-3 License
Publisher: Dr. Gwilherm Kerherve
Version published: 1.54
Code repository: GitHub
Persistent identifier: https://github.com/KherveFitting/KherveFitting
Licence: BSD-3 License
Publisher: Dr. Gwilherm Kerherve

Author Contributions

Gwilherm Kerherve: (Imperial College London) - project lead, software design, and development. **William Skinner:** (Imperial College London) - software development and testing. **Mark A. Isaacs:**(University College London / HarwellXPS) - software testing (Windows). **David J. Morgan:** (Cardiff University / HarwellXPS) - software testing (Windows). **Benjamen P. Reed:** (National Physical Laboratory) - software testing (Windows). **Arthur Graf:** (Cardiff University / HarwellXPS) - software testing (Windows). **Julian A. Hochhaus:** (TU Dortmund University) - software writing and testing (from source - Linux). **Hideki Nakajima:** (Synchrotron Light Research Institute / Thailand) - software testing (from source - macOS). **David J. Payne:** (Imperial College London) - conceptualization and proof reading.

Acknowledgments

We thank the members of the Surface Analysis Laboratory at Imperial College London for their valuable feedback during the beta testing phase and Dr Alex Shard and Dr David Cant from the National Physical Laboratory for their review of the manuscript.

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. T. Scientific, *Avantage Data System* (Thermo Fisher Scientific Inc., 2023).

2. B. V. Crist, "The "Spectral Data Processor" for Windows 3.1," *Journal of Surface Analysis.* 6, no. 1 (1999).

3. R. Hesse, T. Chassé, and R. Szargan, "Peak Shape Analysis of Core Level Photoelectron Spectra Using UNIFIT for WINDOWS," *Fresenius Journal of Analytical Chemistry.* 365, no. 1 (1999): 48–54.

4. R. W. M. Kwok. XPS Peak Fitting Program for WIN95/98 XPSPEAK, ver. 4.1. Software 2000. XPSPEAK4.1.

5. XPSOasis. AAnalyzer: A peak-fitting program for photoemission data. https://xpsoasis.orgAccessed: 2025.

6. N. Fairley, *CasaXPS: Processing Software for XPS* (AES: SIMS and More. Casa Software Ltd., 2009).

7. H. Nakajima, "LG4X," *Zenodo* (2024), https://doi.org/10.5281/zenodo.10477914.

8. J. A. Hochhaus and H. Nakajima, "LG4X-V2," *Zenodo* (2025), https://doi.org/10.5281/zenodo.15223359.

9. J. A. Hochhaus, "Lmfitxps," *Zenodo* (2025), https://doi.org/10.5281/zenodo.15130863.

10. R. Dunn wxPython: A GUI Toolkit for the Python Programming Language 2023. GitHub repository. Available from: https://github.com/wxWidgets/Phoenix.

11. J. D. Hunter, "Matplotlib: A 2D Graphics Environment," *Computing in Science & Engineering.* 9, no. 3 (2007): 90–95, https://doi.org/10.1109/MCSE.2007.55.

12. C. R. Harris, et al., "Array Programming With NumPy," *Nature* 585 (2020): 357–362, https://doi.org/10.1038/s41586-020-2649-2.

13. M. Newville, T. Stensitzki, D. B. Allen, and A. Ingargiola, "LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python," *Zenodo* (2014), https://doi.org/10.5281/zenodo.11813.

14. W. McKinney, "Data Structures for Statistical Computing in Python," in *Proceedings of the 9th Python in Science Conference* (2010): 51–56, https://doi.org/10.25080/Majora-92bf1922-00a.

15. D. A. Shirley, "High-Resolution X-Ray Photoemission Spectrum of the Valence Bands of Gold," *Physical Review B.* 5, no. 12 (1972): 4709–4714.

16. K. Levenberg, "A Method for the Solution of Certain non-Linear Problems in Least Squares," *Quarterly of Applied Mathematics.* 2, no. 2 (1944): 164–168.

17. D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial and Applied Mathematics.* 11, no. 2 (1963): 431–441.

18. S. Tougaard, "Practical Algorithm for Background Subtraction in x-Ray Photoelectron Spectroscopy," *Surface Science* 216, no. 1–2 (1989): 343–360 Available from: https://www.sciencedirect.com/science/article/pii/0039602889906118.

19. S. Doniach and M. Sunjic, "Many-electron Singularity in X-Ray Photoemission and X-Ray Line Spectra From Metals," *Journal of Physics C: Solid State Physics.* 3, no. 2 (1970): 285.

20. S. Tanuma, C. J. Powell, and D. R. Penn, "Calculations of Electron Inelastic Mean Free Paths (IMFPs) IV. Evaluation of Calculated IMFPs and of the Predictive IMFP Formula TPP-2 for Electron Energies Between 50 and 2000 eV," *Surface and Interface Analysis.* 20, no. 1 (1993): 77–89, https://doi.org/10.1002/sia.740200112.

21. M. P. Seah, "Simple Universal Curve for the Energy-Dependent electron Attenuation Length for all Materials," *Surface and Interface Analysis.* 44, no. 10 (2012): 1353–1359, https://doi.org/10.1002/sia.5033.

22. B. P. Reed, J. Radnik, and A. G. Shard, "Ionic Liquid $[PMIM]^+[NTf_2]^-$ (Solarpur);characterized by XPS," *Surface Science Spectra.* 29 (2022): 014001.

23. M. P. Seah, "A System for the Intensity Calibration of electron Spectrometers," *Journal of Electron Spectroscopy and Related Phenomena.* 71 (1995): 191–204.

24. M. P. Seah, I. Gilmore, and S. J. Spencer, "Quantitative XPS: I. Analysis of X-Ray Photoelectron Intensities From Elemental Data in a Digital Photoelectron Database," *Journal of Electron Spectroscopy and Related Phenomena.* 120 (2001): 93–111.

25. A. G. Shard, B. P. Reed, and D. J. H. Cant, "Surface Analysis Insight Note: Uncertainties in XPS Elemental Quantification," *Surface Science Spectra.* 57 (2025): 389–395.

26. D. J. Morgan, "Resolving Ruthenium: XPS Studies of Common Ruthenium Materials," *Surface and Interface Analysis.* 47 (2015): 1072–1079.

27. D. J. Morgan, "The Utility of Adventitious Carbon for Charge Correction: A Perspective From a Second Multiuser Facility," *Surface and Interface Analysis.* 57 (2025): 28–35.

28. E. Paparazzo, "Use and mis-Use of x-Ray Photoemission Spectroscopy Ce3d Spectra of Ce_2O_3 and CeO_2," *Journal of Physics: Condensed Matter.* 30 (2018): 343003.

29. E. Paparazzo, "On the Curve-Fitting of XPS Ce(3d) Spectra of Cerium Oxides," *Materials Research Bulletin.* 46 (2011): 323–326.

30. M. A. Isaacs, C. Drivas, R. Lee, R. Palgrave, C. M. Parlett, and D. J. Morgan, "XPS Surface Analysis of Ceria-Based Materials: Experimental Methods and Considerations," *Applied Surface Science Advances.* 18 (2023): 100469.

31. D. Necas and P. Klapetek, "Gwyddion: An Open-Source Software for SPM Data Analysis," *Central European Journal of Physics.* 10 (2012): 181–188, https://doi.org/10.2478/s11534-011-0096-2.

32. B. H. Toby and R. B. Von Dreele, "GSAS-II: The Genesis of a Modern Open-Source all Purpose Crystallography Software Package," *Journal of Applied Crystallography.* 46, no. 2 (2013): 544–549.

33. C. T. Rueden, J. Schindelin, M. C. Hiner, et al., "ImageJ2: ImageJ for the Next Generation of Scientific Image Data," *BMC Bioinformatics* 18, no. 1 (2017): 529.

34. G. H. Major, N. Fairley, P. M. A. Sherwood, et al., "Practical Guide for Curve Fitting in x-Ray Photoelectron Spectroscopy," *Journal of Vacuum Science & Technology a.* 38 (2020): 061203.

35. M. C. Biesinger, B. P. Payne, A. P. Grosvenor, L. W. M. Lau, A. R. Gerson, and R. S. C. Smart, "Resolving Surface Chemical States in XPS Analysis of First row Transition Metals, Oxides and Hydroxides: Cr, Mn, Fe, co and Ni," *Applied Surface Science* 257 (2011): 2717–2730.