PyXAS – an open-source package for 2D X-ray near-edge spectroscopy analysis

Mingyuan Ge and Wah-Keat Lee*

National Synchrotron Light Source II (NSLS-II), Upton, NY 11973, USA. *Correspondence e-mail: wklee@bnl.gov

In the synchrotron X-ray community, X-ray absorption near-edge spectroscopy (XANES) is a widely used technique to probe the local coordination environment and the oxidation states of specific elements within a sample. Although this technique is usually applied to bulk samples, the advent of new synchrotron sources has enabled spatially resolved versions of this technique (2D XANES). This development has been extremely powerful for the study of heterogeneous systems, which is the case for nearly all real applications. However, associated with the development of 2D XANES comes the challenge of analyzing very large volumes of data. As an example, a single 2D XANES measurement at a synchrotron can easily produce $\sim 10^6$ spatially resolved XANES spectra. Conventional manual analysis of an individual XANES spectrum is no longer feasible. Here, a software package is described that has been developed for high-throughput 2D XANES analysis. A detailed description of the software as well as example applications are provided.

1. Introduction

X-ray absorption near-edge spectroscopy (XANES) is a versatile approach that measures the local atomic structure of specific elements, which can provide valuable information about the oxidation states and local atomic coordination environment (Koningsberger & Prins, 1988; Rehr & Anku-dinov, 2005). A typical XANES measurement is conducted at a synchrotron light source by scanning the X-ray energy of the incident beam onto the sample across a specific element absorption edge. Depending on the setup, a variety of different signals can be used to construct the XANES spectrum, allowing the technique to be either surface sensitive (*e.g.* total electron yield, Auger electron yield) (Bianconi, 1980; Kitajima *et al.*, 1989; Isomura *et al.*, 2015; Frazer *et al.*, 2003) or bulk averaged (transmission mode, fluorescence yield) (Espinosa *et al.*, 2012; Yang *et al.*, 2005). Traditionally, a single XANES spectrum provides ensemble-averaged information of the sample, integrated over the area or volume that was exposed to the X-rays. Recently, by incorporating XANES into scanning X-ray microscopy and full-field transmission X-ray microscopy, spatially resolved XANES imaging (2D XANES) can be performed to analyze the sample heterogeneity by mapping out, for example, the valence states of the element of interest (Ade *et al.*, 1992; Fayard *et al.*, 2013; Liu *et al.*, 2011). Specifically, 2D XANES analysis enables spatially resolved maps of the valence state or local coordination fractions of the element of interest. In addition, the large penetration depth of X-rays enables samples to be measured under *in situ/in operando* conditions, enabling dynamic system studies that track morphology and valence-state evolution (Wang *et al.*, 2013). XANES has been used to gain a funda-

mental understanding of materials in many research fields such as condensed matter physics (Bianconi *et al.*, 1987; Lee *et al.*, 2004), energy storage (Lu *et al.*, 2017; Wang *et al.*, 2017; Yang *et al.*, 2019), catalysts (Andrews & Weckhuysen, 2013; Puga, 2018), *etc.*
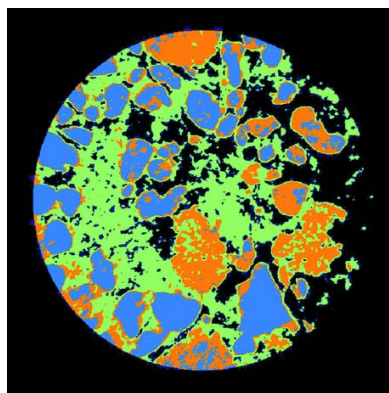
Although the interpretation of XANES spectrum analysis has been well documented and there are several packages with comprehensive analysis capabilities (Ravel & Newville, 2005), batch processing a large number of spectra for spatially resolved XANES imaging remains a challenge (Liu *et al.*, 2012). Generally, a 2D XANES measurement taken with a full-field synchrotron X-ray microscope consists of $>10^6$ spectra. A brute force approach directly looping over individual spectrums is not practical because of the processing time. The problem is further exacerbated for time-resolved (*in operando*) studies (Wang *et al.*, 2014) or 3D XANES tomography (Meirer *et al.*, 2011; Nelson *et al.*, 2011). For example, at the FXI beamline at National Synchrotron Light Source II (NSLS-II), a typical 2D XANES measurement with 100 energy points will take less than 10 minutes (Coburn *et al.*, 2019; Ge *et al.*, 2018). Coupled with the desire to perform *in operando* 2D or 3D XANES, it is not unusual for a user to generate $\sim 10^9$ spectra over 24 h. In order to analyze the data efficiently, it is necessary to streamline the image-analysis procedure so that one can retrieve the material information in a timely fashion.

Here, we describe a Python software package optimized for 2D XANES analysis with a graphic user interface (GUI) and demonstrate its functionalities through a few examples. Although this package can also be used for XANES analysis using fluorescence data, we will focus on data from a full-field transmission X-ray microscope. This package is deployed at the FXI beamline and is used for all analysis of XANES data collected at the beamline (Yang *et al.*, 2019; Zhang *et al.*, 2019).

2. Package structure for 2D XANES analysis

The general approach for XANES analysis is to fit the measured spectrum to a linear combination of reference spectra. The major functional components of the package are: (i) data loading, (ii) preprocessing, (iii) edge normalization, (iv) XANES fitting, (v) post-analysis and (vi) miscellaneous. The components are modular and their specific functions can be easily updated or replaced; for example, to input different file structures or to incorporate alternate optimization algorithms. Fig. 1(*a*) illustrates the layout of the analysis pipeline and Fig. 2 shows the GUI of the related modules. The GUI is written using *PyQT*. In the following, we will describe each section in detail.

3. Data loading

For 2D XANES analysis, we need a 2D XANES image stack and an energy list associated with the images. The current *PyXAS* package supports two input file formats: a simple TIFF image stack or a HDF5 (.h5) file. For the .h5 format, the attribute name of the data entry for the image stack is required. The associated energy list can be imported either from an ascii file (.txt) or from a .h5 file with the data-entry attribute specified. A manual command-line input of the energy list is also supported.
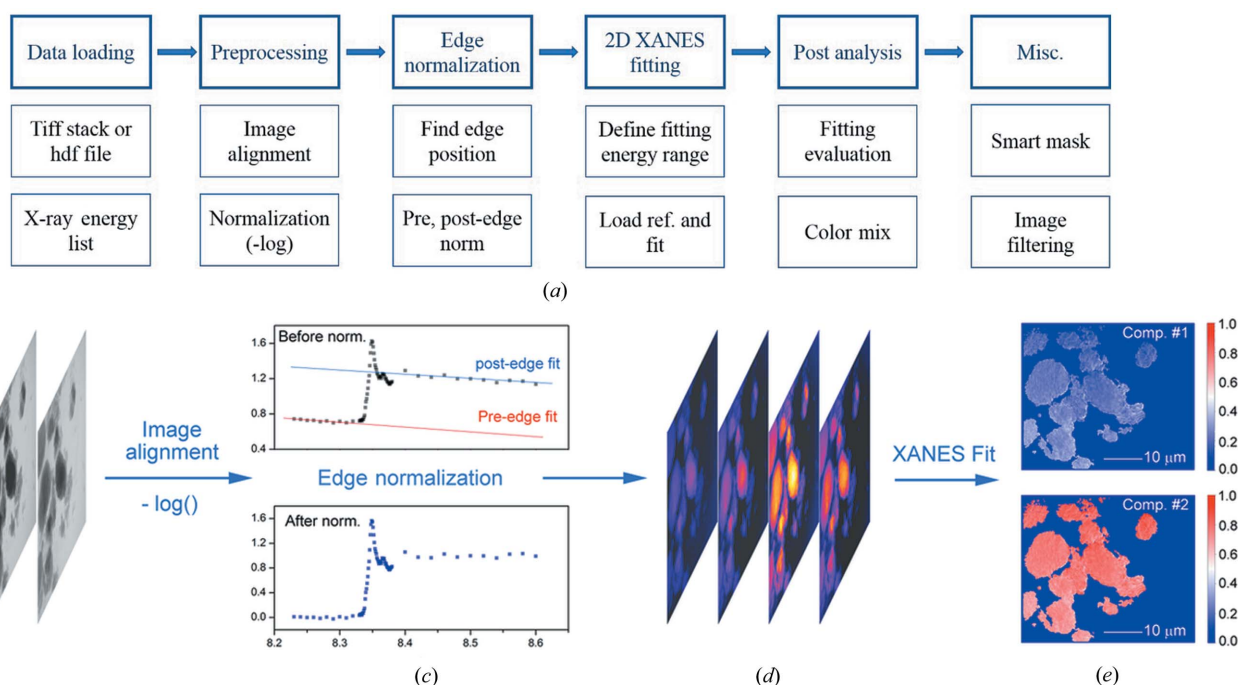

Figure 1
Overview of 2D XANES fitting. (*a*) Workflow of the XANES analysis package. (*b*) An example of raw TXM images data taken at different X-ray energies. (*c*) Pre- and post-edge normalization. (*d*) Edge-normalized image data. (*e*) Fitted results showing the 2D distribution of two components. Regions without particles have been masked out. The color bars on the right show the fractions.
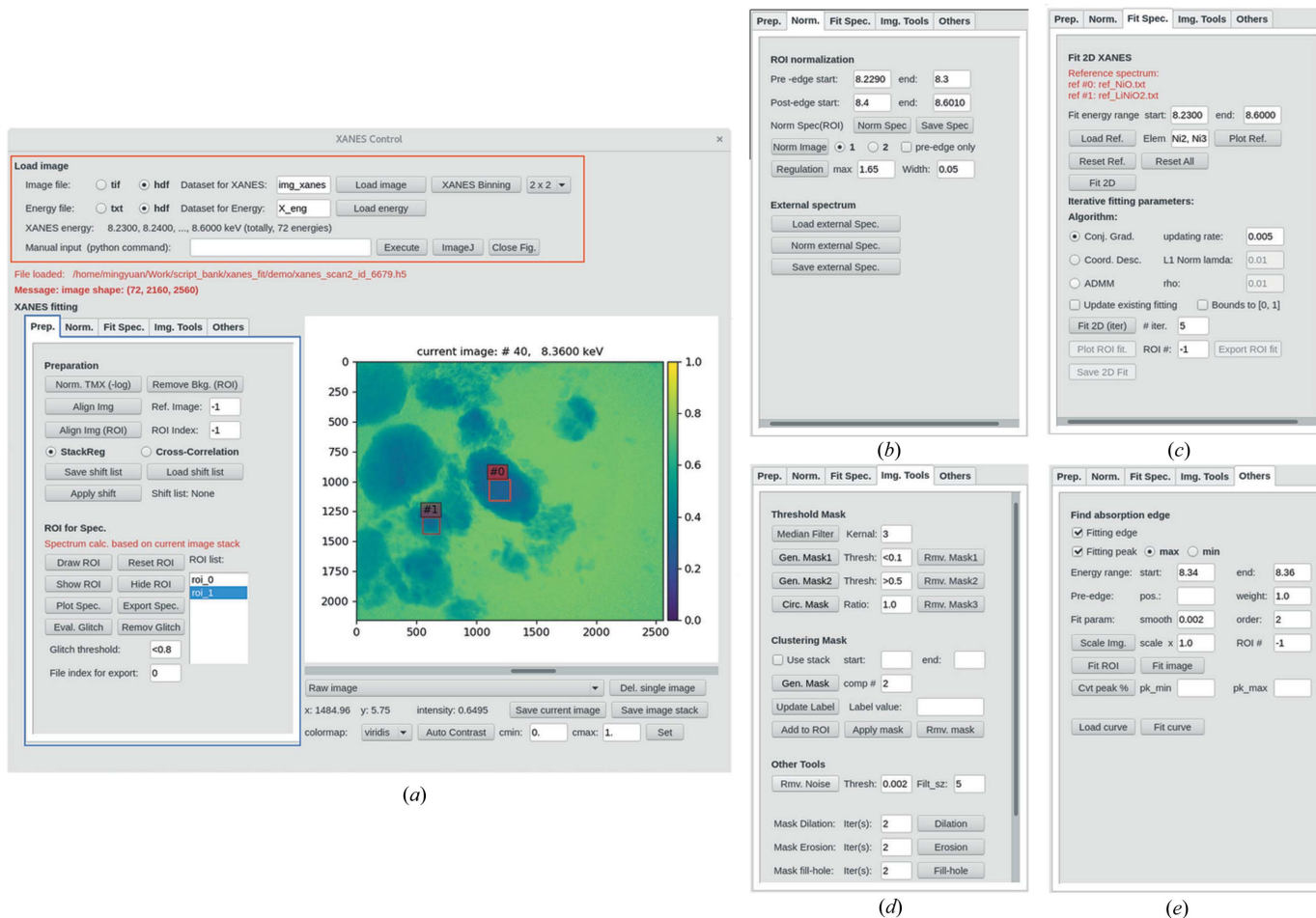
Figure 2
Screenshot of the *PyXAS* GUI. (*a*) Main screen. Data loading and preprocessing are indicated by the areas enclosed by red and blue lines. (*b*) Edge normalization, (*c*) XANES fitting, and (*d*) and (*e*) post-analysis.

4. Preprocessing

The first step for most X-ray image analysis is to perform what is normally called 'flat-field correction' where the 'dark' field is subtracted and the inhomogeneity of the incident beam ('white' or 'background' field) is removed by normalization. The program loads all these files and computes the flat-field corrected images. Once this is done, it is necessary to align the flat-field corrected images to correct for possible sample/ image drifts that occur during data acquisition. Such drifts can come from imperfect beamline optics' alignment and/or sample motions during the energy scans. For transmission X-ray microscopy (TXM) images, for example, this step corrects for small image shifts caused by the required motions of the optics during changes in the incident X-ray energy. The package has implemented two image-registration algorithms ['StackReg' (Thevenaz *et al.*, 1998) and 'Cross-correlation'] to align the XANES image stack without any prior knowledge about the physical motor positions used in the experiment. Specifically, 'StackReg' is capable of correcting a variety of image motions/distortions such as translation, rotation and scaling. 'Cross-correlation' can only correct for image translations in the current version. In general, 'StackReg' performs

better than 'Cross-correlation' if the image object has an abundance of feature details. In the package, one can specify any image from the image stack as a reference image [either using the full frame or a region of interest (ROI)] and align the rest of the images with respect to it. After the image stack has been aligned, the user can choose whether to take the negative natural log of the pixel value of the image. The negative log is required for transmission X-ray images (*e.g.* TXM-XANES) but not required for fluorescent XANES images. Explicitly, for TXM images,

$$I_{\text{TXM}}^{E} = \int_{z} \mu_{i}^{E}(z)\,\mathrm{d}z = -\ln\!\left(\frac{I^{E} - I_{\text{dark}}}{I_{\text{bkg}}^{E} - I_{\text{dark}}}\right), \qquad (1)$$

where I^{E} is the measured 2D intensity image of the sample with incident X-ray at energy E, I_{dark} is the measured 2D dark image without X-ray illumination, I_{bkg}^{E} is the measured 2D background image without the sample at X-ray energy E and I_{TXM}^{E} is the normalized (*i.e.* flat-field corrected) TXM intensity image. μ_{i}^{E} is the linear attenuation coefficient of different components i inside the sample at X-ray energy E and z is the thickness of the sample.

5. Edge normalization

For analysis, a single XANES spectrum from a material mixture is considered as a linear superposition of the spectrum from individual components inside the sample (*e.g.* a metal oxide with different oxidation states). Given the spectrum for each component (denoted as reference spectrum), one of the main tasks in XANES analysis is to calculate the percentage (ratio) of each component. Because the reference spectra are usually normalized to zero and one in the pre-edge and post-edge regions, respectively, similar normalization is required for the measured data to enable subsequent analysis. Furthermore, the pixel values in the flat-field corrected images as indicated in equation (1) reflect the integrated absorption coefficients over the sample thickness, which includes contributions from all material components. For optimal fitting to the reference spectrum of the element of interest, it is necessary to be able to remove the absorption contributions from the other elements. The edge-normalization procedure addresses both these aspects: removes the contribution from the other elements and normalizes the spectrum so that it can be fitted against references.

The edge normalization starts with linear fits to the pre-edge and post-edge regions [see Fig. 1(*c*)]. For this to work, it is assumed that both the attenuation contributions to pre-edge and post-edge regions are linear in energy. This is a reasonable assumption because the energy ranges involved here are in the tens to hundreds of eV range and assumed to be away from any other absorption edges. Pre-edge normalization is a method to remove the contribution from other elements and focus on the element of interest only. The difference between the post-edge and pre-edge fitted lines is then attributed to the 'thickness' of the element in question. Since the slopes of these two lines are generally not identical, a single thickness value is calculated by averaging over a range of energy points in the post-edge region (see Section 7.1).

The following equation describes the edge-normalization process,

$$I_{\text{edge-norm}}^{E} = \frac{I_{\text{TXM}}^{E} - I_{\text{TXM}}^{E_{\text{pre-edge}}}}{I_{\text{TXM}}^{E_{\text{post-edge}}} - I_{\text{TXM}}^{E_{\text{pre-edge}}}}, \qquad (2)$$

where $I_{\text{TXM}}^{E_{\text{pre-edge}}}$ and $I_{\text{TXM}}^{E_{\text{post-edge}}}$ are the calculated intensity image through a linear fit of I_{TXM}^{E} at the pre-edge and post-edge energy ranges and extrapolated to all energies, respectively. Thus, after the edge-normalization process, the value of each pixel in the generated image [$I_{\text{edge-norm}}^{E}$, Fig. 1(*d*)] represents a normalized linear attenuation coefficient for the element in question, which can now be fitted to reference spectra.

6. XANES fitting

After edge normalization, the edge-normalized XANES can be written as

$$[y_1, y_2, \ldots y_m]^T = A^{m \times k} \cdot [x_1, x_2, \ldots x_k], \qquad (3)$$

in which y_i is the edge-normalized spectrum ($I_{\text{edge-norm}}^{E}$) at a single pixel i, x_j is the jth reference spectrum with n energy points, m is the number of pixels and k is the number of reference spectrum. $A^{m \times k}$ is the coefficient matrix.

Explicitly, y_i and x_j can be written as

$$y_i = \left[I_{\text{edge-norm}}^{E_1}[i], I_{\text{edge-norm}}^{E_2}[i], \ldots, I_{\text{edge-norm}}^{E_n}[i] \right]^T$$

and

$$x_j = \left[r_j^{E_1}, r_j^{E_2}, \ldots, r_j^{E_n} \right]^T,$$

where $r_j^{E_n}$ is the intensity of the jth reference spectrum at energy E_n. The coefficient matrix A can be solved using direct linear least-square fitting or the alternating direction method of multipliers (ADMM) – an iterative linear algebra solving engine (Boyd *et al.*, 2011). The choice of the fitting method heavily depends on the data quality. Direct fitting is preferred when the XANES image is of high fidelity with good signal-to-noise ratio. ADMM is more robust for dealing with noisy images. In the package, both methods are optimized (vectorized) for processing large XANES images. Running on a single CPU (*e.g.* a desktop PC with an Intel i7-6800K processor) it takes only a few seconds to fit 100 energy points of 2k × 2k XANES images. The user can define the range of energies over which to fit the spectrum. The resulting values of the coefficient matrix A give the fraction of each reference material for the pixel.

As an illustration, Figs. 1(*b*)–1(*e*) summarize the XANES fitting procedure. Fig. 1(*b*) is the original XANES image taken from the transmission X-ray microscope. Fig. 1(*c*) demonstrates the edge normalization after the image was preprocessed (alignment and negative natural log taken), resulting in the normalized image as shown in Fig. 1(*d*). Fig. 1(*e*) shows the fitting results. In this case, two individual components are distinguished. And their fractions are encoded with the same color scheme. Note that only the region with particles is presented, and the empty region where the fitting gives poor results has been masked out (see Section 8.2).

7. Post-analysis

7.1. 'Thickness' image

The software generates two additional image files for visualization: 'image thickness' and 'fitting error'. The 'image thickness' is derived from absorption-edge normalization (see Section 5). For each pixel, once the pre-edge is normalized, the numerical value of the pixel (μ^E) in the post-edge energy range is largely related to the 'thickness' of the specific element that the X-ray associated with the pixel has traversed. This is defined as the average 'image thickness', which is approximately proportional to the amount of the specific element locally, and it can be written as

$$I_{\text{thickness}} \propto \bar{\mu}_i^E z = \frac{1}{n} \sum_{\text{post-edge energies}} \left(I_{\text{TXM}}^{E_{\text{post-edge}}} - I_{\text{TXM}}^{E_{\text{pre-edge}}} \right),$$

where n is the number of energy points in the post-edge region used for the average and $\bar{\mu}_i^E$ is the averaged attenuation coefficient in the range of post-edge energy for element i.

7.2. Fitting evaluation

Once the XANES fitting is done, it is important to evaluate the goodness of fit. One figure of merit is the fitting error at each pixel: $\mathrm{err} = \sum_i |y_i - \hat{y}_i|^2$, where i denotes the different energy points, y_i is the pixel value at energy i and $\hat{y}_i$ is the fit value. As an example, Fig. 3(a) shows an $\mathrm{LiNi}_x\mathrm{Co}_y\mathrm{Al}_z\mathrm{O}_2$ particle in an Li ion battery cathode material that has been partially charged. XANES imaging of Ni is obtained by recording the X-ray transmission images at different energies from 8.23 to 8.60 keV with a total of 72 energy points. Two Ni reference spectra are used to fit the results. Specifically, the reference spectrum 'ref1' is the Ni oxidation state derived from the same material at its pristine state (mostly Ni^{3+}) and 'ref2' is derived from the material at its fully charged state (mostly Ni^{4+}). Figs. 3(b) and 3(c) are the fitting results showing the spatial distribution of N^{3+} and Ni^{4+}, respectively. As expected, outside the particle, the fitting distribution is very noisy, with large fitting errors [Fig. 3(d)] because this region does not contain characteristic features of reference spectrums. The package provides a tool to explicitly check the fitting results at a specified ROI. As illustrated in Fig. 3(e), a good match between the fitting curve and the raw data in ROI #1 validates the high fitting accuracy, which is in sharp contrast to the poor fitting in ROI #2 as shown in Fig. 3(f). This is useful for masking out regions that are not relevant for subsequent analysis (see Section 8).

8. Miscellaneous

8.1. Peak finding

The package provides functions to find the peak positions in XANES spectra. For screening purposes, for example, peak positions can be used as rough indicators to show the composition differences (e.g. oxidation state) at different locations without performing a complete XANES fitting analysis. In the package, we can take either the raw images (I_{TXM}^E) or the edge-normalized images ($I_{\mathrm{edge-norm}}^E$) as the input. Given the range of energy we would like to fit, the package uses a third-order spline curve to fit individual spectrum at each pixel and display the peak-position map in a color scale.

8.2. Segmentation (mask)

From the above fitting routines (Section 6), oxidation-state fractions can be obtained for every individual image pixel. However, the resulting fraction map can have regions that are very noisy, for example, from regions where the relevant element is not present and/or within very small particles that have very large fitting errors (see Fig. 2). It is desirable to mask out (or segment) these regions so that subsequent analysis can be performed only on regions with good fits. Another goal in XANES imaging is to be able to identify different material compositions, for example, regions where there may be small differences in stoichiometry or regions where the oxidation state of an element may be distinctly different.

In the package, we provide two simple yet versatile segmentation methods for this purpose: threshold masking and clustering masking. In the threshold-masking method, we can use any image (raw or processed) inside the package to generate a mask by choosing a proper pixel value as a threshold. For example, we can choose a threshold according to the error map [Fig. 3(d)] to mask out the parts where the fitting error is greater than a certain value. We can also use the 'image thickness' to effectively generate a mask. It is easy to set a threshold to remove the regions with less mate-
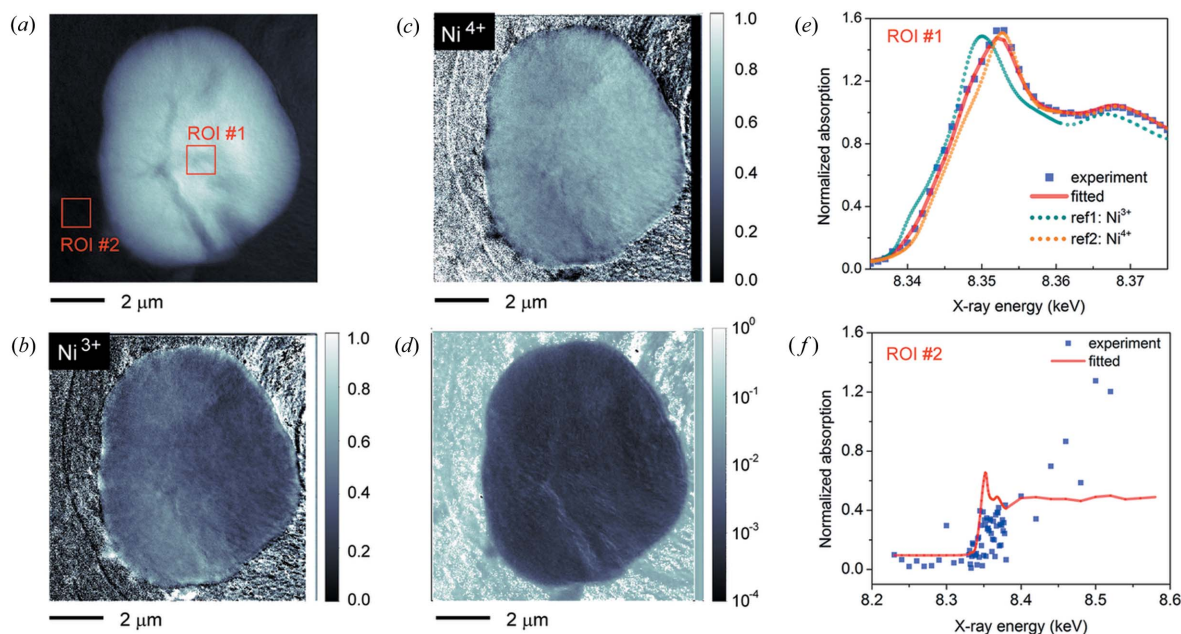

Figure 3
2D XANES fitting evaluation. (a) Flat-field corrected TXM image of a partially charged $\mathrm{LiNi}_x\mathrm{Co}_y\mathrm{Al}_z\mathrm{O}_2$ particle taken at 8.36 keV. (b) 2D distribution of Ni^{3+} after fitting. (c) 2D distribution of Ni^{4+} after fitting. (d) Fitting error on a log scale. (e) Comparison of fitting curve and raw data extracted from ROI #1 shown in (a). (f) Comparison of fitting curve and raw data extracted from ROI #2 shown in (a).
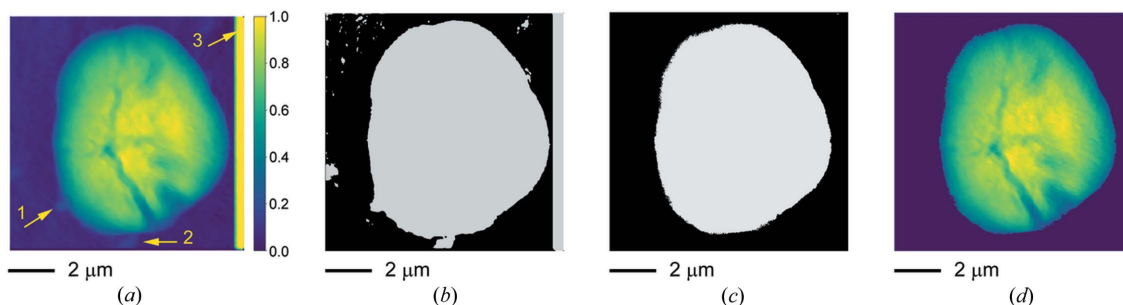
Figure 4
Example of threshold mask. (a) Thickness image of the LiNi$_x$Co$_y$Al$_z$O$_2$ particle. The arrows indicate the regions to be masked out. Specifically, arrow 1 and arrow 2 indicate the two small particles, and arrow 3 indicates the error part (a long stripe of yellow color) that comes from image alignment. (b) Mask generated from thickness image with a threshold set to 0.2. (c) Mask generated from 'error' image [Fig. 2(d)] with a threshold set to 1. (d) Thickness image after applying the masks in (b) and (c).

rial. The threshold masks generated from different images can be sequentially propagated and applied to all images available in the package. For example, Fig. 4(a) show the 'thickness' image. The arrows in Fig. 4(a) show the regions we would like to mask out. Fig. 4(b) shows the mask image that removes the pixels with values <0.2 in the 'thickness' image. Fig. 4(c) shows a second mask that removes pixels with values >1 in the 'Error' image [Fig. 3(d)], which gives the result shown in Fig. 4(d).

The second segmentation/masking tool included in the software is based on cluster analysis (O'Sullivan, 1995; Pedregosa et al., 2011). There are several different algorithms for cluster analysis, among them are: connectivity based, centroid based (k means), distribution based and density based. Two early uses of cluster analysis for X-ray applications have been for X-ray diffraction (Liao & Chen, 1992) and soft X-ray spectromicroscopy (Lerotic et al., 2005). More recently, for hard X-ray TXM, density-based cluster analysis has been used to identify chemically different compounds (Duan et al., 2016), and k-means-based cluster analysis has been used to look for minority phases (Zhang et al., 2017) and outliers (Mao et al., 2019). Currently, in our package, the cluster analysis is based on the k-means algorithm (Pedregosa et al., 2011). The references above and our examples below clearly demonstrate the great potential of cluster analysis for XANES imaging.

An application example of this algorithm (Fig. 5) is applied to a sintered mixture of Li$_7$La$_3$Zr$_2$O$_{12}$ (LLZO) and LiNi$_{0.8}$Mn$_{0.1}$Co$_{0.1}$O$_2$ (NMC), which holds the promise for the next generation of all-solid lithium

ion battery. In the XANES images, across the Ni-K edge, NMC undergoes significant contrast change, while LLZO only shows a gradual contrast change owing to mass attenuation. Here, 101 3D tomographic datasets were taken from 8.212 keV (below the Ni-K edge) to 8.777 keV (above the Ni-K edge). All the data were reconstructed and one cross-sectional slice of the reconstructed data at 8.32 keV is shown in Fig. 5(a), and the same slice at 8.35 keV is shown in Fig. 5(b).
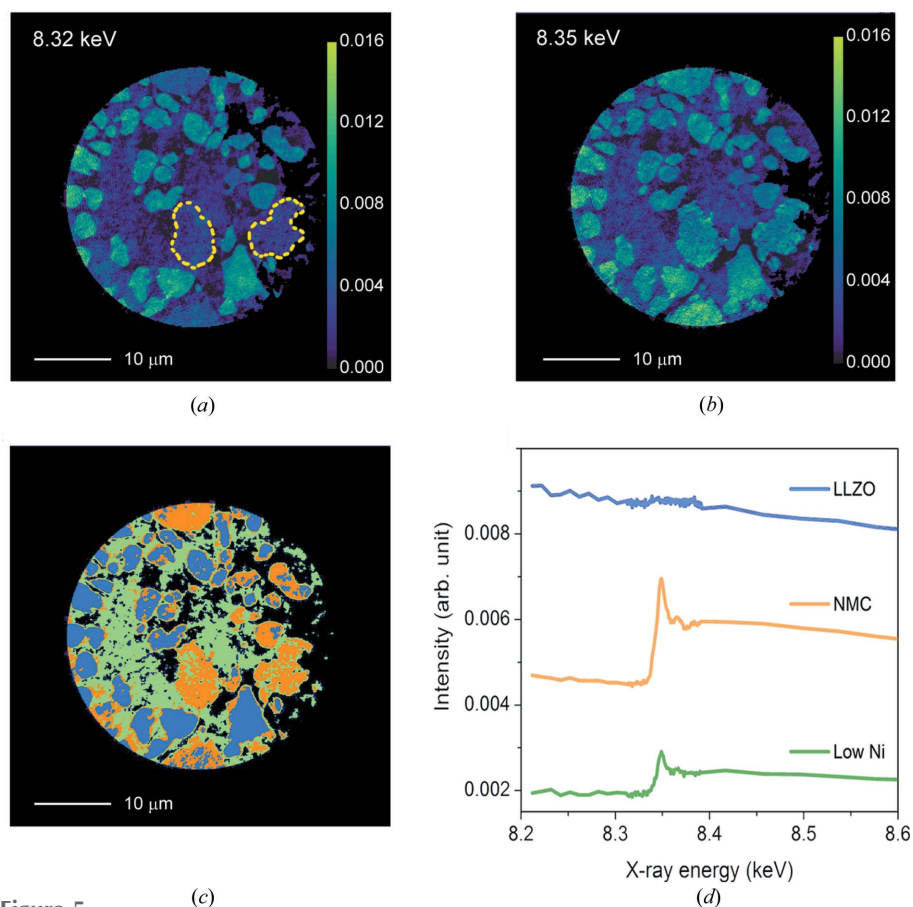

Figure 5
Example of clustering mask. (a) Image of the LLZO-NMC at 8.32 keV. (b) Image of the LLZO-NMC at 8.34 keV. (c) Separation of the LLZO-NMC material into three different material groups based on the characteristic features of XANES spectrum at individual pixels using the clustering-mask algorithm. (d) Spectra derived from the area of the three material groups with the same color scheme as shown in (c).

The dramatic changes in the absorption coefficient in the dashed regions in Fig. 5(a) compared with similar regions in Fig. 5(b) clearly indicates Ni-rich regions (NMC particles). However, the changes in other regions are more subtle, and whether there is a separate material component is not obvious. Cluster analysis can be a useful tool to identify such subtle differences. To demonstrate this, the entire 2D XANES reconstructed image-slice stack (*i.e* the same slice for all the 101 energies) was provided as input to the clustering algorithm. The algorithm requires the user to input the number of clusters desired; in this case, four (three material phases plus empty regions). Fig. 5(c) shows the results from the clustering analysis of the $540 \times 640 \times 100$ individual pixel spectrum. It shows that there are indeed three different material components, not two. To investigate the validity of the cluster analysis as shown in Fig. 5(c), averaged spectra can be calculated from each of the three different identified material components. Fig. 5(d) clearly shows that the three different groups have very different spectra. Note that because in this example the input to the cluster analysis is from 3D reconstructed data, material thickness has already been accounted for. The difference in spectrum between the low-Ni component and NMC is not because of thickness. We hypothesize that the low-Ni component is an inter-diffusion region of Ni (NMC) and LLZO. The spatial distribution of these different components revealed by the analysis will help to understand the sintering process at high temperature and provide guidance for researchers in the synthesis of the materials with desired properties and functionalities. Note that in the above example the cluster analysis was able to differentiate the three component materials without performing edge normalization (Section 5) or fitting (Section 6) on the data.

The above example demonstrates a way to differentiate materials using the clustering-analysis package. Clustering analysis can also help to distinguish different oxidation states in a material without having to do XANES fitting. This is especially useful in cases where reference spectra are not available. We demonstrate this application on a $LiNi_x$-$Co_yAl_zO_2$ sample (same material as in Figs. 3 and 4). As mentioned before, a full 2D XANES image stack (72 energy points, Ni-K edge) was collected and edge normalized as described in equation (2). The edge-normalized 2D image stack was provided as an input to the cluster algorithm and, as before, the desired number of phases was assigned to be four. Fig. 6(a) shows the output from the cluster analysis where the three identified components have been grouped based on the difference of XANES spectrum at individual pixel level. Fig. 6(b) compares the spectrum (8.345–8.360 keV) that is averaged from each of the three identified components. Although subtle, the differences in the spectrum for each component can be seen. We note that the cluster analysis did not require reference spectra.

As a way to investigate the validity of this cluster analysis, we performed a complete XANES fitting analysis on the data using the procedure described in Section 7. As a comparison, Fig. 6(c) shows the actual fitting results of Ni^{3+} (red) and Ni^{4+} (green) distribution. Taking the Ni^{4+} distribution image as the input and applying the clustering method with four components assigned, Fig. 6(d) shows the three-component separation derived from the fitting results using the same color scheme as used in Fig. 6(a). It is worth noting that applying the cluster mask or fundamentally the k-mean cluster algorithm to a single image is equivalent to grouping the pixels according to the pixel value, which in this case is the amount (fraction) of Ni^{4+}. The similarities between Figs. 6(a) and 6(d) demonstrate the ability of the cluster analysis to differentiate oxidation states for unknown material systems when reference spectra are not available. However, the difference between Figs. 6(a) and 6(d) indicates that regular XANES fitting prior to cluster analysis yields more accurate quantitative results. The clustering mask backed by the k-mean clustering algorithm is more sensitive to the noise, especially when the image signal is weak.

As a final remark about the cluster-analysis algorithm, the proper input into the algorithm is important. For example, in the case of Fig. 6, it is noted that correct edge normalization is necessary to achieve reasonable results. To show this, Figs. 6(e) and 6(f) are the output from the cluster analysis when the input images are not edge normalized as described in equation (2). Without proper edge normalization, the values in the input image matrix to the algorithm are highly correlated to the amount of material each ray traverses before it hits the pixel and, thus, the 'largest change' in the cluster parameter space is caused by thickness effect instead of features of the spectrum such as peak shape and position. That is why the resulting Fig. 6(e) looks like a 'thickness' map. In other words, without proper edge normalization, the sample thickness information dominates the signal and the cluster-analysis output reflects the depth integrated elemental concentration (or 'thickness'). The need to perform proper edge normalization here compared with the example in Fig. 5 is because of the fact that in Fig. 5 thickness has already been accounted for because the input images come from reconstructed 3D volumes.

9. 3D application

3D XANES analysis can be realized through a batch fitting of a 2D XANES image stack by calling the functions in the *PyXAS* package in a Python environment. In a typical 3D XANES experiment, 3D tomography at different energies of incident X-rays are collected, reconstructed and aligned. As an example, the material system we are demonstrating in Fig. 7 is $LiNi_5Mn_3Co_2O_2$ particles after being charged/discharged for 72 cycles. In the experiment, we acquired 56 3D tomography datasets across the Ni absorption edge and then performed image registration of the 2D projection collected at the same angle at different energies. Once the projections are aligned, the reconstructed 3D objects are fully aligned automatically. The next step was re-assembling the slice of the 3D object at the same height to form a 2D XANES image stack, which can be batched fitted using the functions provided in the *PyXAS* package. The results are shown in Fig. 7. Specifically, Fig. 7(a) represents the 3D distribution of Ni^{3+} concentration. Fig. 7(b)
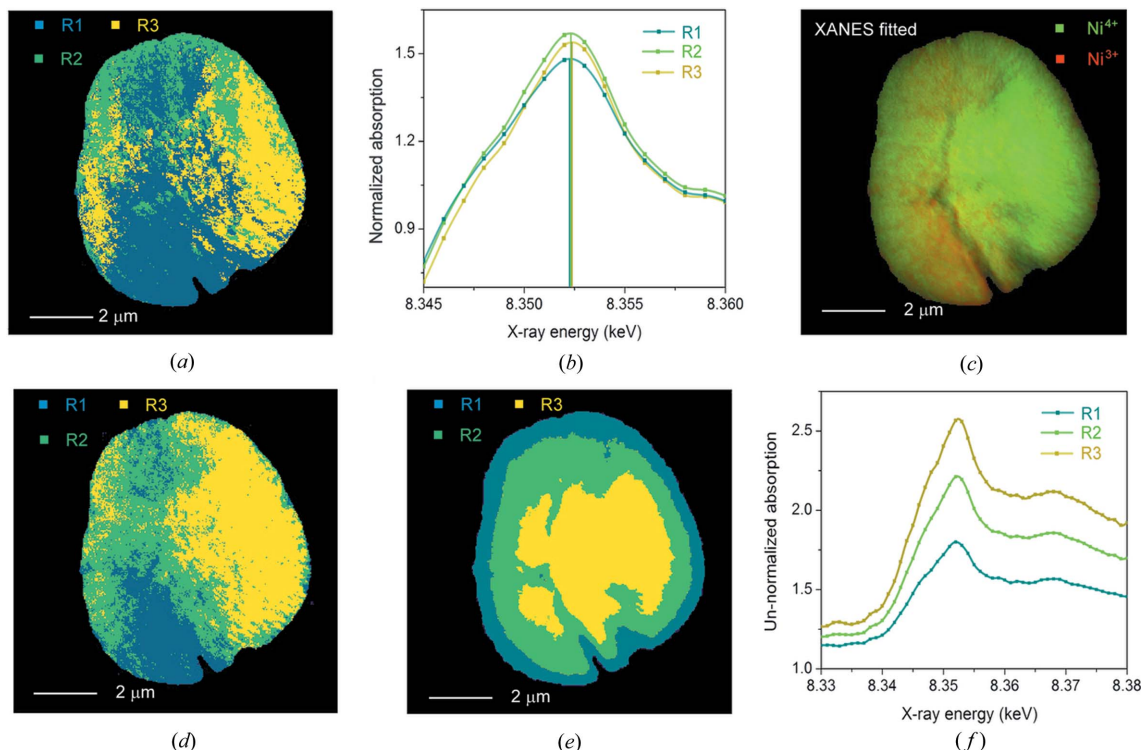
Figure 6
Comparison of clustering masks generated from different input datasets. (a) Clustering-mask generated map with an edge-normalized XANES image of the LiNi$_x$Co$_y$Al$_z$O$_2$ particle as the input. The averaged spectrum from R1, R2 and R3 are found to have Ni^{4+} concentrations of 0.542, 0.573 and 0.641, respectively, after we do the regular XANES fitting. (b) XANES spectra derived from the area of the three groups with the same color scheme as shown in (a). (c) Colormix of the fitting: red and green represent the distribution of Ni^{3+} and Ni^{4+}, respectively. (d) Clustering-mask generated map with the fitted Ni^{4+} fraction as the input (linear least-square fitting with reference spectra given). The averaged spectrum from R1, R2 and R3 are found to have Ni^{4+} concentrations of 0.472, 0.577 and 0.683, respectively. (e) Clustering-mask generated map from XANES image without edge normalization as the input. See main text for explanation. (f) XANES spectra derived from the area of the three groups with the same color scheme as shown in (e).

is a cross-section view of the results, where red represents Ni^{2+} and green represents Ni^{3+}. It is clear to see the enrichment of Ni^{2+} (red) around the crack area, which has arisen from the particle degradation after battery cycling.

10. Summary and future development

We introduced an efficient 2D XANES analysis package that is currently deployed at the NSLS-II FXI beamline. Generally, a regular XANES image (2k × 2k) with tens to hundreds of energy points collected at a synchrotron imaging beamline can be analyzed within a few minutes using a mid-level PC. The package is organized as a collection of modularized functions that can be easily modified or updated. A GUI has been implemented for 2D XANES analysis. The package is freely distributed with instructional documentation available on github: https://github.com/gmysage/pyxas. Currently, all the functions in the package can be called in the Python environment externally by importing the *PyXAS* package. This is especially useful for batch processing of 2D XANES image slices in order to build up a full 3D XANES model. In the future, we will incorporate the batch-process capability and other functions related to 3D XANES analysis into the GUI interface.

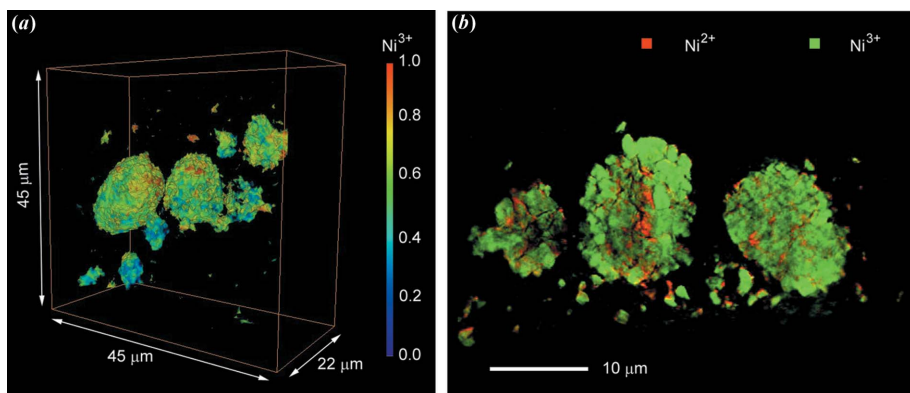**Figure 7**
3D XANES of Ni for material LiNi$_5$Mn$_3$Co$_2$O$_2$ after being charged/discharged for 72 cycles. (a) 3D volume rendering of Ni^{3+} concentration. (b) One slice of a cross-section view of Ni^{2+} (red) and Ni^{3+} (green) distribution.

Acknowledgements

thank Dr Hanfei Yan for the discussion about ADMM implementation.

References

Ade, H., Zhang, X., Cameron, S., Costello, C., Kirz, J. & Williams, S. (1992). *Science*, **258**, 972–975.

Andrews, J. C. & Weckhuysen, B. M. (2013). *Chem. Phys. Chem.* **14**, 3655–3666.

Bianconi, A. (1980). *Appl. Surf. Sci.* **6**, 392–418.

Bianconi, A., Castellano, A. C., De Santis, M., Rudolf, P., Lagarde, P., Flank, A. M. & Marcelli, A. (1987). *Solid State Commun.* **63**, 1009–1013.

Boyd, S., Parikh, N., Chu, E., Peleato, B. & Eckstein, J. (2011). *Mach. Learn.* **3**, 1–122.

Coburn, D. S., Nazaretski, E., Xu, W., Ge, M. Y., Longo, C., Xu, H., Gofron, K., Yin, Z., Chen, H. H., Hwu, Y. & Lee, W. K. (2019). *Rev. Sci. Instrum.* **90**, 053701.

Duan, X., Yang, F., Antono, E., Yang, W., Pianetta, P., Ermon, S., Mehta, A. & Liu, Y. (2016). *Sci. Rep.* **6**, 34406.

Espinosa, A., Serrano, A., Llavona, A., Jimenez de la Morena, J., Abuin, M., Figuerola, A., Pellegrino, T., Fernández, J. F., Garcia-Hernandez, M., Castro, G. R. & Garcia, M. A. (2012). *Meas. Sci. Technol.* **23**, 015602.

Fayard, B., Pouyet, E., Berruyer, G., Bugnazet, D., Cornu, C., Cotte, M., Andrade, V. D., Chiaro, F. D., Hignette, O., Kieffer, J., Martin, T., Papillon, E., Salomé, M. & Sole, V. A. (2013). *J. Phys. Conf. Ser.* **425**, 192001.

Frazer, B. H., Gilbert, B., Sonderegger, B. R. & De Stasio, G. (2003). *Surf. Sci.* **537**, 161–167.

Ge, M. Y., Coburn, D. S., Nazaretski, E., Xu, W. H., Gofron, K., Xu, H. J., Yin, Z. J. & Lee, W. K. (2018). *Appl. Phys. Lett.* **113**, 083109.

Isomura, N., Soejima, N., Iwasaki, S., Nomoto, T., Murai, T. & Kimoto, Y. (2015). *Appl. Surf. Sci.* **355**, 268–271.

Kitajima, Y., Yokoyama, T., Funabashi, M., Ohta, T. & Kuroda, H. (1989). *Physica B*, **158**, 668–669.

Koningsberger, D. C. & Prins, R. (1988). *X-ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS and XANES.* New York: John Wiley and Sons.

Lee, C. H., Matsuhata, H., Yamaguchi, H., Sekine, C., Kihou, K., Suzuki, T., Noro, T. & Shirotani, I. (2004). *Phys. Rev. B*, **70**, 153105.

Lerotic, M., Jacobsen, C., Gillow, J. B., Francis, A. J., Wirick, S., Vogt, S. & Maser, J. (2005). *J. Electron Spectrosc. Relat. Phenom.* **144–147**, 1137–1143.

Liao, B. & Chen, J. (1992). *J. Appl. Cryst.* **25**, 336–339.

Liu, Y., Andrews, J. C., Meirer, F., Mehta, A., Gil, S. C., Sciau, P., Mester, Z. & Pianetta, P. (2011). *AIP Conf. Proc.* **1365**, 357–360.

Liu, Y., Meirer, F., Williams, P. A., Wang, J., Andrews, J. C. & Pianetta, P. (2012). *J. Synchrotron Rad.* **19**, 281–287.

Lu, J., Wu, T. P. & Amine, K. (2017). *Nat. Energy*, **2**, 17011.

Mao, Y., Wang, X., Xia, S., Zhang, K., Wei, C., Bak, S., Shadike, Z., Liu, X., Yang, Y., Xu, R., Pianetta, P., Ermon, S., Stavitski, E., Zhao, K., Xu, Z., Lin, F., Yang, X. Q., Hu, E. & Liu, Y. (2019). *Adv. Funct. Mater.* **29**, 1900247.

Meirer, F., Cabana, J., Liu, Y., Mehta, A., Andrews, J. C. & Pianetta, P. (2011). *J. Synchrotron Rad.* **18**, 773–781.

Nelson, G. J., Harris, W. M., Izzo, J. R., Grew, K. N., Chiu, W. K. S., Chu, Y. S., Yi, J., Andrews, J. C., Liu, Y. J. & Pianetta, P. (2011). *Appl. Phys. Lett.* **98**, 173109.

O'Sullivan, F. (1995). *Ann. Statist.* **23**, 1267–1300.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011). *J. Mach. Learn. Res.* **12**, 2825–2830.

Puga, A. V. (2018). *Catal. Sci. Technol.* **8**, 5681–5707.

Ravel, B. & Newville, M. (2005). *J. Synchrotron Rad.* **12**, 537–541.

Rehr, J. J. & Ankudinov, A. L. (2005). *Coord. Chem. Rev.* **249**, 131–140.

Thevenaz, P., Ruttimann, U. E. & Unser, M. (1998). *IEEE Trans. Image Process.* **7**, 27–41.

Wang, J. J., Chen-Wiegart, Y. K. & Wang, J. (2013). *Chem. Commun.* **49**, 6480–6482.

Wang, J. J., Chen-Wiegart, Y. K. & Wang, J. (2014). *Nat. Commun.* **5**, 4570.

Wang, L. G., Wang, J. J., Zhang, X. Y., Ren, Y., Zuo, P. J., Yin, G. P. & Wang, J. (2017). *Nano Energy*, **34**, 215–223.

Yang, H. P., Wu, H. H., Ge, M. Y., Li, L. J., Yuan, Y. F., Yao, Q., Chen, J., Xia, L. F., Zheng, J. M., Chen, Z. Y., Duan, J., Kisslinger, K., Zeng, X. C., Lee, W. K., Zhang, Q. B. & Lu, J. (2019). *Adv. Funct. Mater.* **29**, 1808825.

Yang, L. C., McRae, R., Henary, M. M., Patel, R., Lai, B., Vogt, S. & Fahrni, C. J. (2005). *P. Natl Acad. Sci. USA*, **102**, 11179–11184.

Zhang, K., Ren, F., Wang, X., Hu, E., Xu, Y., Yang, X. Q., Li, H., Chen, L., Pianetta, P., Mehta, A., Yu, X. & Liu, Y. (2017). *Nano Lett.* **17**, 7782–7788.

Zhang, M. J., Hu, X. B., Li, M. F., Duan, Y. D., Yang, L. Y., Yin, C., Ge, M. Y., Xiao, X. H., Lee, W. K., Ko, J. Y. P., Amine, K., Chen, Z. H., Zhu, Y. M., Dooryhee, E., Bai, J. M., Pan, F. & Wang, F. (2019). *Adv. Energ. Mater.* **9**, 1901915.