Accurate EELS background subtraction – an adaptable method in MATLAB

Kayleigh L.Y. Fung[a,*], Michael W. Fay[b], Sean M. Collins[c], Demie M. Kepaptsoglou[d,e], Stephen T. Skowron[a], Quentin M. Ramasse[d,f,*], Andrei N. Khlobystov[a,b]

[a] *Department of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom*
[b] *Nanoscale and Microscale Research Centre, Cripps South, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom*
[c] *School of Chemical and Process Engineering and School of Chemistry, University of Leeds, Leeds LS2 9JT, United Kingdom*
[d] *SuperSTEM Laboratory, SciTech Daresbury Campus, Keckwick Lane, Daresbury WA4 4AD, United Kingdom*
[e] *Department of Physics, University of York, Heslington, York YO10 5DD, United Kingdom*
[f] *School of Chemical and Process Engineering and School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, United Kingdom*

1. Introduction

Incident electrons passing through a thin volume of matter in an electron microscope lose energy due to a variety of interactions that are dependent on the nature of the sample. The resulting energy distribution of electrons can be measured using electron energy-loss spectroscopy (EELS), and used to study structure, chemical composition or bonding. Specifically, EELS can for instance probe sample thickness using the integrated intensity of the zero-loss peak relative to the total integrated intensity of the spectrum [1], valence or conduction electron density using the plasmon peaks [2], elemental composition using core-loss edges, bonding and oxidation state using energy loss near edge structure (ELNES) [3], as well as band structure [4,5] and vibrational modes [6] using near zero-loss features. Accurate background fitting is essential to extract this information from raw EELS data. An excellent overview of EELS theory, instrumentation, and applications is presented in Ref. 1.

The majority of electrons passing through a suitable sample lose little to no energy, resulting in a high intensity zero-loss peak (ZLP). This peak is often in practice asymmetric, especially in the case of a cold field emitter, and often exhibits a tail extending significantly beyond the full-width at half-maximum (FWHM) of the ZLP which must be subtracted from the low-loss region (< 50 eV) before any quantification can be made. The background will also have contributions from any excitations of lower binding energy. In the core-loss region ($>$

50 eV), the major contributions to the large background come from several processes such as the plasmon tail, lower energy core-edges, multiple scattering, and instrument noise. Because of this, it can be difficult to extract quantitative elemental composition information from EEL spectra, even though EELS is particularly useful for quantification of light elements and for high spatial resolution [7]. In contrast, this information can be more readily obtained using the complementary technique of energy dispersive X-ray spectroscopy (EDXS) where background removal for non-overlapping transition metals and heavier elements causes fewer issues.

There is an approximately linear relationship between the logarithm of the energy loss and the logarithm of the intensity of the background of a core-loss EEL spectrum, where the largest contribution is from excitation of plasmons [8]. An ideal background can be fitted using a power law function in the form of $I(E) = AE^{-r}$ where I is the spectral intensity, E is the energy loss, and A and r are the coefficients. A description of the theoretical explanation for this behaviour can be found in Ref. 9. However, in practice, the spectral background deviates significantly from the ideal. Because of this, power laws may not accurately describe the backgrounds of low-loss ionisation edges below 50 eV or the majority of edges around 100 eV in medium-thick samples such as some K (Li, Be, B), L (Mg, Al, Si, P, S), M (Fe), and even N or O (Th, U) edges [10]. Power laws are more useful for background subtraction of non-overlapping edges at higher energy losses [11] and have limited application to ultralow-loss excitations close to the ZLP [6,12].

Recently there has been much progress in monochromators and aberration-correction for EELS which decreases the energy distribution of the ZLP and reduces the electron probe to the size of an atom. As a result, in state-of-the-art monochromated instruments the intense tail of the ZLP is now reduced to the point that it no longer obscures ultralow-loss features such as band gaps, optical or even phonon excitations. Typically, the FWHM of the ZLP is used as the energy-resolution indicator and an energy resolution of below 10 meV has been achieved [6,12]. Statistically good fitting is particularly important for background subtraction of the ultralow-loss region because the signal-to-background ratios are much smaller and the signals tend to be broader than core-loss EELS (> 50 eV). As a result, the peak position and edge integration can be greatly affected by the model and fitting window chosen for background subtraction [12]. A statistically worse fit can also lead to incorrect extrapolation of the background in the edge region. Background fitting can be further complicated when only small pre-edge regions can be used to model the background, such as when there are overlapping edges, because the width of the pre-edge window used can affect the fitting result [10]. It is important to process data with clarity and accuracy because curve fitting is often an essential, but not always appropriately used, tool across many areas of research (see the recent discussion in Ref. 13 on curve fitting in the analysis of X-ray photoelectron spectroscopic data).

The ubiquitous Gatan Microscopy Suite (GMS) 3 is the most commonly used software for EELS background fitting, offering users a straightforward non-linear least squares estimation (NLLSE) method to fit the background to a power law function. Many alternative methods have also been developed in the literature for quantification of EELS data such as the ImageJ plugin Cornell Spectrum Imager (CSI) developed as an open source software package for spectroscopic analysis [14,15], a stand-alone program for EELS quantification called EELSMODEL [16], quantitative multiple least-squares fitting [17], improved algorithms for fitting a power law that places constraints on the background above and below the ionisation threshold energy [18] which was subsequently expanded upon to provide a robust background subtraction method for noisy spectra containing multiple core-loss edges [19], high-throughput automated extraction of band gap maps using EELS mapping scripts written in MATLAB [20], as well as a signal-from-background separation algorithm written for GMS 3 that uses subspace division [21]. All of these methods have their merits and uses, and the consensus in the EELS community is that no single method is generally appropriate for every kind of EELS dataset.

In this paper we present a series of scripts written in MATLAB v. R2019b that can be applied to background fitting and subtraction of both the low-loss region as well as more generic uses of EELS, such as core-loss ionisation edges. We demonstrate several examples of their use for different applications of EELS and compare the quality of the results obtained to typical background subtraction in GMS 3. The primary focus of the scripts is to ensure accuracy in the background subtraction, allowing the user to manually assess the goodness-of-fit and estimate the effect of the background subtraction on the extracted data. The scripts are written in the MATLAB programming language – widely used in research and industry for data analysis and post-processing – and have a general format applicable to many types of EELS data. This allows users to easily modify the code if needed to suit their specific research needs. The three MATLAB scripts can be downloaded alongside the Supplementary Information which contains step-by-step instructions on usage (Section 3, pages 5–14). These instructions have been written to be user-friendly for researchers new to MATLAB. The MATLAB scripts are also available from GitHub at https://github.com/Fungussy/EELS-background-subtraction-in-MATLAB.

2. MATLAB fitting and fit analysis scripts

2.1. Functions used for fitting

The MATLAB scripts use the Curve Fitting Toolbox provided by MathWorks and have extensive commenting (text beginning with '%' or '%%') within the scripts to describe how the code works. There is also comprehensive MATLAB documentation by MathWorks as well as many sources of information on the Internet for researchers who wish to modify the scripts to their purposes. Alternatively, the scripts could also be ported over to an open source language such as Python since MATLAB has similarities with Python scripting.

There are three scripts described: EELS_fitting.m (fitting script), EELS_fit_analysis.m (fit analysis script), and EELS_subtracted_spectrum.m (plotting script). The first two scripts enable the researcher to find and analyse the best fit for the background while the third is to guide newer users of MATLAB in plotting their subtracted spectra. The fitting script gives a broad overview of a model fitted to a range of windows for background subtraction to estimate the best fit visually before using the fit analysis script to properly assess the goodness-of-fit. The x and y values of the subtracted spectrum can then be saved as separate MATLAB files and plotted using the plotting script. This third script also allows users to save the subtracted spectrum as a .txt for plotting in other software such as Origin.

The models available for fitting in the Curve Fitting Toolbox include one- and two-term power laws and exponentials. The models are as follows, with the MATLAB nomenclature in parentheses.

One-term power model ('power1′)

$$f(x) = ax^b$$

Two-term power model ('power2′)

$$f(x) = ax^b + c$$

One-term exponential model ('exp1′)

$$f(x) = ae^{bx}$$

Two-term exponential model ('exp2′)

$$f(x) = ae^{bx} + ce^{dx}$$

Exponential models can be more useful when the behaviour of the background deviates from a power law, such as in the low-loss region or when there are overlapping edges.

2.2. Statistical information on fitting

MATLAB uses an NLLSE method to iteratively fit the power law or exponential function to the data and can also give detailed information on the fit. For example, the output of running the script can contain the equation and coefficients of the fit as well as goodness-of-fit information such as the sum of squares due to error (SSE) and the adjusted R-square value. The SSE measures the total deviation of the fit values to the original data. The smaller the SSE value, the smaller the random error component in the model, and therefore the better the model is for prediction. For EELS background subtraction, there may be cases where only a small window can be used for fitting and therefore the SSE will be quite a large value and not necessarily useful for analysing the goodness-of-fit.

The adjusted R-square value, v, is the R-square value adjusted based on the residual degrees of freedom. This is the number of response values, n, minus the number of fitted coefficients, m, estimated from the response values.

$$v = n - m$$

The adjusted R-square value is a statistical measure of how successful the fit is in explaining the variation of the data. A value closer to 1 indicates that a greater proportion of variance is accounted for by the
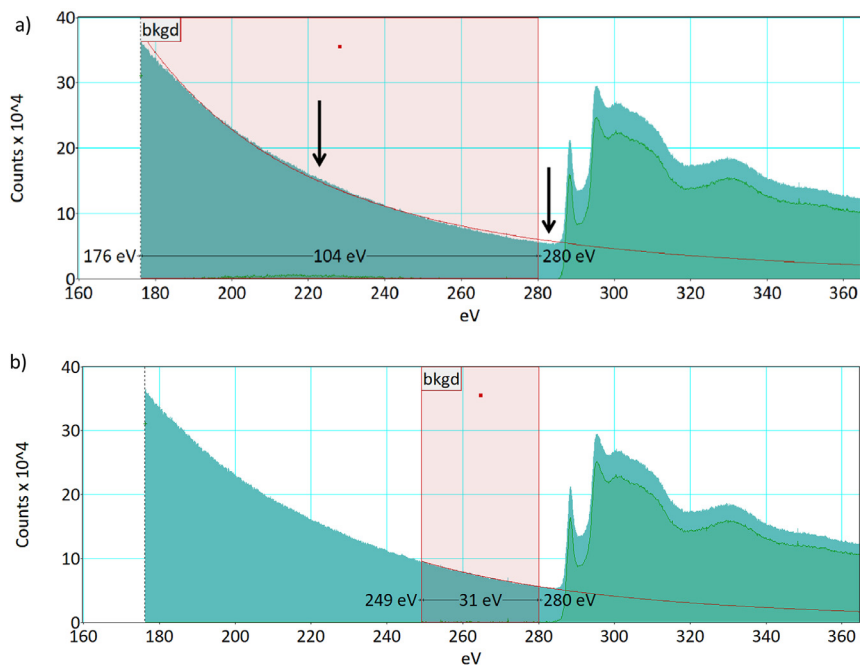
Fig. 1. Background fitting and subtraction for the carbon K-edge of single-walled carbon nanotubes using a one-term power law in Gatan Microscopy Suite (GMS 3) with a wide pre-edge window of 176–280 eV a) showing a processing artefact around 220 eV from the fitted curve dipping below the background signal (left arrow) as well as an over-estimation of the background in the edge region (right arrow). A narrow pre-edge window of 249–280 eV b) was also used, giving a more reasonable carbon K-edge signal. Original spectra in teal; fitting windows indicated with red rectangles; subtracted spectra overlaid in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

model, and therefore the closer to 1, the better the fit is for the data. It is worth noting that the R-square value provides little distinction between systematic errors in the fit and errors in the fit due to noise in the data.

The root mean squared error (RMSE) is an estimate of the standard deviation of the random component in the data. A smaller RMSE indicates a better fit for prediction.

In addition to the fit information, the fit analysis script also plots the prediction bounds and the residuals. The observation prediction bounds take into account the mean square error and therefore accounts for random errors in predicting the fitted curve above the edge threshold energy. The residuals are the difference between an observed value and the fitted value provided by a model – these should be randomly distributed if the fit is good. The user will be able to inspect the residuals for features such as artefacts and extra signals, which can affect background fitting.

Finally, the fit analysis script calculates the "h-parameter" and the signal-to-noise ratio (SNR) using the signal and background integrals (I_k and I_B respectively) which can be used to evaluate the background subtraction. Extrapolating the background fit under the signal can introduce errors into the extracted signal, leading to a lower SNR.

$$SNR = \frac{I_k}{\sqrt{I_k + hI_B}}$$

The h-parameter accounts for any statistical uncertainties in the background-dependant part of the signal integral that arise from fitting and extrapolation errors (i.e. the variance in I_B) [7,22].

$$h = \frac{I_B + \text{var}(I_B)}{I_B}$$

All the information about the fit should be considered when determining whether a fit for the background is statistically good or not. The MATLAB scripts we provide show this information which allows the user to objectively assess the goodness of their background fit and compare the suitability of different fits. This gives accurately subtracted EEL spectra with peak positions and integrals that can be used to describe the sample in a more quantitative way. Other statistical estimators of the goodness-of-fit are available in MATLAB, such as the chi-squared test, and can be implemented in addition to the examples demonstrated in the current set of scripts depending on the specific scientific questions the user may have.

3. Results and discussion

In this section we present four example cases using the scripts. In Section 3.1 we compare the method to the commonly used process available in GMS 3, using a simple example of background subtraction for the carbon K-edge of single-walled carbon nanotubes. The information that can be obtained in this case is comparable to GMS, but the presented MATLAB scripts offer the user increased transparency and flexibility. Section 3.2 demonstrates the increased accuracy enabled by an improved background fitting process, resulting in the elimination of artefacts in the subtracted spectrum of the carbon K-edge in diamond. Section 3.3 subsequently gives an example of when this kind of quantitatively accurate fitting is essential: independent background extraction for multiple overlapping peaks of different elements, in order to determine elemental composition of a sample. The spectrum recorded for this system – composed of carbon nanotubes inside boron nanotubes – also has a large degree of noise, demonstrating the resilience of our method. Section 3.4 considers low loss and vibrational/ultra-low loss EELS, for which accurate background subtraction is vital even for qualitative analysis; the choice of background and model can lead to substantially different peak positions, intensities, and integrations.

3.1. Comparison to GMS 3: background subtraction of the carbon K-edge of single-walled carbon nanotubes

The Gatan Microscopy Suite (GMS) 3 software contains a highly automated signal extraction method where the edge intensity is extracted by fitting a model to a single pre-edge region and extrapolating this fit to the entire background intensity below the edge signal. The most commonly used model for background subtraction in the core-loss region is a power law. It is recommended that the pre-edge fitting region precedes the edge (for example by 5 eV) and should be as wide as possible to reduce statistical error. However, at the same time, the fit region should only be around 10–30% of the edge energy. Comprehensive information on EEL spectral acquisition and post-processing using GMS 3 can be found at the website www.eels.info. The example used here for comparison between GMS 3 and MATLAB is a sample of single-walled carbon nanotubes (SWNTs) which has minimal multiple scattering since the SWNT bundles are, at most, only a few tens of nanometres thick. As can be seen in Fig. 1a, a wide pre-edge fitting

window of the carbon K-edge of SWNTs results in a worse background fit than a narrow pre-edge window using the default one-term power law model. There is a processing artefact around 220 eV from the fitted curve dipping below the background signal (left arrow in Fig. 1a) as well as an over-estimation of the background in the edge region (right arrow in Fig. 1a). Although a narrower pre-edge window gives a more reasonable carbon K-edge signal in Fig. 1b, the inclusion of fewer data points in the fitting process results in a statistically less accurate fit. The decreased suitability of fitting a one-term power law to the background with a larger range of data is an indication of the inadequacy of the power law in describing the entire background accurately.

While NLLSE methods are appropriate for EELS background fitting, especially for EEL spectra containing a small number of counts [23], GMS 3 fitting models are limited to one-term power laws and polynomials. No goodness-of-fit information is given which can be useful when quantifying a fit and determining whether a background fit is appropriate or not.

In comparison, fitting a two-term power law model to the same carbon K-edge background in MATLAB using EELS_fitting.m resulted in a much better fit across the entire pre-edge background. All data below several user-defined values of (i) eV were fitted, giving a range of background fits of windows from 176–200 eV to 176–280 eV. The background-subtracted spectra were then plotted in a rainbow of colours to aid finding the most appropriate window for background subtraction (Fig. 2a). In contrast to GMS 3, the widest window of 176–280 eV seemed to give the best fit – the subtracted spectrum had a flat baseline. Afterwards, this chosen background fit was analysed using EELS_fit_analysis.m (Fig. 2b) to show the equation used for fitting, the coefficient values, the goodness-of-fit, and the confidence bounds on the coefficients amongst other statistical information (see the output of this script in the Supplementary Information, Section 3, pages 5–14, for all this information as well as a guide on how to use the scripts). The prediction bounds (Fig. 2b; left plot) are particularly narrow in Fig. 2b which suggests low uncertainty in predicting the background under the signal. Most of the residuals of the fit (Fig. 2b; right plot) are distributed evenly which further suggests that this background fit is good. The script also outputs the h-parameter (which should typically be between 2 and 30 for low extrapolation errors) and the signal-to-noise ratio (SNR). Using signal and background integrals between 284 and 300 eV, we can see that the h-parameter is reasonable at a value of 25 which gives a good SNR of 507. The energy width of the integrals used to calculate the h-parameter and the SNR will vary depending on the elements in the sample and whether there are any overlapping edges. The background fitting window should be a similar width or wider than the energy width of the integrals in order to minimise extrapolation errors. The energy width should also be wide enough to have adequate signal but not too wide as this will introduce significant extrapolation errors [7,22].

After finding and analysing the best fit for removing the background, the subtracted spectrum can then be plotted separately using EELS_subtracted_spectrum.m (Fig. 2c). This demonstrates MATLAB's ability to produce post-processed spectra similar to GMS 3 but with additional statistical information on the fit. Understanding the goodness-of-fit allows one to extrapolate this fit across the entire background with higher accuracy.

It is straightforward to determine whether or not a fit is appropriate – an example of the results of deliberate poor fitting is shown in Fig. 3. In this case, a smaller fitting window of 176–200 eV was used for the carbon K-edge of the SWNTs instead of the larger window shown in Fig. 2. While the two-term power law fits that small window relatively well (adjusted R-square value of 0.9982), the prediction bounds of the fit are not as good and this is shown visually by the bounds curving away from the actual fit (Fig. 3a). The h-parameter and SNR are part of the output from MATLAB and can used to quantitatively assess the goodness-of-fit above the edge threshold energy. Using the same energy width as in Fig. 2b (between 284 and 300 eV), the h-parameter is much

larger at a value of 351 which gives a smaller SNR of 467. These two values also show that the power law fit is less suitable than the exponential fit for this spectrum. Fitting a smaller data set means that extrapolating the fit under the signal has a larger degree of uncertainty and leads to an under-estimation of the background in the edge region and therefore incorrect background subtraction (Fig. 3b).

Having a range of models available for fitting in the provided scripts, in contrast to GMS 3, allows the user to find the most applicable model for background subtraction that uses as many data points as possible. This gives a statistically more accurate fit with lower error which is useful for many situations, such as accurate quantification of elemental composition, extracting signals from noisy spectra, or background subtraction in the ultralow-loss region.

3.2. Fitting the background of the carbon K-edge of a polycrystalline diamond composite containing additional signals

The background of the carbon K-edge of a polycrystalline diamond (PCD) composite sample from Pacella et al. [24] was fitted using the same series of steps as above (Fig. 4). This is an example of EELS analysis of a sample approximately 100–150 nm in thickness, produced using focused ion beam (FIB) milling. The PCD composite sample is thicker than the above carbon nanotube sample prepared using a drop-casting technique and is therefore more likely to contain regions with significant multiple scattering. Several background windows below 280 eV were fitted using a one-term exponential function. Since all the windows gave similar subtracted spectra (Supplementary Information, Fig. 1, page 1), the largest window was chosen to include as much data in the fit as possible. The residuals are not evenly distributed (Fig. 4a; right plot); in the subtracted spectrum (Fig. 4c), there is a signal peaking at 170 eV which is in the region of the silicon L_1-edge, as well as a phosphor response artefact at 250 eV (a dip in intensity caused from having the ZLP focused on the spectrometer). Many EEL spectra may have backgrounds that contain additional signals or multiple scattering effects from thicker samples, and therefore non-evenly distributed residuals do not always indicate inappropriate fitting. These scripts are robust and not easily confused by irregularities in the background. Due to the thickness of the sample and the possible resulting tail of the Si L_1-edge, the h-parameter is comparatively large in this case (3447), with a small SNR of 83. Analysis of the absolute values of h and the SNR may be of limited relevance for such samples, however a relative comparison when varying fit parameters may still be instructive and help in determining the most suitable background fit. The published EEL spectrum of this PCD sample was fitted in GMS 3 using a one-term power law and there is a residual 'hump' at the start of the spectrum which can be replicated in MATLAB (Fig. 4b). The prediction bounds for the one-term power law (Fig. 4b) are also much wider than those for the one-term exponential (Fig. 4a; left plot), suggesting the exponential fit is more suitable for this spectrum. The power law fit in Fig. 4b has a large h-parameter of 2085 and a low SNR of 22. The exponential fit has a comparatively larger SNR of 83 which suggests the exponential fit is more suitable for signal extraction. Regardless of the fit, the SNR will be quite low because the original dataset is of a thick sample and thus has a significantly large background contribution. It should be noted that the fit of a one-term power law from Pacella et al. [24] did not affect their chemical analysis of this PCD sample. However, the presence of fitting artefacts gives unrealistic ratios of peak intensities, preventing accurate elemental quantification in samples containing multiple elements.

3.3. Fitting the background of a carbon nanotube inside boron nitride nanotube sample to extract the signals of the carbon, boron, and nitrogen K-edges

The MATLAB scripts can also be used to fit EELS data containing multiple elements. When choosing a fitting window for overlapping
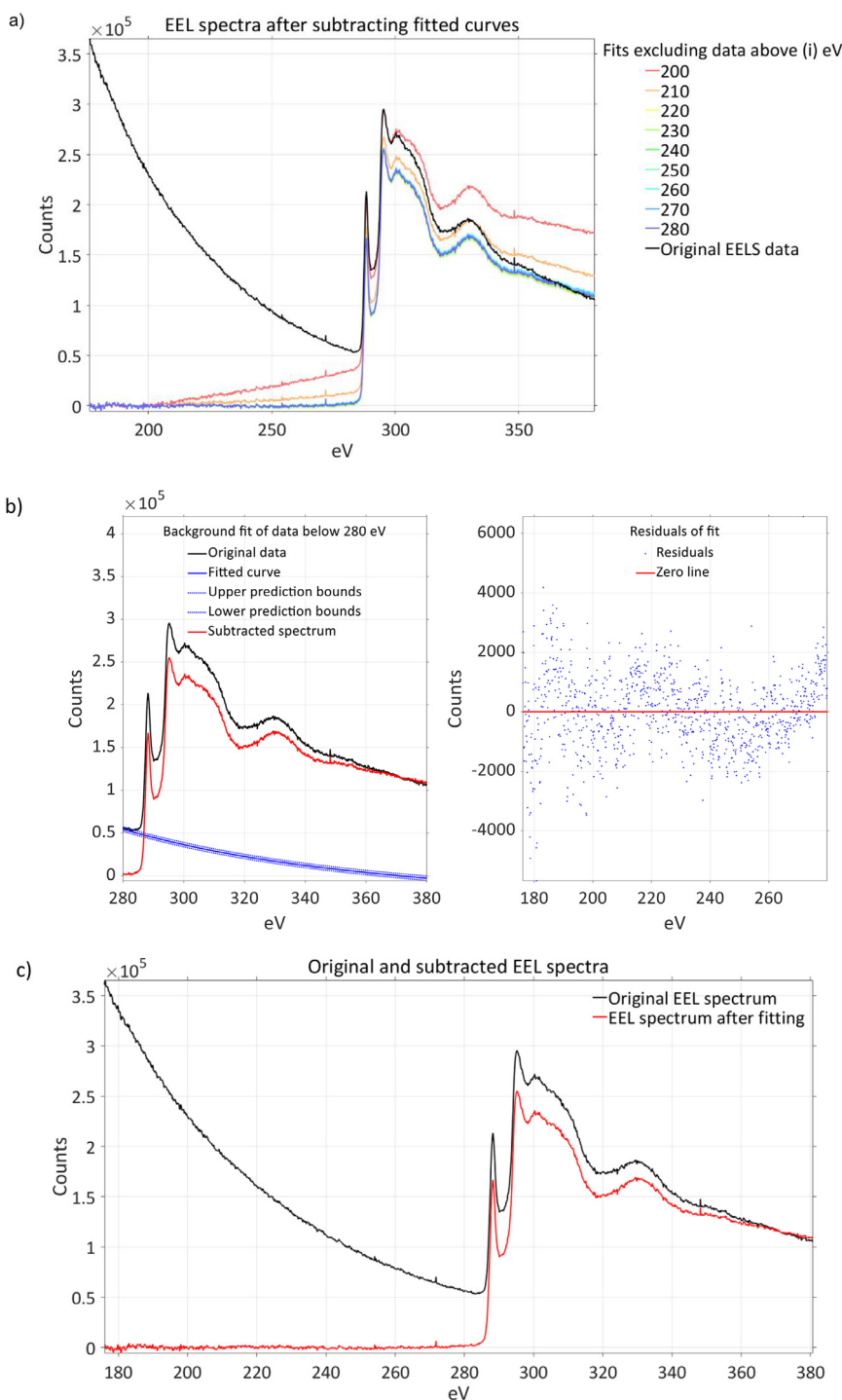
Fig. 2. a) The pre-edge background of the carbon K-edge of single-walled carbon nanotubes (SWNTs) was fitted using a two-term power law of the form $f(x) = ax^b + c$ in the EELS_fitting.m MATLAB script. Several fitting windows were chosen below (i) eV, from 176–200 eV to 176–280 eV, and plotted in a rainbow of colours to visually aid finding the best fit. b) The best fit using the widest window of 176–280 eV was analysed using EELS_fit_analysis.m which produced two plots as well as statistical information on the fit. This information can be found in the Supplementary Information (Section 3, pages 5-14) which contains a step-by-step guide on how to use these scripts. The guide uses the SWNTs EEL spectrum as an example to illustrate how carry out background subtraction and obtain the same data shown in this figure. The left plot shows that the model closely matches the background (the prediction bounds are shown as dotted lines). The right plot shows that the residuals after subtracting the fitted curves from the original EELS data are evenly distributed which suggests that the fit is good. The fit has an adjusted R-square value of 0.9998. The h-parameter and SNR (calculated using integrals between 284 and 300 eV) are 25 and 507 respectively. c) The original and background-subtracted spectra were plotted on the same axes for comparison using EELS_subtracted_spectrum.m.

edges, the widest possible window may not be the most appropriate as this may include weak signals from previous edges. The integral of the edges can be obtained from the EELS_fit_analysis.m script using the trapezoidal rule from which absolute quantification of the element without plural scattering can be calculated, provided the required parameters are known, using the equation below.
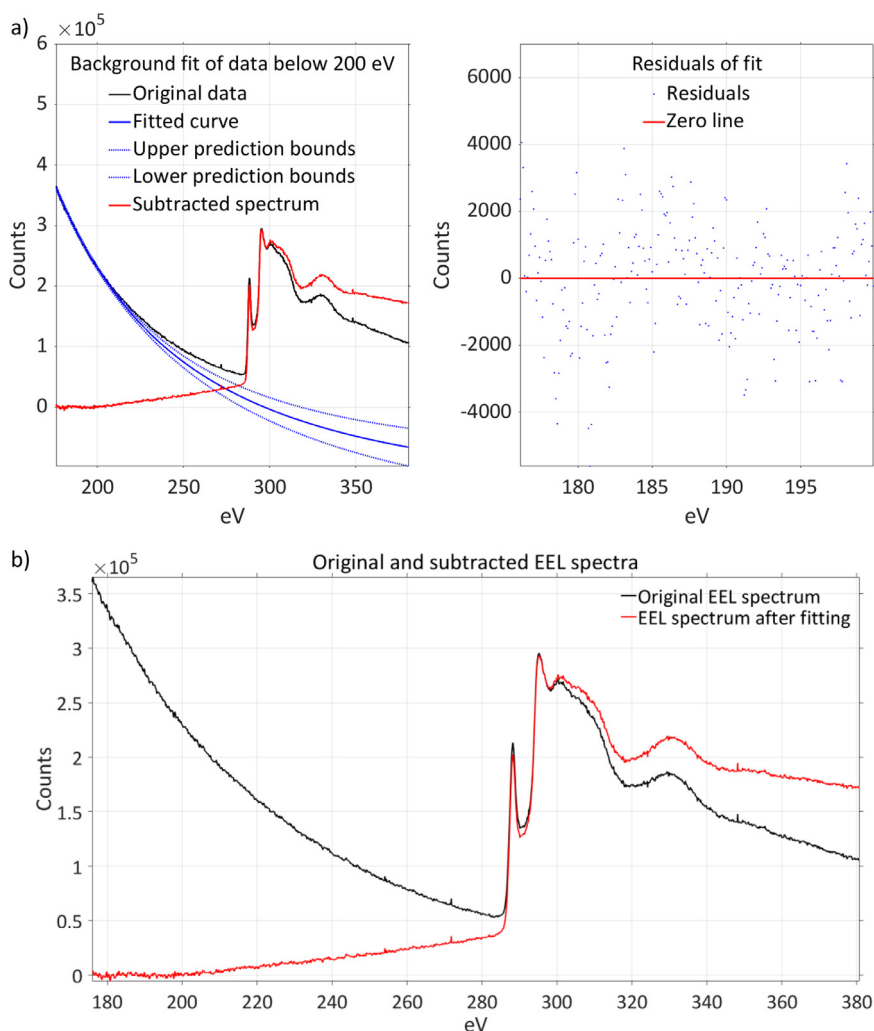
Fig. 3. a) A deliberate example of statistically bad fitting for the pre-edge background of the carbon K-edge of single-walled carbon nanotubes (SWNTs) to illustrate what a bad fit might look like. A two-term power law of the form $f(x) = ax^b + c$ was used, but with a smaller fitting window of 176–200 eV compared to the wider 176–280 eV window used previously in Fig. 2. The fit was analysed using EELS_fit_analysis.m which produced two plots as well as statistical information on the fit. The left plot shows that the model deviates significantly from the background after the window (the prediction bounds curve away from the fitted curve). The right plot shows that the residuals after subtracting the fitted curves from the original EELS data are evenly distributed which suggests that the fit is good for the small window but this is not the case for the entire background. The fit has an adjusted R-square value of 0.9982. The h-parameter and SNR (calculated using integrals between 284 and 300 eV) are 351 and 467 respectively. b) The original and background-subtracted spectra were plotted on the same axes for comparison using EELS_subtracted_spectrum.m. The baseline of the background-subtracted spectrum is not flat due to underestimation of the background in the edge region.

$$N = \frac{I_k^1(\beta, \Delta)}{I_o(\beta)\sigma_k(\beta, \Delta)}$$

N is the projected areal density of the atoms of element k; $I_k^1(\beta, \Delta)$ is the integral edge of element k with no plural scattering, integrated over a window Δ; I_o is the zero-loss integral; β is the effective collection angle; Δ is the signal integration width; and $\sigma_k(\beta, \Delta)$ is the partial inelastic scattering cross-section [8].

Obtaining accurate edge integrals relies on having a good fit for the background, and this is particularly pertinent for spectra where the signal-to-noise ratio is low. It is useful to note here that the contributions to the background may not stay identical under an ionisation edge, especially a few eV past the onset. Therefore, in terms of quantification, even if the best possible background fit is achieved using pre-edge fitting windows, there may still remain some unknown errors that could affect analysis.

The example we chose to illustrate this section was boron nitride nanotubes containing carbon nanotubes within, synthesised by Walker et al. (EELS data from Ref. 25). The deposited electron dose was minimised to reduce beam damage to the sample, so the EEL spectrum is noisy and has a low signal-to-background ratio. The boron edge can be fitted first (Supplementary Information, Fig. 2, page 2), then the carbon edge (Supplementary Information, Fig. 3, page 3), and then the nitrogen edge to give better elemental quantification (Supplementary Information, Fig. 4, page 4).

The models used for fitting varied between all three elements (see Supplementary Information Figs 2, 3 and 4, pages 2–4) and the adjusted R-square values get progressively smaller as the signal-to-noise ratio decreases. The prediction bounds also get wider as the data becomes noisier (Supplementary Information Figs 2b, 3b, and 4b). The h-parameters are all quite low, giving large SNR values (Supplemetary Information Figs 2, 3 and 4) which suggest that the models used for fitting are appropriate. After the backgrounds from all the edges have been subtracted, the edges can then be plotted against one another as well as the original EEL spectrum (Fig. 5).

3.4. Low loss and ultra-low loss peaks

The final example uses the MATLAB scripts to subtract the background of ultralow-loss spectra of glass blends of metal-organic
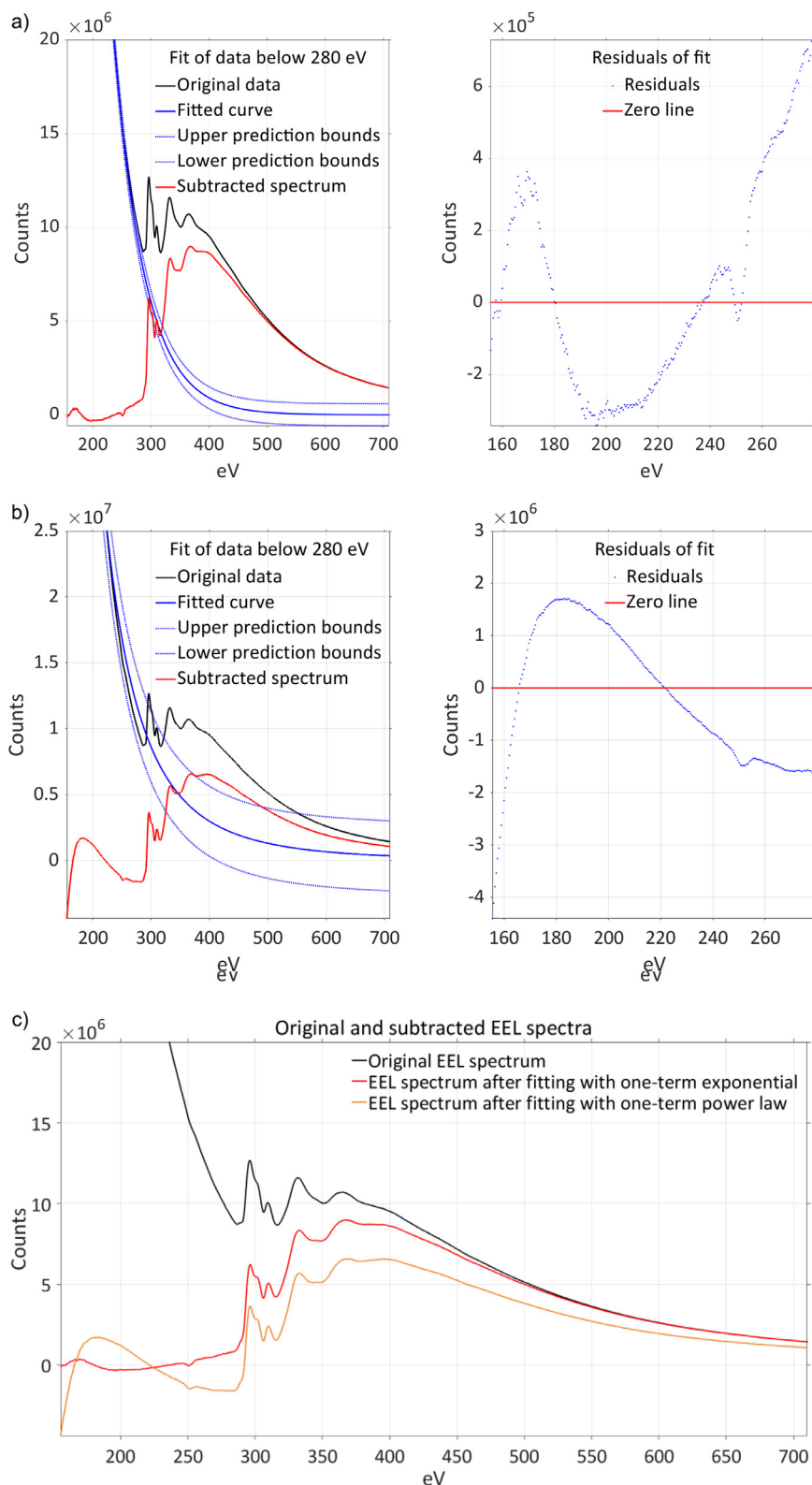
Fig. 4. a) The best fit using the widest window of 155.5–280 eV was analysed using EELS_fit_analysis.m which produced two plots as well as statistical information on the fit. The left plot shows that the model closely matches the background (the prediction bounds are shown as dotted lines). The right plot shows that the residuals after subtracting the fitted curves from the original EELS data are not evenly distributed due to a possible silicon L_1-edge at around 170 eV as well as a phosphor response at 250 eV. The fit has an adjusted R-square value of 0.9998. The h-parameter and SNR (calculated using integrals between 284 and 300 eV) are 3447 and 83 respectively. b) An unsuitable model of a one-term power law of the form $f(x) = ax^b$ was used instead with the same fitting window of 155.5–280 eV to demonstrate that the power law does not fit the background well. A one-term power law was used to fit the background by Pacella et al. [24]. The fit was analysed using EELS_fit_analysis.m which produced two plots as well as statistical information on the fit. The left plot shows that the fit both over- and under-estimates the background. Note that the prediction bounds are much wider than for the fit using a one-term exponential. The right plot shows that the residuals after subtracting the fitted curves from the original EELS data are not evenly distributed. The adjusted R-square value is 0.9966; however, this may not be meaningful. The h-parameter and SNR (calculated using integrals between 284 and 300 eV) are 2085 and 22 respectively. c) The original and both background-subtracted spectra were plotted on the same axes for comparison using EELS_subtracted_spectrum.m. The baseline of the EEL spectrum after fitting with a one-term power law contains a residual 'hump' peaking around 180 eV, as seen in the original paper.

frameworks (MOFs) acquired using scanning transmission electron microscopy (STEM) EELS [26,27]. Electron beam monochromators were used to achieve an incredibly low energy spread of several tens to hundreds of meV (two different microscopes were used in these publications). The electron beam can lose energy to delocalised electronic excitations and cause low energy excitations less than an eV, which is in the infrared region where vibrational modes occur. An atom-sized probe allows the study of vibrational modes of many types of nanostructures with high spatial resolution [6]. The dominant feature in the experimental data is the ZLP which must be removed in order to analyse the vibrational peaks. In particular, the ZLP from a cold field emitter exhibits strong asymmetry due to the physics of the emission process and once the electron beam has passed through the sample, there can be further asymmetry between the gain and loss sides due to
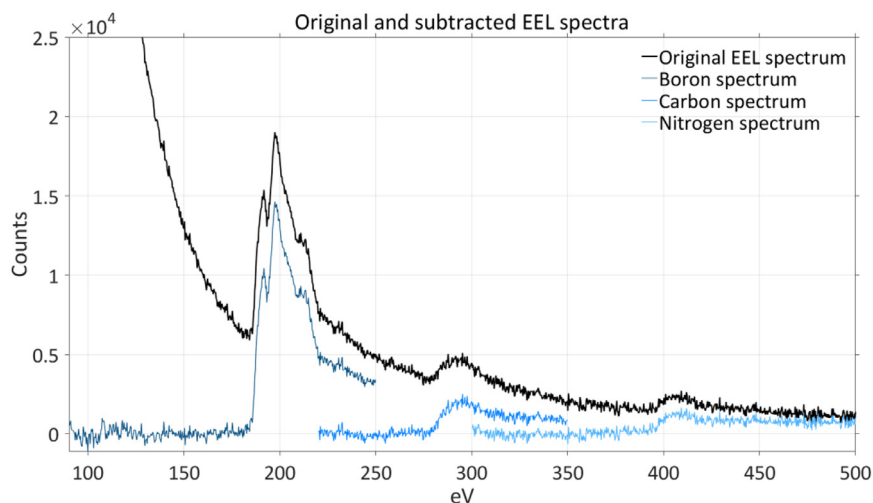
Fig. 5. The subtracted spectra of the boron, carbon, and nitrogen K-edges of carbon nanotube in boron nitride nanotube (CNT@BNNT) (EELS data from Ref. 25) was plotted with the original EEL spectrum on the same axes for comparison using EELS_subtracted_spectrum.m.

electron-sample interactions.

All of the example spectra are of element-rich regions of particles of glass blends or composites. Fig. 6 shows the low-loss plasmon spectra of particles with Co- and Zn-rich regions [26]. All new red spectra in Fig. 6 are similar to the published blue spectra and have the added benefit of having statistical information on the goodness-of-fit. Due to the data being in the low-loss region, the background contribution from the ZLP is significant and the SNR values are correspondingly small compared to the core-loss example of SWNTs above. At this point, it is worth discussing the camera technology which can affect how background

fitting is carried out. The digital signals given by the CCD camera have an offset and gain applied by the pre-amplifier. The offset is a value added or subtracted to the analogue signal while the gain is the analogue signal multiplied by a value to boost the signal. Generally, these values are corrected by calibration so when there is no analogue signal, the digital signal is at zero. However, calibrations can drift so incorrect offset values can give negative values, which can therefore affect the fitting. The offset and gain can either be assumed to be correct or a correction can be applied to the fit. The published spectra took into account the camera offset and incorporated that into their fitting, which
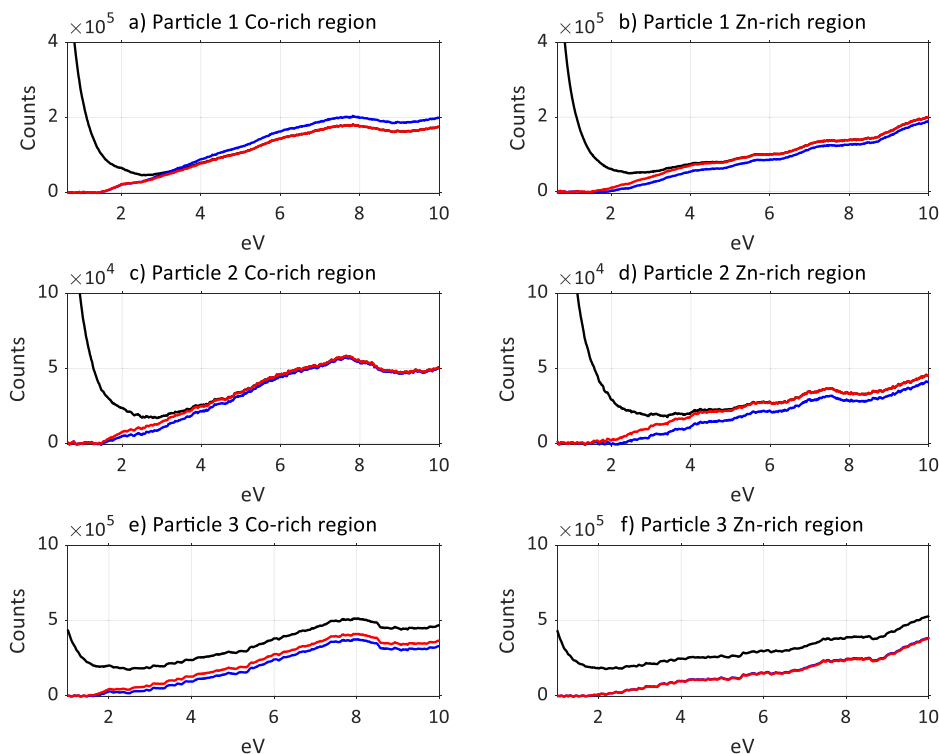

Fig. 6. The pre-edge backgrounds of metal-organic framework glass blends of $(ZIF-4-Co)_{0.5}(ZIF-62-Zn)_{0.5}$ (EELS data from Fig. S1 from the Supplementary Information of Ref. [26]) were fitted using either two-term exponential functions of the form $f(x) = ae^{bx} + ce^{dx}$ (top and middle rows) or two-term power laws of the form $f(x) = ax^b + c$ (bottom row) depending on which function gave the best goodness-of-fit. Data below 1.5 eV was used as the fitting window based on the original background subtractions carried out by Collins et al. All fits had adjusted R-square values between 0.9988 and 1.000. The SNR values were between 26 and 151. Data pre-background subtraction (black plots) as well as the original background subtraction by Collins et al. (blue plots) were plotted against the new background subtractions carried out in MATLAB (red plots).
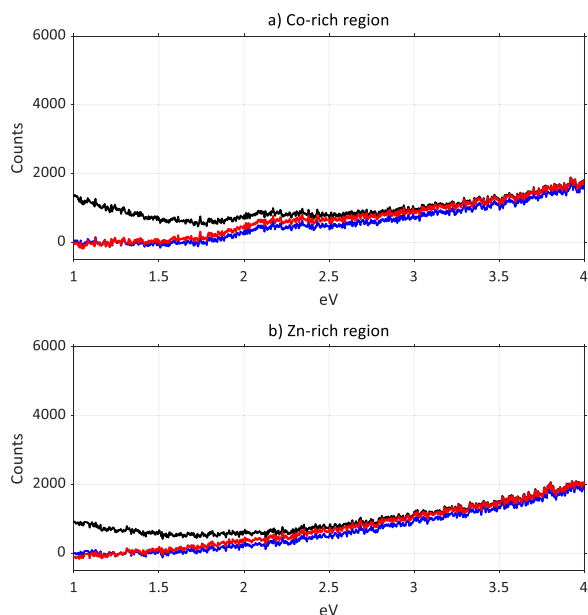
Fig. 7. The pre-edge backgrounds of metal-organic framework glass blends of $(ZIF\text{-}4\text{-}Co)_{0.5}(ZIF\text{-}62\text{-}Zn)_{0.5}$ (EELS data from Fig. S4 from the Supplementary Information of Ref. [26]) were fitted using two-term exponential functions of the form $f(x) = ae^{bx} + ce^{dx}$. Data below 1.5 eV was used as the fitting window based on the original background subtractions carried out by Collins et al. The fits had adjusted R-square values of 0.9979 and 0.9965, and SNR values of 10 and 12, for the top and bottom spectra respectively. Data pre-background subtraction (black plots) as well as the original background subtraction by Collins et al. (blue plots) were plotted against the new background subtractions carried out in MATLAB (red plots).
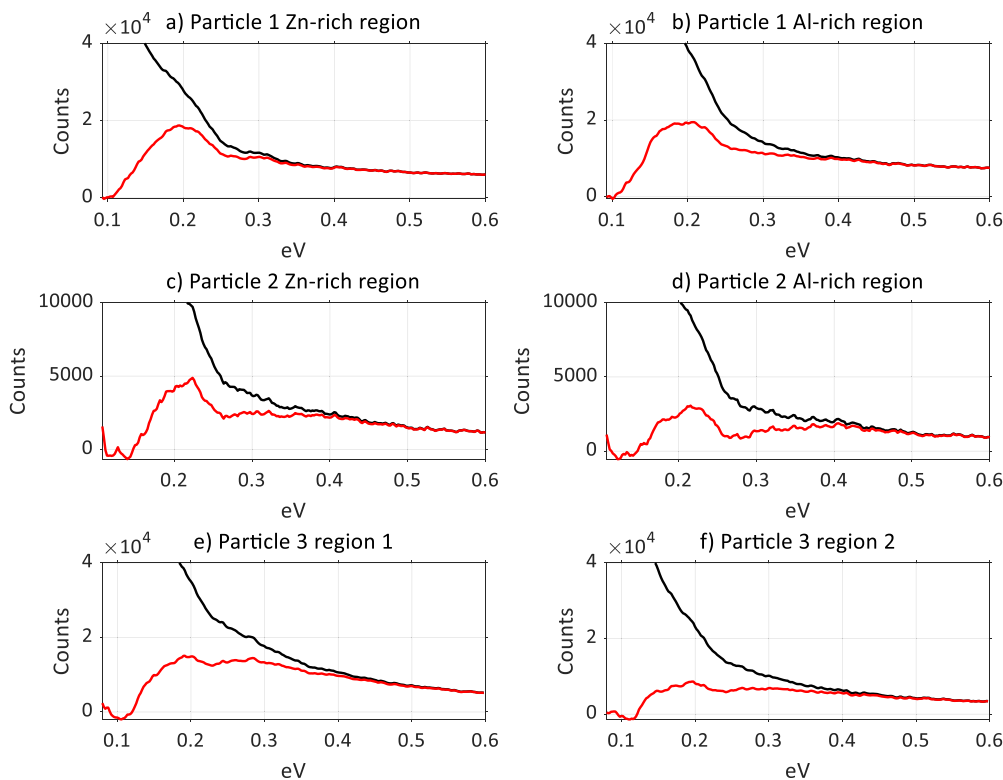
is why some of the blue spectra have a higher intensity than the black pre-background subtraction spectra. There will be trade-offs in any background fitting and this fitting routine provides one alternative which may be suitable in many applications.

The d-d ligand field transition spectra of the same type of particle with Co- and Zn-rich regions are shown in Fig. 7 [26]. This was acquired using high resolution monochromated STEM-EELS. The same colour scheme has been used where the pre-background subtraction spectra are black, the original published background subtracted spectra are blue, and the new background subtracted spectra are red. Similar to Fig. 6, all new red spectra in Fig. 7 are similar to the published blue spectra and have goodness-of-fit information. However, the SNR may not be a particularly useful metric because the signals are relatively weak in this region and signals may overlap with the ZLP. The SNR does point to intrinsic challenges in robust background subtraction in this spectral range, especially given a prominent ZLP tail or when limited spectral channels are available preceding the signals of interest. Statistical parameters nevertheless enable optimisation of the fitting and background subtraction steps.

The following datasets were acquired for crystal-glass composite particles with Zn- and Al-rich regions [27]. Background subtraction was not carried out on the ultralow-loss vibrational spectra in Fig. 8 (the spectra for similar particles was published with background subtraction) so there are no blue plots for the original published background subtracted spectra. The new red spectra have realistic intensities with the exception of some negative values close to 0 eV where irregularities in the ZLPs meant the background fittings occasionally went above the acquired spectra. This does not appear to affect the rest of the background subtraction, i.e. the fits are statistically good and the adjusted R-square values are close to 1.

Finally, the background subtracted vibrational spectra of a similar particle is shown in Fig. 9 [27]. The original spectra in blue were
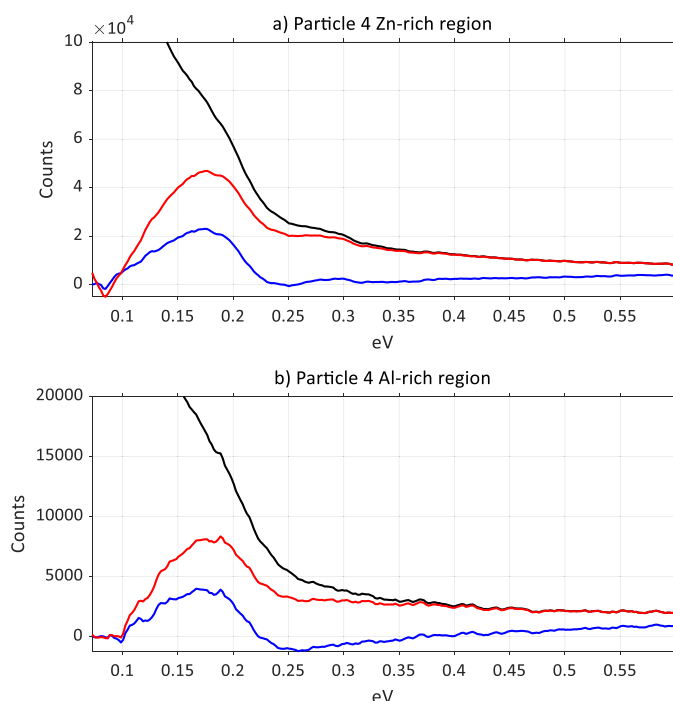

Fig. 8. The pre-edge backgrounds of metal-organic framework crystal-glass composite particles of $(MIL\text{-}53)_{0.25}(a_gZIF\text{-}62)_{0.75}$ (EELS data from Fig. S14 from the Supplementary Information of Ref. [27]) were fitted using one-term exponential functions of the form $f(x) = ae^{bx}$. Data below 0.13 eV was used as the fitting window. All fits had adjusted R-square values between 0.9925 and 0.9994. Due to intrinsically low signal levels in this energy range, the calculated SNR values appear comparatively low (below 1). Data pre-background subtraction (black plots) were plotted against the new background subtractions carried out in MATLAB (red plots).

Fig. 9. The pre-edge backgrounds of metal-organic framework crystal-glass composite particles of (MIL-53)$_{0.25}$(a_gZIF-62)$_{0.75}$ (unpublished EELS data from Ref. [27]) were fitted using one-term exponential functions of the form $f(x) = ae^{bx}$. Data below 0.10 eV was used as the fitting window. The fits had adjusted R-square values of 0.9939 and 0.9998 for the top and bottom spectra respectively. Due to intrinsically low signal in this energy range, the calculated SNR values appear comparatively low (below 1). Data pre-background subtraction (black plots) were plotted against the new background subtractions carried out in MATLAB (red plots). Unpublished background subtracted spectra by Collins et al. are shown on the same axes for comparison (blue plots).

unpublished due to difficulties with background subtraction (note the unrealistic negative values in the Fig. 9b) blue spectrum). The new red spectra fitted in MATLAB, however, have realistic intensities and the goodness-of-fit information suggest the fits are statistically good.

In summary, the MATLAB scripts presented can offer a different method for background subtraction in the low-loss region that allows the user to analyse the fit. This is particularly important for vibrational spectra where the peak positions and intensities are greatly affected by the background fit chosen. Accurate background subtractions provide more information on the sample that may otherwise be obscured by the huge contribution from the ZLP.

4. Conclusions

We have shown that the MATLAB scripts can be used to extract signals from a variety of spectra, ranging from core-loss to low and ultra-low loss EELS. The focus of this approach gives users the opportunity to quantitatively assess the goodness of their fit using the statistical information provided in order to get accurate fitting for the background and signal extraction. In addition to this, the large database of fitting models in the MATLAB Curve Fitting toolbox allows users to apply many different fits to backgrounds. This will be particularly useful for background subtraction of the low-loss region where the signals are quite weak compared to the core-loss region and accurate background fitting with different models is needed to extract the signals. While these scripts have been designed specifically for analysis of EELS data, they could very well be applied to any analysis that requires background subtraction. The MATLAB scripts provide an alternative to the widely-used GMS 3 EELS fitting software and could be ported into GNU Octave, an open-source program which has many features similar

to MATLAB, or indeed Python as mentioned previously.

5. Methods

The single-walled carbon nanotubes (SWNTs) were purchased from Carbon Solutions, Inc. (product code P2-SWNT). The lacey carbon film copper (200 mesh) TEM grid that the SWNTs were deposited onto for EELS analysis were purchased from Agar Scientific Ltd.

The SWNTs EEL spectrum was recorded on a JEOL 2100Plus operating at an accelerating voltage of 80 kV at the University of Nottingham Nanoscale and Microscale Research Centre (nmRC), equipped with a Gatan Enfinium. The spectrum was acquired in imaging mode with an effective convergence angle of < 3 mrad and a collection angle of > 100 mrad. A total of 20 individual exposures of 0.05 s was accumulated with a dispersion of 0.1 eV/channel.

5.1. Processing and cleaning of single-walled carbon nanotubes

Single-walled carbon nanotubes (50 mg) were refluxed in nitric acid (3 M, 50 mL) for 2 hr, then washed with deionised water (500 mL) and dried under vacuum. The dried nanotubes (11 mg) were heated at their oxidation temperature of 600 °C for 17 min to give cleaned single-walled carbon nanotubes as a black powder (2.3 mg).

5.2. Preparing the EELS sample of single-walled carbon nanotubes

The processed single-walled carbon nanotubes were sonicated in HPLC-grade isopropyl alcohol and drop-cast, or deposited, onto a lacey carbon film copper TEM grid.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ultramic.2020.113052.

References

[1] R.F. Egerton, Electron energy-loss spectroscopy in the TEM, Rep. Prog. Phys. 72 (2009) 016502.
[2] H.R. Daniels, R. Brydson, A. Brown, B. Rand, Quantitative valence plasmon mapping in the TEM: viewing physical properties at the nanoscale, Ultramicroscopy 96 (2003) 547–558.
[3] P. Rez, D.A. Muller, The Theory and Interpretation of Electron Energy Loss Near-Edge Fine Structure, Annu. Rev. Mater. Res. 38 (2008) 535–558.
[4] B. Rafferty, L.M. Brown, Direct and indirect transitions in the region of the band gap using electron-energy-loss spectroscopy, Phys. Rev. B 58 (1998) 10326–10337.
[5] K. Kimoto, G. Kothleitner, W. Grogger, Y. Matsui, F. Hofer, Advantages of a monochromator for bandgap measurements using electron energy-loss spectroscopy, Micron 36 (2005) 185–189.

[6] O.L. Krivanek, T.C. Lovejoy, N. Dellby, T. Aoki, R.W. Carpenter, P. Rez, E. Soignard, J. Zhu, P.E. Batson, M.J. Lagos, R.F. Egerton, P.A. Crozier, Vibrational spectroscopy in the electron microscope, Nature 514 (2014) 209–212.

[7] R.F. Egerton, in *Electron Energy-Loss Spectroscopy in the Electron Microscope*, Springer US, Boston, MA, 1996, pp. 245–299.

[8] R. BrydsonRoyal Microscopical Society (Great Britain), Electron Energy Loss Spectroscopy, Taylor & Francis, Oxford, 2001.

[9] R.F. Egerton, Electron Energy-Loss Spectroscopy in the Electron Microscope, Plenum Press, New York, 1986.

[10] H. Tenailleau, J.M. Martin, A new background subtraction for low-energy EELS core edges, J. Microsc. 166 (1992) 297–306.

[11] G. Balossier, X. Thomas, J. Michel, D. Wagner, P. Bonhomme, E. Puchelle, D. Ploton, A. Bonhomme, J.M. Pinon, Parallel EELS elemental mapping in scanning transmission electron microscopy: use of the difference methods, Microsc. Microanal. Microstruct. 2 (1991) 531.

[12] J.A. Hachtel, A.R. Lupini, J.C. Idrobo, Exploring the capabilities of monochromated electron energy loss spectroscopy in the infrared regime, Sci. Rep. 8 (2018) 5637.

[13] P.M.A. Sherwood, The use and misuse of curve fitting in the analysis of core X-ray photoelectron spectroscopic data, Surf. Interface Anal. (2019) 1–22.

[14] R. Hovden, P. Cueva, J.A. Mundy, D.A. Muller, The Open-Source Cornell Spectrum Imager, Micros. Today 21 (2013) 40–44.

[15] P. Cueva, R. Hovden, J.A. Mundy, H.L. Xin, D.A. Muller, Data Processing for Atomic Resolution Electron Energy Loss Spectroscopy, Microsc. Microanal. 18 (2012) 667–675.

[16] J. Verbeeck, S. Van Aert, Model based quantification of EELS spectra, Ultramicroscopy 101 (2004) 207–224.

[17] R. Door, D. Gängler, Multiple least-squares fitting for quantitative electron energy-loss spectroscopy — An experimental investigation using standard specimens, Ultramicroscopy 58 (1995) 197–210.

[18] R.F. Egerton, M. Malac, Improved background-fitting algorithms for ionization edges in electron energy-loss spectra, Ultramicroscopy 92 (2002) 47–56.

[19] J.T. Held, H. Yun, K.A. Mkhoyan, Simultaneous multi-region background subtraction for core-level EEL spectra, Ultramicroscopy 210 (2020) 112919.

[20] C.S. Granerød, W. Zhan, Ø. Prytz, Automated approaches for band gap mapping in STEM-EELS, Ultramicroscopy 184 (2018) 39–45.

[21] S. Lu, K.J. Kormondy, A.A. Demkov, D.J. Smith, An EELS signal-from-background separation algorithm for spectral line-scan/image quantification, Ultramicroscopy 195 (2018) 25–31.

[22] G. Kothleitner, F. Hofer, Optimization of the signal to noise ratio in EFTEM elemental maps with regard to different eonization edge types, Micron 29 (1998) 349–357.

[23] T. Pun, J.R. Ellis, M. Eden, Weighted least squares estimation of background in EELS imaging, J. Microsc. 137 (1985) 93–100.

[24] M. Pacella, P.W. Butler-Smith, D.A. Axinte, M.W. Fay, FIB/TEM/EELS micro/nanometric investigations of the effects of laser ablation on the diamond/binder structure in polycrystalline diamond composites, J. Mater. Process. Technol. 214 (2014) 1153–1161.

[25] K.E. Walker, G.A. Rance, Á. Pekker, H.M. Tóháti, M.W. Fay, R.W. Lodge, C.T. Stoppiello, K. Kamarás, A.N. Khlobystov, Growth of Carbon Nanotubes inside Boron Nitride Nanotubes by Coalescence of Fullerenes: toward the World's Smallest Coaxial Cable, Small Methods 1 (2017) 1700184.

[26] S.M. Collins, D.M. Kepaptsoglou, K.T. Butler, L. Longley, T.D. Bennett, Q.M. Ramasse, P.A. Midgley, Subwavelength Spatially Resolved Coordination Chemistry of Metal–Organic Framework Glass Blends, J. Am. Chem. Soc. 140 (2018) 17862–17866.

[27] S.M. Collins, D.M. Kepaptsoglou, J. Hou, C.W. Ashling, G. Radtke, T.D. Bennett, P.A. Midgley, Q.M. Ramasse, Functional Group Mapping by Electron Beam Vibrational Spectroscopy from Nanoscale Volumes, Nano Lett 20 (2020) 1272–1279.