Received: 9 May 2019 | Revised: 12 August 2019 | Accepted: 12 August 2019

DOI: 10.1002/sia.6702

RESEARCH ARTICLE

WILEY SURFACE and INTERFACE ANALYSIS

Automatic spike correction using UNIFIT 2020

Ronald Hesse[1] | Carsten Bundesmann[2] | Reinhard Denecke[1]

[1] Wilhelm-Ostwald-Institute for Physical and Theoretical Chemistry, Leipzig University, Leipzig, Germany

[2] Department of Precision Surfaces, Ion Source Development and Applications Group, Leibniz Institute of Surface Engineering (IOM), Leipzig, Germany

Correspondence
Ronald Hesse, Wilhelm-Ostwald-Institute for Physical and Theoretical Chemistry, Leipzig University, Linnéstr.2, D-04103 Leipzig, Germany.
Email: rhesse@rz.uni-leipzig.de

The improvement of the software UNIFIT 2020 from an analysis processing software for photoelectron spectroscopy (XPS) only to a powerful tool for XPS, Auger electron spectroscopy (AES), X-ray absorption spectroscopy (XAS), and Raman spectroscopy requires new additional programme routines. Particularly, the implementation of the analysis of Raman spectra needs a well-working automatic spike correction. The application of the modified discrete Laplace operator method allows for a perfect localization and correction of the spikes and finally a successful peak fit of the spectra. The theoretical basis is described. Test spectra allow for the evaluation of the presented method. A comparison of the original and spike-corrected real measurements demonstrates the high quality of the method used.

KEYWORDS

Auger electron spectroscopy AES, photoelectron spectroscopy XPS, Raman spectroscopy, spectrum analysis software, spike correction, UNIFIT, X-ray absorption spectroscopy XAS

1 | INTRODUCTION

The long way of the development of the UNIFIT software to a powerful tool for photoelectron spectra (XPS)[1] requires a continuous improvement of the programme. The high complexity of the modelling of the spectral-background function in combination with the peak fit of the photoelectron lines requires innovative processing methods, spectrometer calibration, and validation methods of the received results. Some worldwide standalone features are the following:

1. Estimation of the uncertainties of the fit results using two different statistic methods[2];
2. Offering of three different model functions for the modelling of peaks using relative or absolute fit parameters[3];
3. Calculation of the spectral background using the improved Tougaard method and fittable parameters[4];
4. Estimation of the transmission function of spectrometers with two different methods for a higher accuracy of the quantification[5];
5. Advanced Tougaard background calculation method for the analysis of laterally inhomogeneous samples.[6]

The expansion of the software to other spectroscopic techniques has two sides: on the one hand, a lot of existing programme routines may be used for all techniques. On the other hand, different analysis methods require additional programme tools for an appropriate support for the analysis of the data.

For the extension of the software to the analysis of the Auger electron spectra (AES), we have developed and implemented six elementary new subroutines.

1. Tool for semi-empirical calculation of AES sensitivity factors with the following input data: atomic number, value for K-, L-, M-, and N-electrons, number of electrons in the subshell, excitation energy, and the binding energy of the target atom subshell.[7]
2. Implementation of the database of the binding energies of the target atom subshells.
3. Implementation of a database of energies of Auger electron transitions.
4. Implementation of empirical sensitivity factors of frequently used spectrometer types (eg, PHI 700).
5. Support of the quantification using the peak-to-peak values.

6. Generation of an input routine for SEM images for the comparison with SAM data.

The software support for X-ray absorption spectra (XAS) requires the following two completely new routines.

1. Development of a method for the correct modelling of the spectral background of XA spectra (integral of the corresponding XP spectra),[8] and
2. because the UNIFIT software works generally with equidistant step widths, input routines have to be developed for the conversion of non-monotonic energy scale and non-equidistant step widths to a monotonic energy scale and equidistant step widths.

The idea of the implementation of the analysis of Raman spectra was suggested by UNIFIT users. In all cases, a peak fit of Raman data could be generated using the UNIFIT software only. In order to fit a large number of spectra affected with spikes (not rarely in Raman) of a batch measurement, an automatic spike correction is necessary. Because former versions of the UNIFIT software offer a manual spike correction only, an automatic spike correction had to be developed, tested, and implemented into the new version. The following new routines were implemented:

1. Integration of an automatic spike correction.
2. Integration of input routines with an automatic conversion of non-equidistant step widths to equidistant step width. The loading of reference spectra from the RRUFF database[9] is possible.

In general, Raman spectrometers utilize a charge-coupled device (CCD) as detector, because of its high quantum efficiency, great sensitivity, high dynamic range, linear response to photons, high reliability, and low dark current.[10,11] CCDs are based on generation of electrons in a sensor element by photons and subsequent readout of the accumulated charge. A CCD consists of a line or array of sensor elements (pixels) for spectroscopic or imaging applications. Electrons cannot only be generated by photons but also by other so-called "noise sources." One of them are cosmic rays, such as, muons and other ionizing particles.[10,12] If cosmic rays hit the CCD, several thousand electrons can be generated in a sensor element or neighbouring sensor elements. In case of Raman spectroscopy, the detected intensity of the sample signal is low, and, therefore, the electrons related to cosmic noise give rise to sharp lines in the Raman spectra, so-called "cosmic spikes."

Removing cosmic spikes is necessary for accurate data analysis of the Raman signal. At first glance, spike removal appears to be a simple task. Basically, there are several hardware or software-based approaches[11,13]: (a) additional acquisition-based methods; (b) methods involving hardware modification; and (c) correction of single-scan data by filtering or smoothing. Because of time and costs, approach (c) is desirable if large number of data sets are to be processed.

Correction-based methods include, for instance, application of wavelet transforms,[11,12] weighted moving window filtering,[13] and median or polynomial filters.[10,14,15] All approaches can handle the problem quite well. However, none of the existing approaches can treat all aspects perfectly and fully automated, especially, if the linewidth of the cosmic spikes is similar to that of certain Raman lines. Limitations are related, for instance, to the complexity of the approach and the need of properly chosen pre-settings, such as, threshold parameters for minimizing distortion or wavelet transform for correct identification of cosmic spikes.[11,12]

2 | THEORETICAL BACKGROUND

Before the development of the automatic spike correction was started, several suggestions in literature were tested.[11,16] We have developed software codes for testing and using the recommended methods. However, we could not confirm the correct working of the suggested methods. The recommended methods could not give successful results. Additionally, no information regarding the boundary conditions

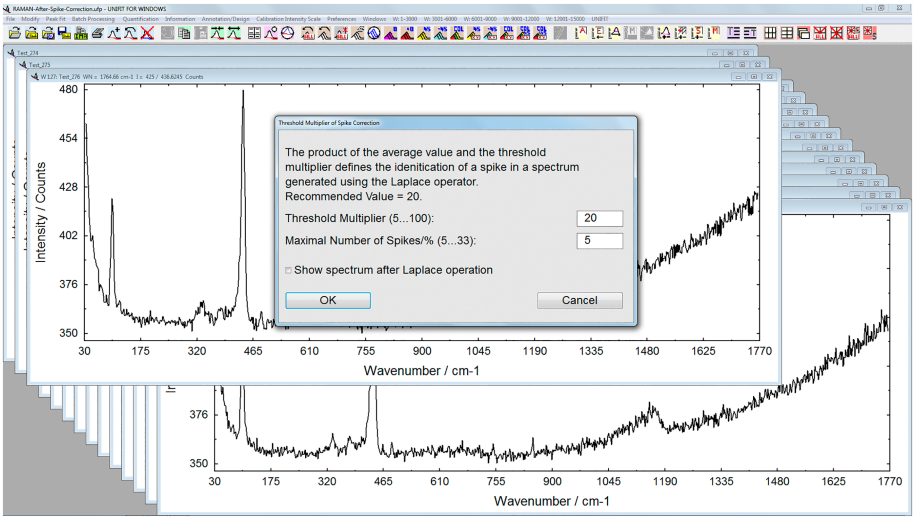

FIGURE 1 Input dialogue for the definition of the threshold multiplier T and the maximal number of spikes N'

(eg, first spectrum/first channel, last spectrum/last channel, ...) were available. In the presented paper, all possible boundary conditions are discussed.

The theoretical basics of the automatic spike correction of a two-dimensional data set (a series of spectra) is a modified Laplace method suggested by Ryabchykov.[16] The spikes are corrected using the intensity of the lower measurement channel at the same spectrum and the intensities of neighbouring spectra consecutively. The maximum number of spikes (in percentage, with respect to all measuring points, 5% ... 33%) and the threshold multiplier (5 ... 100) for the identification of the spikes have to be defined by the user. The general form of the discrete Laplace operator is given by

$$L(i,\ j) = 4{\cdot}M(i,\ j){-}M(i{-}1,\ j){-}M(i+1,\ j){-}M(i,\ j{-}1){-}M(i,\ j+1), \quad (1)$$

with the measured spectra $M(i, j)$ recorded at N energy/wavenumber values corresponding to channels i and spectrum j. All spectra must have the same number of channels N. The number of the recorded spectra is O. The spike correction is carried out in five steps.

2.1 | Step 1: Definition of the threshold multiplier

In order to determine the existence of a spike as compared with a real spectral feature, the value of the Laplace operator has to be considered. A threshold value (here given by the threshold multiplier and an average value) has to be defined to distinguish. The definition of the threshold multiplier T (typically 20) and the maximal number of spikes N' per spectrum (typically 5% of all data points, $N' = N{\cdot}0.05$) has to be carried out manually using the programme presenting (see Figure 1). Optionally, the calculated spectra modified with the Laplace operator (Equation 1) can be displayed (see Figure 2).

2.2 | Step 2: Generation of the Laplace spectra

The location of spikes has to be determined. Therefore, the experimental data are subjected to a discrete Laplace operator (see, eg, equation 3 in Ryabchykov et al[16]), strongly emphasizing abrupt intensity changes in the spectra. Special attention has to be paid to data points at the edges of the two-dimensional data set. The generation of the modified spectra using the Laplace operator $L(i, j)$ (Laplace spectra) for all nine possiblse cases (see Figure 3) is defined by (Equation 2):

$$L(i,\ j) = \begin{cases} 4{\cdot}M(i,\ j){-}M(i{-}1,\ j){-}M(i+1,\ j){-}M(i,\ j{-}1){-}M(i,\ j+1) & 1 < i < N \quad 1 < j < O \\ 4{\cdot}M(i,\ j){-}M(i{-}1,\ j){-}M(i+1,\ j){-}2{\cdot}M(i,\ j+1) & 1 < i < N \quad j = 1 \\ 4{\cdot}M(i,\ j){-}M(i{-}1,\ j){-}M(i+1,\ j){-}2{\cdot}M(i,\ j{-}1) & 1 < i < N \quad j = O \\ 4{\cdot}M(i,\ j){-}2{\cdot}M(i+1,\ j){-}M(i,\ j{-}1){-}M(i,\ j+1) & i = 1 \quad 1 < j < O \\ 4{\cdot}M(i,\ j){-}2{\cdot}M(i{-}1,\ j){-}M(i,\ j{-}1){-}M(i,\ j+1) & i = N \quad 1 < j < O \\ 4{\cdot}M(i,\ j){-}2{\cdot}M(i+1,\ j){-}2{\cdot}M(i,\ j+1) & i = 1 \quad j = 1 \\ 4{\cdot}M(i,\ j){-}2{\cdot}M(i+1,\ j){-}2{\cdot}M(i,\ j{-}1) & i = 1 \quad j = O \\ 4{\cdot}M(i,\ j){-}2{\cdot}M(i{-}1,\ j){-}2{\cdot}M(i,\ j+1) & i = N \quad j = 1 \\ 4{\cdot}M(i,\ j){-}2{\cdot}M(i{-}1,\ j){-}2{\cdot}M(i,\ j{-}1) & i = N \quad j = O \end{cases} \quad (2)$$

2.3 | Step 3: Determination of the average values *A(j)*

The actual threshold value depends on the average intensity of the respective Laplace spectra $L(i, j)$ without the highest intensities, lowest intensities, and negative values. The calculation of the reference-average values $A(j)$ of every Laplace spectrum $L(i, j)$, which is then multiplied with the defined threshold multiplier T for comparison with the Laplace spectra $L(i, j)$, is carried out using five steps:

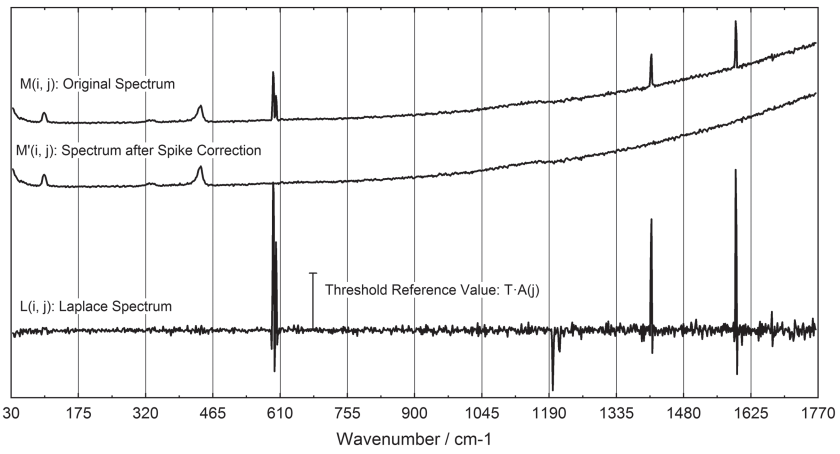

FIGURE 2 Illustration of the original spectrum $M(i, j)$, the spectrum after spike correction $M'(i, j)$, the Laplace spectrum $L(i, j)$, and reference value $T{\cdot}A(j)$

1. Determination of the channel numbers $P(k, l)$ ($k = 1...N', l = 1...O$) of the $N'{\cdot}O$ positions of the elements with the highest intensities of the two-dimensional field $L(i, j)$.
2. Determination of the channel numbers $Q(k, l)$ ($k = 1...N', l = 1,,,O$) of the $N'{\cdot}O$ positions of the elements with the lowest intensities of the two-dimensional field $L(i, j)$.
3. Determination of the number of elements with negative intensities $N_{neg}(j)$ of every spectrum j of the field $L(i, j)$ with exception of the elements of $P(i, j)$ and $Q(i, j)$:

$$N_{neg}(j) = \sum_{i=1}^{N} \begin{cases} 1 & for \quad L(i,\ j) < 0, \quad i \notin P(k,\ j), \quad i \notin Q(k,\ j) \\ 0 & for \quad L(i,\ j) \geq 0, \quad i \in P(k,\ j), \quad i \in Q(k,\ j) \end{cases}. \tag{3}$$

4. Calculation of the number of channels $N_{ave}(j)$ for the calculation of $A(j)$:

$$N_{ave}(j) = N - 2{\cdot}N' - N_{neg}(j). \tag{4}$$

5. Calculation of $A(j)$ with exception of the channels $P(k, l)$, $Q(k, l)$ and the elements with a negative intensity of $L(i, j)$:

$$A(j) = \frac{1}{N_{ave}(j)} \sum_{i=1}^{N} L(i,\ j) \quad for \quad L(i,\ j) \geq 0, \quad i \notin P(k,\ j), \quad i \notin Q(k,\ j). \tag{5}$$

2.4 | Step 4: Determination of the spike positions

Spike positions are found if the values of the Laplace spectra exceed the threshold set by the multiplication of the average value $A(j)$ and the threshold multiplier T. Thus, the definition of the positions of the spikes $P(k, j)$ of every spectrum j is defined by

$$P(k,\ j) = \begin{cases} 0 & for \quad L(P(k,\ j),\ j) \leq T{\cdot}A(j), \quad k = 1...N' \\ P\{k,\ j & for \quad L(P(k,\ j),\ j) > T{\cdot}A(j), \quad k = 1...N' \end{cases}. \tag{6}$$

All channels $P(k, j) > 0$ define a spike position in the spectrum j.

2.5 | Step 5: Spike correction

Finally, the data points identified as spikes need to be corrected. If the channel i_S of the spectrum j is an element of $P(k, j)$, and therefore $M(i_S, j)$ is a spike, then the intensities of the five channels (if available), $M(i_S - 2, j)$, $M(i_S - 1, j)$, $M(i_S, j)$, $M(i_S + 1, j)$, and $M(i_S + 2, j)$, are changed. This is an important point because the spikes of the Laplace spectrum $L(i, j)$ are sharper with respect to the same spikes of the original spectrum $M(i, j)$.

The correction is carried out consecutively (by stepping i for each spectrum j) from $M(1, 1)$, $M(2, 1)$, ... , $M(N, 1)$, $M(1, 2)$, ... , to $M(N, O)$. The actual calculation of the corrected points depends on the following two aspects: (a) the position of the spike and (b) the possible presence of a neighbouring spike. In order to distinguish the latter, five different cases with adapted calculations have to be applied. The corrected spectra $M'(i, j)$ are calculated by

Case I. The measuring points i, $i - 1$, and $i - 2$ of the spectrum $j + 1$ are not elements of $P(k, j + 1)$, $j < O$, ($k = 1 \ldots N'$) (Equation 7),

Case II. The measuring point i, $i - 1$, or $i - 2$ of the spectrum $j + 1$ is an element of $P(k, j + 1)$, $j < O$, ($k = 1 \ldots N'$) (Equation 8):

$$M'(i,\ j) = \begin{cases} (M(i{-}1,\ j) + M(i{-}1,\ j{-}1) + M(i{-}1,\ j + 1) + M(i,\ j{-}1) + M(i,\ j + 1))/5, & i > 1, \quad 1 < j < O \\ (M(i{-}1,\ 1) + M(i,\ 2) + M(i{-}1,\ 2))/3, & i > 1, \quad j = 1 \\ (M(1,\ j{-}1) + M(1,\ j + 1))/2, & i = 1, \quad 1 < j < O \\ M(1,\ 2), & i = 1, \quad j = 1 \end{cases}. \tag{7}$$

$$M'(i,\ j) = \begin{cases} (M(i{-}1,\ j) + M(i{-}1,\ j{-}1) + M(i,\ j{-}1))/3, & i > 1, \quad 1 < j < O \\ M(i{-}1,\ 1), & i > 1, \quad j = 1 \\ M(1,\ j{-}1), & i = 1, \quad 1 < j < O \\ M(1,\ 1), & i = 1, \quad j = 1 \end{cases}. \tag{8}$$

Case III. The spectrum number is $j = O$ (last spectrum of the series):

$$M'(i,\ O) = (M(i{-}1,\ O) + M(i,\ O{-}1) + M(i{-}1,\ O{-}1))/3, \quad i > 1, \quad j = O \tag{9}$$

$$M'(1,\ O) = M(1,\ O{-}1), \quad i = 1, \quad j = O \tag{10}$$

$$M'(i,\ j) = (M(i{-}1,\ j) + M(i,\ j{-}1) + M(i{-}1,\ j{-}1))/3. \quad i = N, \quad j = O \tag{11}$$

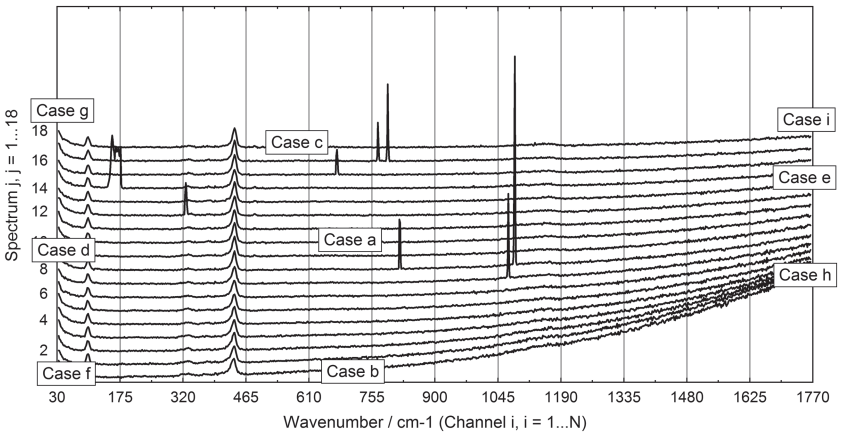

FIGURE 3 The nine cases for the calculation using the discrete Laplace operator: case a: $1 < i < N$, $1 < j < O$; case b: $1 < i < N$, $j = 1$; case c: $1 < i < N$, $j = O$; case d: $i = 1$, $1 < j < O$; case e: $i = N$, $1 < j < O$; case f: $i = 1$, $j = 1$; case g: $i = 1$, $j = O$; case h: $i = N$, $j = 1$; case i: $i = N$, $j = O$

Case IV. The spectrum number is $j = 1$ (first spectrum of the series) and the measuring point $i = N$ (last point in the spectrum):

$$M'(N, 1) = M(N-1, 1). \quad i = N, \quad j = 1 \tag{12}$$

Case V. The measuring points $i - 2$, $i - 1$, i, $i + 1$, and $i + 2$ of the spectrum j are not elements of $P(k, j)$, ($k = 1 \ldots N'$) and, therefore, the measuring point i is not a spike:

$$M'(i, j) = M(i, j). \tag{13}$$

Figure 2 illustrates the different spike correction steps and Figure 3 the positions of the different spike-correction cases.

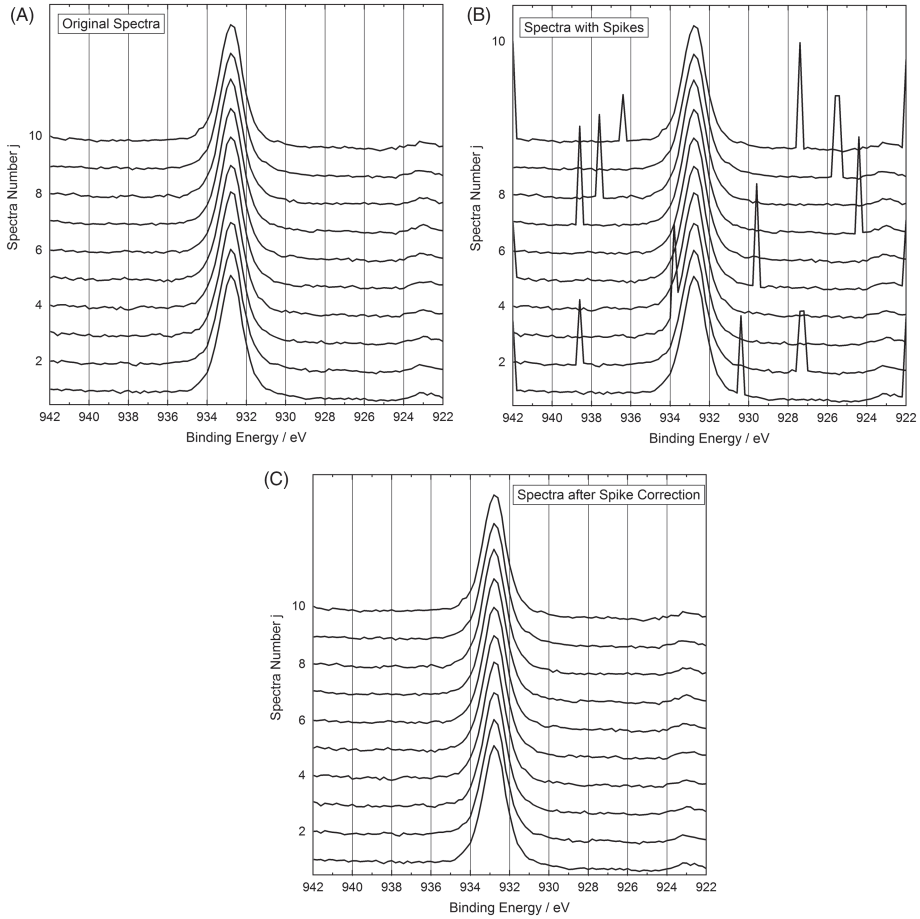

FIGURE 4 A, Ten original XP spectra $M_{Ori}(i, j)$ represent a part of a batch measurement for the test of the stability of the photoelectron spectrometer ESCALAB 220i. B, Ten XP spectra $M(i, j)$ (see A) modified with 19 additional spikes. The positions of the spikes represent all possible cases a to i. C, Ten spectra $M'(i, j)$ after automatic spike correction using the spectra of B

3 | VALIDATION OF THE METHOD USING TEST SPECTRA

In order to verify the calculated new intensities of the spikes, we have generated XP spectra $M(i, j)$ with spikes (see Figure 4B) using XP spectra $M_{Ori}(i, j)$ without spikes (see Figure 4A). The spikes were manually implemented into the spectra at all positions of the different cases a to i (see Figure 4B). The number of spikes is 19 and the number of spectra is 10. Table 1 gives the values for the spike positions before and after adding the artificial spikes. After the spike correction, the intensities at the spike positions of the spectra $M'(i, j)$ were compared with respect to the original intensities. This should have the same statistical behaviour as the measured data. In order to verify the calculated intensities, we have calculated the normalized residual R_k for all spikes k = 1 to 19:

$$R_k = \frac{M_{Ori}(i_k, j_k) - M'(i_k, j_k)}{SN_k}, \quad SN_k \neq 0 \tag{14}$$

with the original intensities $M_{Ori}(i, j)$, the corrected intensities $M'(i, j)$, the channel numbers of the spikes i_k of the spectrum j_k, and

$$SN_k = \sqrt{M_{Ori}(i_k j_k)}, \; M_{Ori}(i_k j_k) > 0. \tag{15}$$

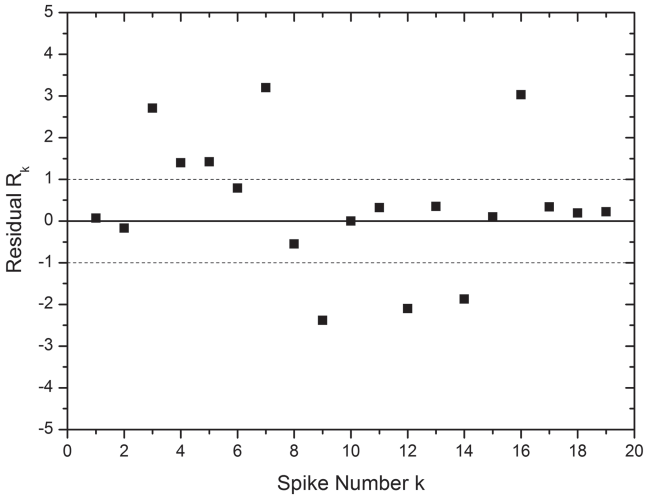

FIGURE 5 The evaluation of the corrected intensities with the residual; 59% of R_K are between 1 and −1. The expected value for a complete statistic behaviour is 68%

Table 1 gives the calculated values. Of the R_k values, 59% are between −1 and +1 (see also Figure 5). The expected value for statistical noise is 68%. Figure 4C illustrates the spectra after the automatic spike correction.

TABLE 1 Comparison of the intensities at the spike channels i_k of the original spectrum $M_{Ori}(i, j)$ without spikes, of the spectrum $M(i,j)$ with spikes, and the spectrum $M'(i, j)$ after spike correction; column 6 gives the value for the statistical noise $SN_k = \sqrt{M_{Ori}(i_k, j_k)}$ for M_{Ori} $(i_k j_k) > 0$ and column 8 gives the residual $R_k = \frac{M_{Ori}(i_k, j_k) - M'(i_k, j_k)}{SN_k}$, see also Figure 5

Spike number k	i_k, j_k (case)	Energy E_B/eV	M_{Ori} $(i_k j_k)$/counts	M (i_k, j_k)/counts	SN_k/counts$^{1/2}$	$M'(i_k j_k)$/counts	R_k counts$^{1/2}$
1	1, 1 (f)	942.0	4484	9162	67	4479	0.07
2	59, 1 (b)	930.4	4300	9498	66	4311	−0.17
3	101, 1 (h)	922.0	3860	9397	62	4028	−2.71
4	18, 2 (a)	938.6	4316	8702	66	4293	1.40
5	73, 2 (a)	927.4	3913	7924	63	3823	1.42
6	74, 2 (a)	927.2	3881	7941	62	3835	0.79
7	42, 3 (a)	933.8	6358	11 910	80	6102	3.20
8	1, 5 (d)	942.0	4449	8333	67	4486	−0.55
9	63, 5 (a)	929.6	3910	10 882	63	4060	−2.38
10	101, 5 (e)	922.0	3993	8431	63	3936	0.90
11	18, 7 (a)	938.6	4320	11 004	66	4299	0.32
12	177, 7 (a)	924.4	3642	10 284	60	3768	−2.10
13	45, 8 (a)	937.6	4343	9916	66	4320	0.35
14	82, 9 (a)	925.6	3691	9297	61	3805	−1.87
15	83, 9 (a)	925.4	3800	9297	62	3794	0.10
16	1, 10 (g)	942.0	4643	11 087	68	4437	3.03
17	29, 10 (c)	936.4	4345	7490	67	4322	0.34
18	74, 10 (c)	927.4	3834	11 004	62	3821	0.19
19	101, 10 (i)	922.0	3947	9866	63	3933	0.22

4 | AUTOMATIC SPIKE CORRECTION OF REAL SPECTRA

The automatic spike correction was additionally tested using real spectra. Spike affected Raman and XP spectra were corrected.

4.1 | Raman spectra

As a test, a set of experimental Raman spectra were used for spike correction. The spectra were measured during the cool down phase of a ZnO bulk sample. Figure 6 shows, exemplarily, one curve of the set. The measurement represents an unpolarised Raman spectrum. Typical features can be identified[17,18]: $E_2^{(1)}$ and $E_2^{(2)}$ mode at about 101 and 439 cm^{-1}, respectively, and multiphonon (MP) structures at 330 and 1155 cm^{-1}. The rise of the signal intensity with increasing wavenumber is related to thermal radiation of the heated sample.

Figure 7A contains all Raman spectra of the data set. Each time index represents a time interval of 60 seconds. At the beginning (time index 196), the sample temperature was about 800°C. At the end (time index 276), it was about 65°C. Because of the decreasing temperature, the intensity rise towards higher wavenumbers gets smaller. Furthermore, the central wavenumber, broadening and intensity of the phonon modes are affected.[19,20]

The curves in Figure 7A are uncorrected, and a large number of spikes are visible. Upon spike correction (threshold multiplier: 20, maximal number of spikes: 5% of the number of data points), these unwanted features can be removed completely without changing the shape of the phonon mode features and contribution of thermal radiation (see Figure 7B).

4.2 | XPS spectra

An XPS batch measurement of the Cu 3p peak (see Figure 8) of the sample manipulator was carried out. Fifty spectra were recorded. The spikes have been implemented by electronic interferences with light switches. In order to generate spikes at different positions into the spectrum, the measured time of one spectrum and the switching frequency were different (measurement time per spectrum: 0.26 min,

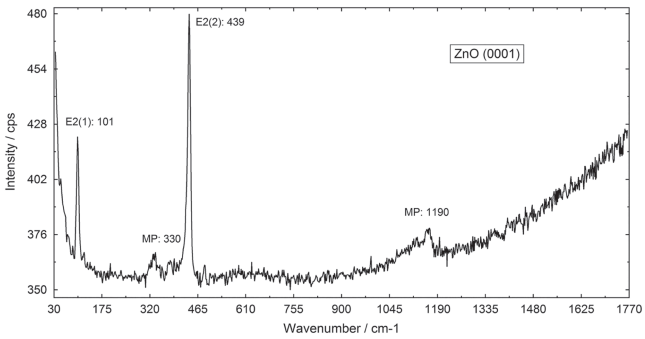

FIGURE 6 Selected, unpolarised Raman spectrum of a heated ZnO (0001) bulk sample with typical features related to $E_2^{(1)}$ and $E_2^{(2)}$ and multiphonon (MP) modes. The signal increase with increasing wavenumber is caused by thermal radiation

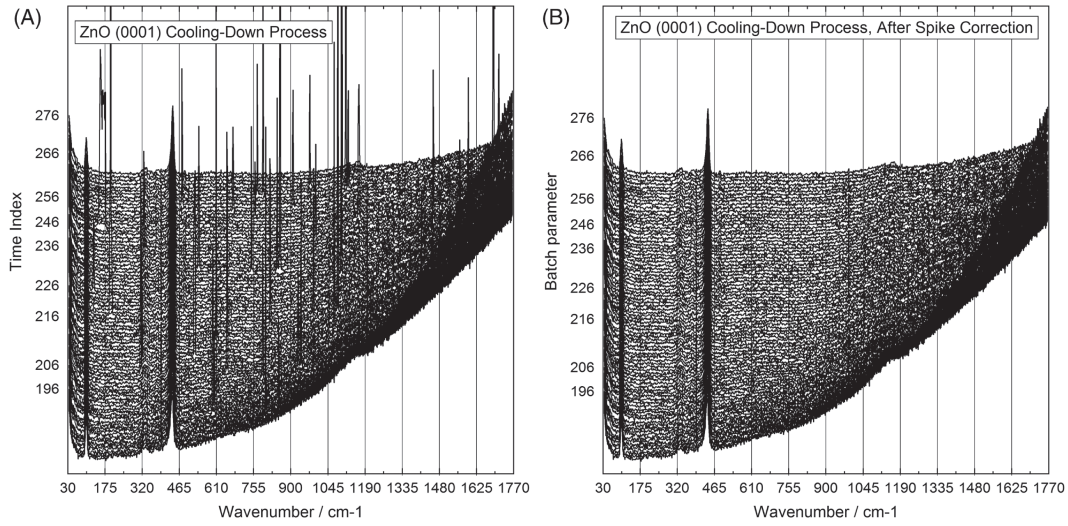

FIGURE 7 A, Presentation of all original spectra of ZnO (0001) with spikes during the cooling-down process. B, Presentation of all original spectra of ZnO (0001) after automatic spikes correction during the cooling down process

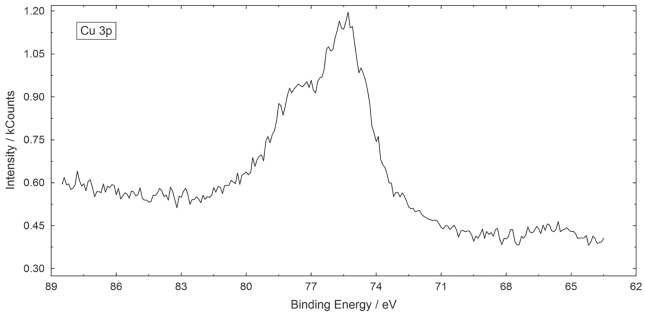

FIGURE 8 Presentation of one Cu 3p XP spectrum of the series without spike

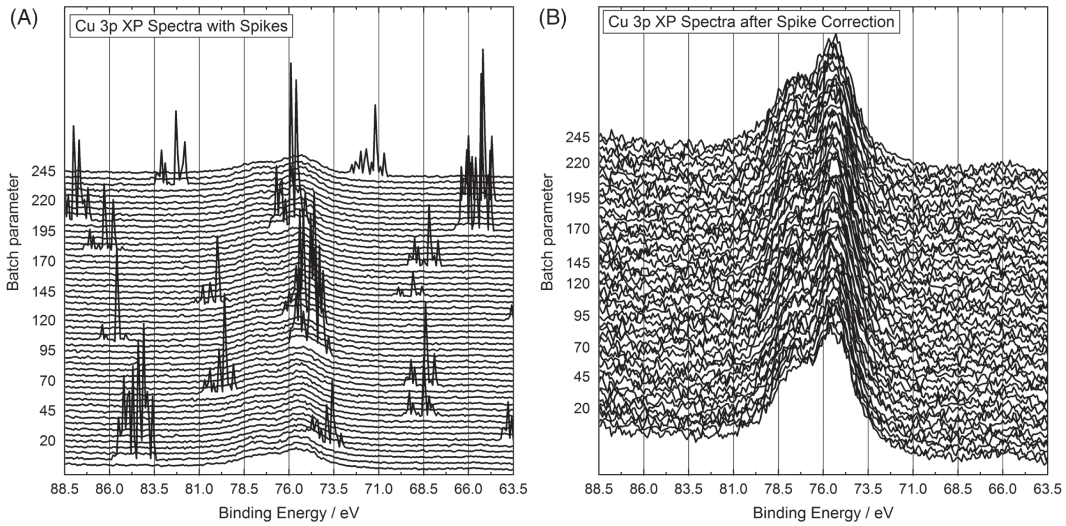

FIGURE 9 A, Presentation of all Cu 3p XP spectra with generated spikes. B, Presentation of all the Cu 3p XP spectra after the automatic spike correction

switching frequency: 0.5 min^{-1}). Figure 9A illustrates the spectra with the generated strong spikes. Because the spectrometer ESCALAB 220 iXL detector has six channeltrons, each switch generates more than one spike.

After the measurement, the automatic spike correction with the default setting (threshold multiplier: 20, maximal fraction of spikes: 5% of the number of data points) was carried out. The intensities of all corrected spectra show a statistic behaviour. Particularly at the energy positons of the spikes, no incorrect features may be found. Figure 9B presents the spectra after the spike correction.

5 | LIMIT TESTS OF THE ALGORITHM

Finally, we have tested the limits of the correct performance of the suggested method.

5.1 | Number of spectra

The minimal number of spectra with affected spikes is two. The maximal number is the number of simultaneously loadable spectra (currently 75 600 spectra). The operation time can be reduced in case of a very large number of spectra by optimization of the UNIFIT software setting (eg, hiding of windows, optimal setting of the general programme parameters) as well as the anti-virus software of the used operating system (eg, definition of the UNIFIT software as excluded process).

5.2 | Spike correction setting

We have tested the limits of the threshold multiplier and the number of spikes. After the tests, we found and defined minimum and maximum values (also integrated in the UNIFIT software):

1. Threshold multiplier: minimum: 5; recommended: 20; maximum: 100.
2. Maximal number of spikes (fraction of total number of data points): minimum: 5%; recommended: 10%; maximum: 33%.

6 | FINAL DISCUSSION AND CONCLUSIONS

In this paper, a method for automatic spike correction in large data sets of spectra such as time-dependent XPS series or Raman scans is presented. The spike finding and correction by interpolated values

are included. The modified Laplace operator method is used. All nine different conditions of the calculation of the Laplace operator, depending on the spike position in the spectra, were discussed. The higher sharpness of the spikes of the Laplace spectra with respect to the original spectra was respected. Not shown here, the tests of an automatic estimated threshold[9] could not deliver satisfactory results. Therefore, we propose a manual setting of the threshold multiplier. The calculation time of the spike correction of a large number of spectra can be reduced by the optimal setting of the maximal number of spikes. For a better understanding of the method, the operator can optionally plot the calculated Laplace spectra.

The new algorithm was tested using test spectra and real Raman and XP spectra. The validation of the spike correction results of the test spectra using the residuals show a mainly statistic behaviour of the new calculated intensities at the spike positions. At the original spike positions of the spike-corrected real Raman and XP spectra, no untypical behaviour of the intensities could be observed.

ORCID

Ronald Hesse https://orcid.org/0000-0001-8631-6687
Carsten Bundesmann https://orcid.org/0000-0003-0599-5038
Reinhard Denecke https://orcid.org/0000-0003-1065-5791

REFERENCES

1. Hesse R, Chassé T, Szargan R. Peak shape analysis of core level photoelectron spectra using UNIFIT for WINDOWS. *Fresenius J Anal Chem*. 1999;365(1-3):48-54.
2. Hesse R, Chassé T, Streubel R, Szargan R. Error estimation in peak-shape analysis of XPS core-level spectra using UNIFIT 2003: how significant are the results of peak fits? *Surf Interface Anal*. 2004;36 (10):1373-1383.
3. Hesse R, Streubel R, Szargan R. Product or sum: comparative tests of Voigt, and product or sum of Gaussian and Lorentzian functions in the fitting of synthetic Voigt-based X-ray photoelectron spectra. *Surf Interface Anal*. 2007;39(5):381-391.
4. Hesse R, Denecke R. Improved Tougaard background calculation by introduction of fittable parameters for the inelastic electron scattering cross-section in the peak fit of photoelectron spectra with UNIFIT 2011. *Surf Interface Anal*. 2011;43(12):1514-1526.
5. Hesse R, Streubel R, Szargan R. Improved accuracy of quantitative XPS analysis using predetermined spectrometer transmission functions with UNIFIT 2004. *Surf Interface Anal*. 2005;37(7):589-607.
6. Hesse R, Weiß R, Streubel R, Szargan R, Denecke R. Improved peak-fit procedure for XPS measurements of inhomogeneous samples—Development of the advanced Tougaard background method. *J Electr Spectrosc Related Phenomena*. 2015;205:29-51.
7. Mroczkowski S, Lichtman D. Quantitative AES of binary alloys: A comparison between handbook and pseudo-first principles corrections. *Surf Sci*. 1983;127(1):119-134.
8. Outka DA, Stöhr J, Rabe JP, Swalen JD. The orientation of Langmuir–Blodgett monolayers using NEXAFS. *J Chem Phys*. 1988;88 (6):4076-4087.
9. http://rruff.info/
10. Hill W, Rogalla D. Spike-correction of weak signals from charge-coupled devices and its application to Raman spectroscopy. *Anal Chem*. 1992;64(21):2575-2579.
11. Tian Y, Burch KS. Automatic Spike Removal Algorithm for Raman Spectra. *Appl Spectrosc*. 2016;70(11):1861-1871.
12. Ehrentreich F, Sümmchen L. Spike Removal and Denoising of Raman Spectra by Wavelet Transform Methods. *Anal Chem*. 2001;73 (17):4364-4373.
13. Whitaker DA, Hayes K. A simple algorithm for despiking Raman spectra. *Chemom Intel Lab Syst*. 2018;179:82-84.
14. Phillips GR, Harris JM. Polynomial filters for data sets with outlying or missing observations: application to charge-coupled-device-detected Raman spectra contaminated by cosmic rays. *Anal Chem*. 1990;62 (21):2351-2357.
15. Georg Schulze H, Turner RFB. A Fast, Automated, Polynomial-Based Cosmic Ray Spike–Removal Method for the High-Throughput Processing of Raman Spectra. *Appl Spectrosc*. 2013;67(4):457-462.
16. Ryabchykov O, Bocklitz T, Ramoji A, et al. Automatization of spike correction in Raman spectra of biological samples. *Chemom Intel Lab Syst*. 2016;155:1-6.
17. Damen TC, Porto SPS, Tell B. Raman Effect in Zinc Oxide. *Phys Rev*. 1966;142(2):570-574.
18. Ashkenov N, Mbenkum BN, Bundesmann C, et al. Infrared dielectric functions and phonon modes of high-quality ZnO films. *J Appl Phys*. 2003;93(1):126-133.
19. Samanta K, Bhattacharya P, Katiyar RS. Temperature dependentE2Raman modes in theZnCoOternary alloy. *Phys Rev B*. 2007;75(3):035208.
20. Bundesmann C, Schmidt-Grund R, Schubert M, Optical properties of ZnO and related compounds; In *"Transparent Conductive Zinc Oxide"*; K. Ellmer, A. Klein, B. Rech (Eds.), New York Springer, 2008. ISBN: 978-3540736110,

How to cite this article: Hesse R, Bundesmann C, Denecke R. Automatic spike correction using UNIFIT 2020. *Surf Interface Anal*. 2019;51:1342–1350. https://doi.org/10.1002/sia.6702