

R. Hesse · T. Chassé · R. Szargan

Unifit 2002—universal analysis software for photoelectron spectra

Received: 16 September 2002 / Accepted: 15 November 2002 / Published online: 29 January 2003

© Springer-Verlag 2003

Abstract The current level of development of the spectra analysis software Unifit for Windows is presented and evaluated by checking with test spectra. The program is characterized and the correctness of the numerical routines is demonstrated for the particular cases of the Shirley type background model and the Gaussian–Lorentzian product model functions. Different approaches to an optimum fit result are suggested. A convenient analysis of the assessment of the peak fit procedure is proposed. All results are presented in tabular form too, to make the data more comprehensible.

Keywords Unifit for Windows · Spectrum analysis software · Photoelectron spectroscopy · Shirley type background model · Gaussian–Lorentzian product model

Introduction

The demands on software for analysis of photoelectron spectra become more and more complex. An adequate theoretical modelling and accurate treatment of the numerical problems has to be combined with intuitive and convenient handling. The development of new or improved codes proceeds continuously. Nevertheless, careful checking and testing of the code remain a challenge. There are two typical approaches for testing the applicability of the software for the analysis. The results and the performance of the software may be checked by analysing synthesized spectra with well-known parameters for the model peaks. The comparison of the ideal input parameters and the result of the evaluation of the parameters by

application of the code are straightforward and thus problems may be identified easily. However, a variety of such model spectra covering quite different possible cases of complexity should be analysed. On the other hand, it is strongly advised that developers may test their code by analysing test spectra available from key references or provided by institutions like the NIST.

During recent years we have developed the software UNIFIT based on 32-bit code for the analysis of core level photoemission spectra. It permits the fit of the spectra by using Gaussian–Lorentzian product or convolution-type functions as well as several different background models. Excellent convergence by minimizing chi-squared and numerically correct calculations were demonstrated for synthesized model peaks. Here we will further characterize the software and demonstrate the performance by analysing test spectra recently suggested by Seah and Brown [1].

Characterisation of the UNIFIT code

Version 2002 of “Unifit for Windows” manifests another step from simple peak-fit software towards a more universal spectrum treatment and analysis software for photoemission spectra. The 32-bit code of the software runs under Windows 95/98, Windows NT and Windows 2000. The requirements for running the software successfully are a computer of Pentium or comparable level and at least 32-Mbyte RAM.

The program permits the analysis of up to 30 spectra in separate windows at once. The spectrum information is available to the user in the window titles.

A presetting option allows the user to choose, the program language (German or English), the presentation of the data (kinetic or binding energies, etc.), type and colour of printout, export options, number of points (for, e.g., differentiation) and features of the peak models to be used (convolution or multiplicative coupling-type model functions) and the type of fit parameters (absolute, relative). Five sets of parameters for calculation of satellite-induced

R. Hesse (✉) · R. Szargan
Wilhelm-Ostwald-Institute for Physical and Theoretical Chemistry, University of Leipzig, Linnéstr. 2, 04103 Leipzig, Germany
e-mail: rhesse@rz.uni-leipzig.de

T. Chassé
Institute for Surface Modification, Permoserstrasse 15, 04303 Leipzig, Germany

background are provided, which may be further edited by the user according to his requirements. A dynamic toolbar may be adapted to the needs of the current application.

An input routine permits the treatment of measured data of different origin (e.g. ESCALAB 210/220, PHI 5400/5600, VGX-900, VAMAS,...). The simplest kind of spectrum to be accepted may be represented by an energy column and a related intensity column, independent on direction of scan and kinetic/binding energy.

Measured data in VAMAS- or VGX-900 format may be represented as parameter dependent series. Up to 720 Spectra may be imported together. Stored projects may be imported without using the stored data during measurement.

An export routine ensures quick data transfer to graphics software like Origin or Excel for presentation of the results. Unifit2002 permits a summary of the current stage of data processing in all the open windows into a project. This is very useful for storing the measurement data as well as the results of the analysis conveniently. The printout includes the fit parameter and the residuals, automatically. Measurement parameters may be optionally included, too. A quick presentation of the data obtained may be performed by storing the current window as a bitmap. In this manner the results may be transferred to other presentation programs or documents conveniently.

In a data treatment part the charging correction of the spectra, several kinds of background calculation, satellite subtraction, smoothing, differentiation, or single data correction may be performed. Copy/paste function is available to save or insert results of a partial data treatment procedure temporarily.

The implemented chi-squared minimization procedure proceeds according to the algorithm of Marquardt–Levenberg. The components may be calculated either using a multiplicative coupling of Gaussian and Lorentzian functions or represented by a convolution. The fit parameters may be chosen as absolute or relative ones, the latter in comparison to a leading peak. The peak parameters may be fixed, let free or constrained within a min/max range. The background may be composed of contribution of a third polynomial in energy, a Shirley- or a Tougaard-type contribution. The fit residuals are routinely calculated and presented at any stage of the fit.

A serial treatment of spectra is adapted in order to support the quick analysis of, e.g., depth profile or angular dependent measurement of peaks. All options of the data treatment of a parent peak may be integrated into the treatment of other series of spectra.

The quantitative analysis may be performed within the homogeneous mixing model framework using the overview spectra or by analysis of single spectra. The sensitivity factors may be chosen from the commonly accepted data sets of Wagner or Scofield, which are available in the program, or from user-specified values. The peak components determined from peak fits may be taken into account in the quantitative analysis.

The following data sets have been integrated within the Unifit2002 software:

1. line energies of photoelectrons of elements and chemical shifts [2, 3, 4],
2. Auger parameters of chemical compounds [2, 4],
3. 14 test spectra,
4. several different sets of sensitivity factors (see above),
5. energy separations and intensity ratios of doublet lines.

Database 1 is to support identification and lettering of the spectra. Database 2 may be presented as a Wagner-type plot and may also be easily lettered. The user may edit both databases in a convenient way. The 14 test spectra of database 3 permit the user to check the correctness of different program routines. Database 4 is essential for the quantitative analysis. Database 5 supports the user in choosing the initial parameters for the fit of doublet lines.

At any stage the information on measurement and treatment of the data is accessible to the user. Important comments may be inserted in the presentation of the spectra on screen or in the output.

Program testing with synthesized spectra

An evaluation of fit programs for photoemission spectra has been performed by Seah and Brown [1], recently. According to the tested programs additive and multiplicative coupling of Gaussian and Lorentzian peaks were checked. The details and particular problems of the different codes are discussed in Ref. [1].

Here we will use the same set of test functions or test spectra to evaluate the performance of the UNIFIT software. We focus on the multiplicative coupling for determination of the model functions for the fit. While SBF (Shirley-Background-Function) does not represent any real spectrum, the other synthesized spectra have been established in order to model the following polymer samples [3]:

- PVC: poly(vinyl chloride);
- PMMA: poly(methyl methacrylate);
- PVAc: poly(vinyl acetate); and
- PIB: polyisobutylene.

Throughout this paper we will use the short notations only. The parameters characterising these five test spectra are provided in Table 1. Generally, a peak width (FWHM, full width half maximum) of 1.00 eV and a Gaussian/Lorentzian mixing parameter of 0.50 (except for SBF, G/L: 0.25) were adopted. The energy step width was 0.10 eV. The following tests were performed:

1. acceptance of VAMAS format;
2. correctness of the generation of the model functions;
3. effectiveness of the Shirley-type background fit; and
4. determination of peak separation in strongly overlapping peak structures.

First, a correct I/O of the VAMAS formats in different presentations of the spectra could be demonstrated.

The rate of convergence and the results of the iterative chi-squared minimization may depend on the starting pa-

Table 1 Binding energies and relative intensities of the single peaks applied to generate the test spectra used (from Ref. [4])

Spectrum	Parameter	Peak number			
		1	2	3	4
SBF	Energy (eV)	487.60	486.60		
	Intensity (%)	50	50		
PVC	Energy (eV)	287.00	285.90		
	Intensity (%)	50.00	50.00		
PMMA	Energy (eV)	289.00	286.80	285.70	285.00
	Intensity (%)	16.80	20.80	20.80	41.60
PVAc	Energy (eV)	289.20	286.60	285.50	285.00
	Intensity (%)	24.00	25.00	25.00	26.00
PIB	Energy (eV)	285.60	285.20	285.00	
	Intensity (%)	25.00	50.00	25.00	

rameters sensitively. Thus, we have started the fit procedures with initial values at least about 20% from the ideal values for the parameters intensity, G/L-mixing, and FWHM. The peak energies were shifted by 0.5 eV. All asymmetry parameters were set to zero and kept fixed. No

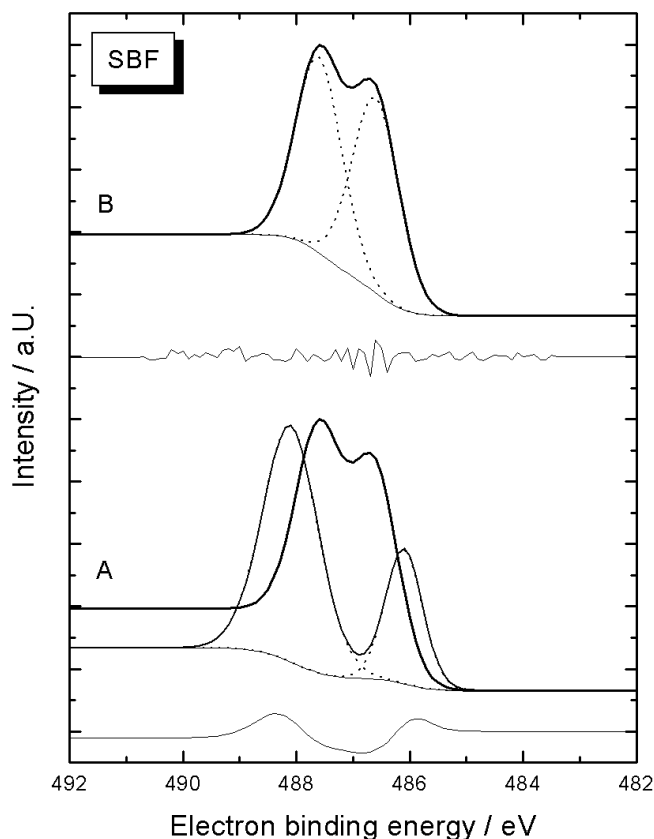


Fig. 1 Fit of the test function SBF including background fit. *A*: Test spectrum and result using the initial parameter values. *B*: Test spectrum and result of the fitting procedure. *Solid line*, component sum curve; *heavy solid line*, test spectrum; *dotted line*, components. Residuals are provided below the curves of spectrum and components

coupling of parameters was set. The background parameters Shirley background intensity and constant background intensity were determined together with the parameters of the model peaks in the fit (except for SBF, method 1). The correct number of single peaks was preset.

As in most peak-fit routines in Unifit 2002 chi-squared is calculated to provide the quantity which measures the error of the fit or the deviation of the original data and the fit function. chi-squared is minimized during the iteration procedure. The expectation value of the reduced chi-squared should decrease towards or at least approach unity in case of a good fit to spectra possessing random noise. However, the present test spectra were analysed as synthesized, i.e. no noise-like contributions were added. Therefore, in case of ideal matching the error of the fit should be mostly of numerical origin and actually approach zero. The so called Abbe criterion provides information on the appearance of correlation of neighbored residuals and thus on possible shortcomings of the fit. The expectation value of the Abbe criterion is unity in case of a good fit with statistically uncorrelated residuals as, e.g., present if the residuals are governed by random noise of the data only. The Abbe criterion should approach zero for strong correlations among neighbored residuals indicating bad fits. Anticorrelation in the residuals as may be anticipated for mostly numerical origin of the residuals is indicated by Abbe values exceeding unity. This should be considered during the discussion of the following results.

Fit of the SBF spectrum

The test spectrum SBF was used to evaluate the effectiveness of the Shirley-background calculation. The function SBF was generated from two model peaks with a FWHM of 1.00 eV, a G/L-mixing of 0.25, and a peak separation of 1.00 eV (see Fig. 1a). The background intensity was coupled to the peak areas per channel at higher energies by a factor of 0.20. This leads to a large background contribution in the test spectrum.

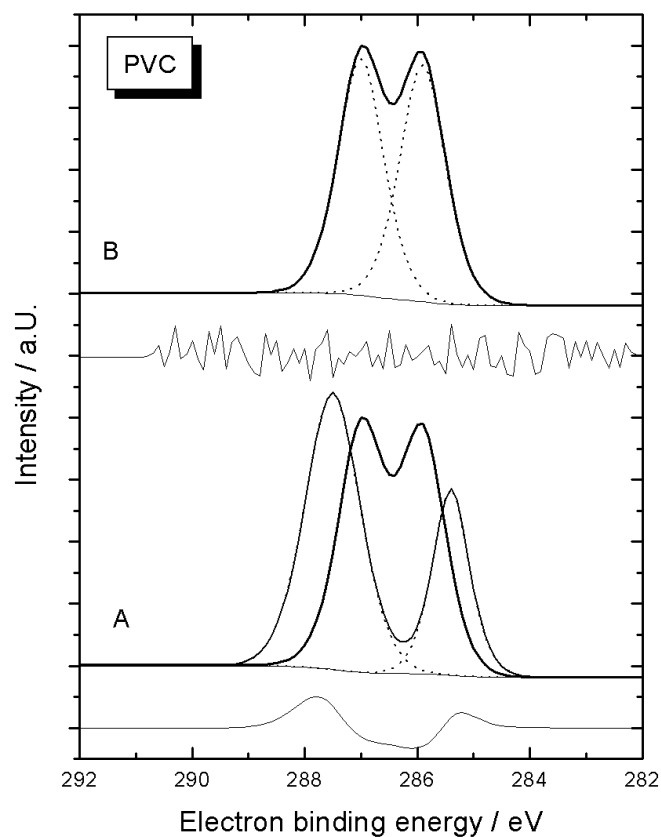
Two basically different approaches may be chosen in Unifit 2002 to calculate background spectral shape. First,

Table 2 Parameters of the model functions for synthesis of the test spectrum SBF before and after the minimization procedure. Parameters shown represent the joint minimization procedure for peak and background

Parameter	Model function			
	Initial parameter		Optimised parameter	
	Peak 1	Peak 2	Peak 1	Peak 2
Height (counts)	116000	66000	96075	96075
G/L ratio	0.30	0.20	0.25	0.25
Energy (eV)	488.10	486.10	487.60	486.60
FWHM (eV)	1.20	0.80	1.00	1.00
Rel. intensity (%)	72.51	27.49	50.00	50.00
Error	>17000		0	
Abbe criterion	0.026		1.310	

Table 3 Comparison of fit parameters of the test function SBF obtained with either background fit or background subtraction

Quantity	Background fit	Background subtraction
Error	0	0
Abbe criterion	1.310	1.283
Height peak1	96075	97036
Height peak2	96075	97036
Intensity peak1	100006	101006
Intensity peak2	100006	101006

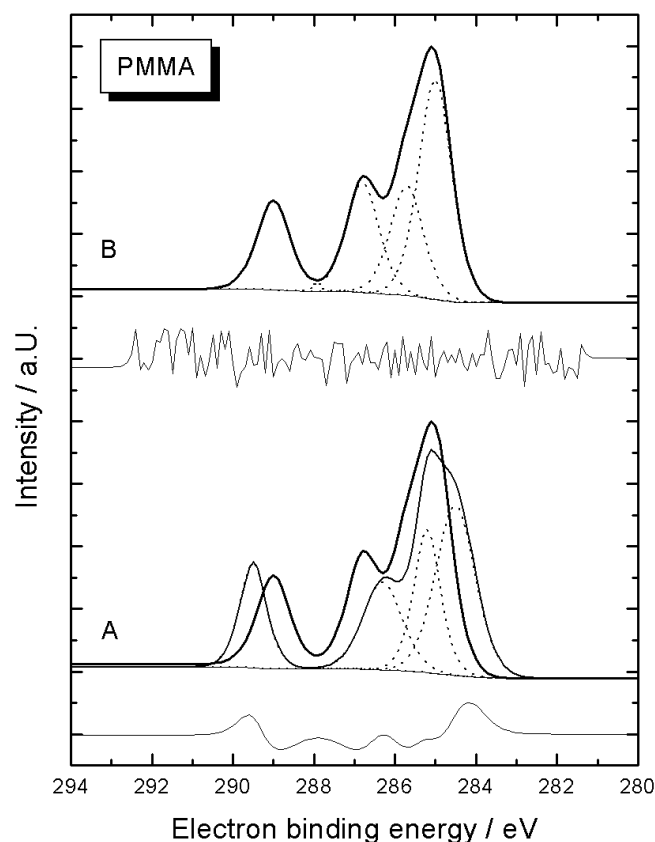
**Fig. 2** Fit of the test function PVC including background fit. A: Test spectrum and result using the initial parameter values. B: Test spectrum and result of the fitting procedure. *Solid line*, component sum curve; *heavy solid line*, test spectrum; *dotted line*, components. Residuals are provided below the curves of spectrum and components

the background may be estimated and then subtracted from the measured spectrum before running the fit procedure. The Shirley background model requires the solution of an integral equation, which is performed iteratively. Nine iterative steps were sufficient to ensure convergence at a stable background peak shape, which was then removed from the original data. The fit procedure was then run with the difference spectrum, which should represent the peaks only, according to the background model.

Alternatively, a free parameter for the constant background and another one for the Shirley background con-

Table 4 Initial and optimised parameters of the model functions representing the test spectrum PVC

Parameter	Model function			
	Initial parameters		Optimised parameters	
	Peak 1	Peak 2	Peak 1	Peak 2
Height (counts)	55000	37000	47082	47083
G/L ratio	0.40	0.60	0.50	0.50
Energy (eV)	287.50	285.40	287.00	285.90
FWHM (eV)	1.20	0.80	1.00	1.00
Rel. intensity (%)	68.35	31.65	50.00	50.00
Error	>7000		0	
Abbe criterion	0.028		1.217	

**Fig. 3** Fit of the test function PMMA including background fit. A: Test spectrum and result using the initial parameter values. B: Test spectrum and result of the fitting procedure. *Solid line*, component sum curve; *heavy solid line*, test spectrum; *dotted line*, components. Residuals are provided below the curves of spectrum and components

tribution were determined in a joint minimization procedure in parallel with the free peak shape parameters (Fig. 1b). Both methods converge to an error indicator of (nearly) zero and lead to an excellent approach to the ideal parameters (Table 2). This clearly proves that the global minimum was reached by the procedure. Note that the residuals in Fig. 1a are mainly determined by the bad representation of the spectra by the initial data, while the

Table 5 Initial and optimised parameters of the model functions representing the test spectrum PMMA

Parameter	Model function							
	Initial parameters				Optimised parameters			
	Peak 1	Peak 2	Peak 3	Peak 4	Peak 1	Peak 2	Peak 3	Peak 4
Height (counts)	19300	15700	25700	30500	16008	19775	19775	39549
G/L ratio	0.60	0.40	0.60	0.40	0.50	0.50	0.50	0.50
Energy (eV)	289.50	286.30	285.20	284.50	289.00	286.80	285.70	285.00
FWHM (eV)	0.80	1.20	0.80	1.20	1.00	1.00	1.00	1.00
Rel. intensity (%)	16.99	20.40	22.98	39.63	16.83	20.79	20.79	41.59
Error	>1900				0			
Abbe criterion	0.033				1.216			

residuals of Fig. 1b are strongly enhanced and represent remaining deviations due to numerical effects, mostly. However, the peak heights and peak intensities differ by 1% between these two methods (Table 3) and the difference in the Abbe criterion of about 2% indicates some difference, too. A possible explanation is that the iterative background procedure might not have completely converged before background subtraction.

Nevertheless, for the following tests the background fit procedure was applied, generally.

Fit of the test spectrum PVC

This test spectrum has been synthesized from two well-separated single lines, which may be clearly distinguished due to the valley between the maxima, as may be judged from Fig. 2. The Shirley background intensity is much weaker than in SBF. Therefore, a straightforward convergence of the fit procedure was to be expected.

A negligibly small error and a value of the Abbe criterion close to the expectation value of 1 indicate successful modelling by the optimised parameters. The optimised parameters in Table 4 reproduce the ideal ones in the leading digits excellently, as demonstrated by comparison to Table 1 (see also Fig. 2). The fit procedure employing the background fit together with peak parameter fit converged quickly to the global minimum (only 12 iteration cycles).

Fit of the test spectrum PMMA

In contrast to the two-peak spectrum PVC the test spectrum PMMA was composed of four single peaks. Two of the model peaks in PMMA are very close and may not at all be resolved by eye in Fig. 3. However, a shoulder at about 285 eV may be clearly distinguished.

Even this more complex test spectrum could be modelled well. Again, an error of nearly zero and an Abbe criterion of nearly unity were obtained. The optimised model peak parameters provided in Table 5 do not exhibit significant deviations from the ideal ones (Table 1) within the leading digits shown. A reasonably fast convergence within 22 iterations was achieved using the chosen starting parameters.

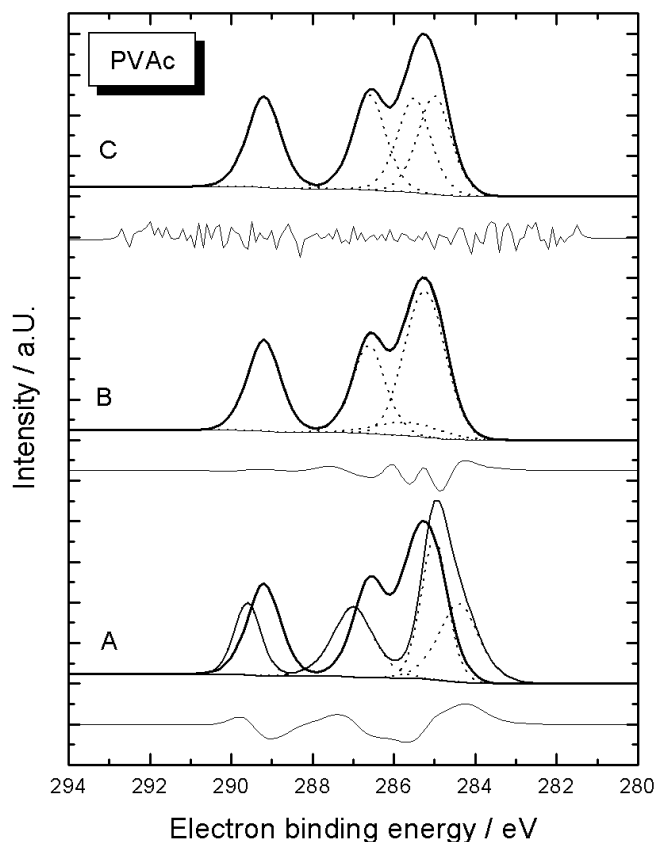


Fig. 4 Fit of the test function PVAc including background fit. *A*: Test spectrum and result using the initial parameter values. *B*: Test spectrum and result of a first fitting procedure. *C*: Test spectrum and result of a second fitting procedure starting with the optimised parameters from procedure *B* modifying only on peak height value. *Solid line*, component sum curve; *heavy solid line*, test spectrum; *dotted line*, components. Residuals are provided below the curves of spectrum and components

Fit of the test spectrum PVAc

Similar to spectrum PMMA the test spectrum PVAc was constructed using four single peaks [1]. But the different distribution of binding energies and relative intensities nevertheless result in a quite different spectral shape as shown in Fig. 4. Thus, peaks 3 and 4 of spectrum PVAc are hardly recognizable. The ideal values for the relative

Table 6 Initial and optimised parameters of the model functions representing the test spectrum PVAc

Parameter	Model function							
	Initial parameters				Optimised parameters			
	Peak 1	Peak 2	Peak 3	Peak 4	Peak 1	Peak 2	Peak 3	Peak 4
Height (counts)	18000	17500	34800	19500	22620	22061	3356	36431
G/L ratio	0.40	0.60	0.40	0.60	0.49	0.53	0.20	0.20
Energy (eV)	289.60	287.00	285.00	284.40	289.20	286.61	285.64	285.23
FWHM (eV)	0.80	1.20	0.80	1.20	0.99	0.98	1.92	1.21
Rel. intensity (%)	16.34	24.62	31.61	27.43	24.08	23.04	7.50	45.38
Error	>2500				0.943			
Abbe criterion	0.026				0.125			

intensities and the peak positions of the individual peaks used to synthesize the spectrum may be compared in Table 1.

As described above, rather bad initialisation values were chosen for the peak parameters. In this case the code converged to the parameter set of Table 6, which resulted in a low, but non-zero, error. From visual inspection of the fit results in Fig. 4 no conclusive statement can be derived on shortcomings of the fit. On the other hand systematic variations of the residuals demonstrate less than optimum representation of the spectrum by the fit results obtained so far. The Abbe criterion deviates much from an expectation value of unity (Table 6), and thus indicates insufficient modelling of the test spectrum by applying the obtained fit parameters for the model peaks, too. In this test case (Table 6) the fit procedure could not come close to the ideal parameters even after a huge number of 1000 iteration cycles, demonstrating that convergence in a local minimum happened. Indeed, after distorting the set of parameters by artificially increasing the height of peak 3 from 3356 to 20000 by operator manipulation the minimization procedure left the region of the local minimum it had approached. Subsequently the fit arrived at a solution characterized by parameter values very close to the ideal ones by approaching the global minimum of chi-squared. Further, not only the error but also the Abbe criterion displayed significantly improved values, too.

Fit of the test spectrum PIB

The test spectrum PIB is the most complex one considered here [1]. Three single lines positioned very close to each other generate a broad structure with no evidence for the presence of individual peaks (Fig. 5).

Due to this complexity an even higher sensitivity to the initialisation values of the fit procedure was to be expected from the beginning and was in fact observed. A decent approach to the test spectrum was already obtained after 22 iterations (Fig. 5b), but systematic deviations in the residual curve indicated significant shortcoming of this fit. However, a very good approach to the ideal parameters was finally achieved (Table 7) but only at the expense of an extremely huge number of >4000 iteration cycles, which would not be accepted by any routine fit pro-

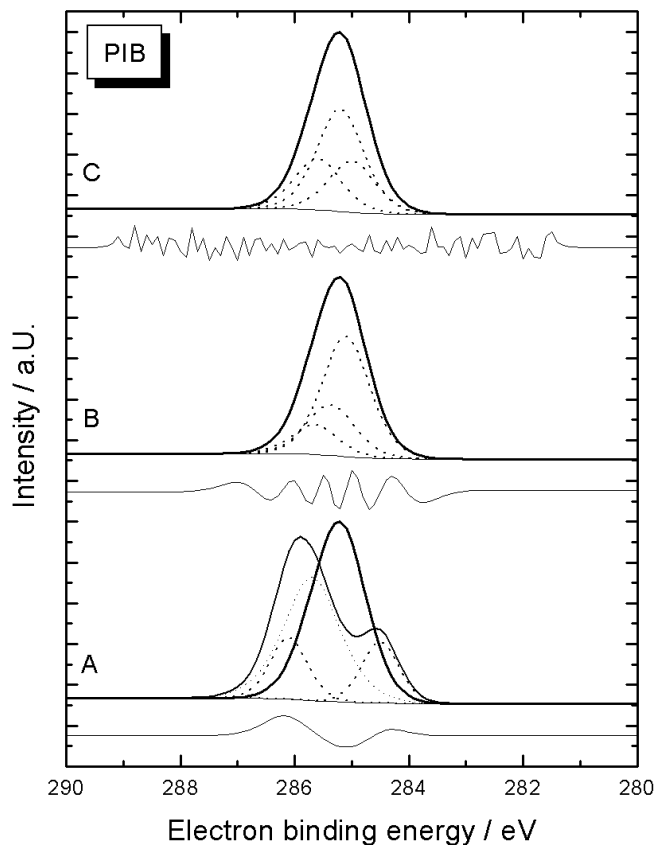


Fig. 5 Fit of the test function PIB including background fit. *A*: Test spectrum and result using the initial parameter values. *B*: Test spectrum and result of a fitting procedure after 22 iterations. *C*: Test spectrum and result of the continued fitting procedure after >4000 iterations (see text). *Solid line*, component sum curve; *heavy solid line*, test spectrum; *dotted line*, components. Residuals are provided below the curves of spectrum and components

cedure for photoemission spectra, typically. Anyway, it is interesting to note that in the considered case no interaction by the operator was necessary to approach the global minimum (except for restarting the procedure after the maximum iteration number was reached), much in contrast to the former example of the fit to PVAc. But, as mentioned above, we arrived at different results after very minor modification of the initial parameter set of Table 8

Table 7 New set of initial and optimised parameters of the model functions representing the test spectrum PVAc following interaction by the operator

Parameter	Model function							
	Second set of parameters				Optimised parameters			
	Peak 1	Peak 2	Peak 3	Peak 4	Peak 1	Peak 2	Peak 3	Peak 4
Height (counts)	22620	22061	20000	36431	22600	23541	23541	24483
G/L ratio	0.49	0.53	0.20	0.20	0.50	0.50	0.50	0.50
Energy (eV)	289.20	286.61	285.64	285.23	289.20	286.60	285.50	285.00
FWHM (eV)	0.99	0.98	1.92	1.21	1.00	1.00	1.00	1.00
Rel. intensity (%)	18.00	17.37	30.24	34.39	24.00	25.00	25.00	26.00
Error	742.9				0			
Abbe criterion	0.004				1.070			

Table 8 Initial and optimised parameters of the model functions representing the test spectrum PIB as chosen and obtained for the particular fit run including 4000 iteration cycles. The interval width for calculations was 291.7–278.6 eV

Parameter	Model function					
	Initial parameters			Optimised parameters		
	Peak 1	Peak 2	Peak 3	Peak 1	Peak 2	Peak 3
Height (counts)	28300	56400	28300	13541	47082	23541
G/L ratio	0.40	0.60	0.40	0.50	0.50	0.50
Energy (eV)	286.10	285.10	284.50	285.60	285.20	285.00
FWHM (eV)	0.80	1.20	1.20	1.00	1.00	1.00
Rel. intensity (%)	19.66	60.69	19.66	25.00	50.00	25.00
Error	>6500			0		
Abbe criterion	0.041			1.328		

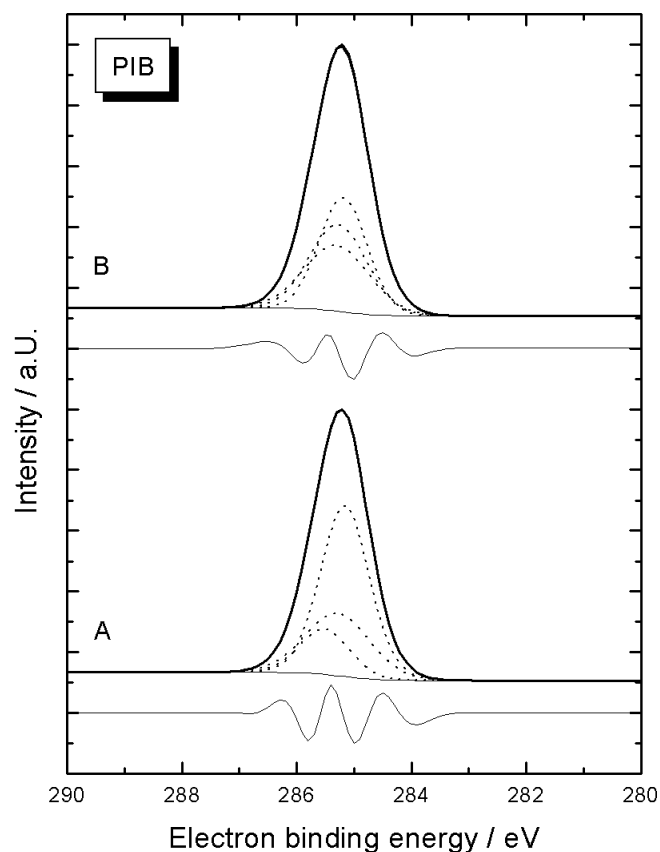


Fig. 6 Fit of the test function PIB including background fit. Demonstration of the effects of slight modification of the initial parameters leading to worse results, compared to the situation in Fig. 5B (see text). *Solid line*, component sum curve; *heavy solid line*, test spectrum; *dotted line*, components. Residuals are provided below the curves of spectrum and components

as is shown in Fig. 6. The problems of these results are again displayed by significant systematic variation in the residuals. It turned out that the width of the energy region used for these calculations (here 291.7–278.6 eV) contributed to the fit result to some extent. In cases of the fits presented in Fig. 6 the error was also close to zero, but values of the Abbe criterion <0.7 indicate some shortcomings of the spectrum description using the current parameter set, in line with the conclusions from the residuals. In case of this complex test spectrum we may conclude that finding the global minimum is not straightforward at all and requires input from as much pre-knowledge as available and very much numerical effort.

Summary

Extending earlier investigations on the correct calculation of Gaussian–Lorentzian convolution functions, good convergence behaviour and correct identification of parameters of simple test functions using the Unifit software [5], here we have shown the accurate generation and treatment of the product model function. The correct acceptance of any VAMAS formats as well as correct calculation of the Shirley-Background model were confirmed, too.

The test spectra SBF, PVC and PMMA could be analysed without problems using Unifit 2002. The global minima of chi-squared were determined quickly. From the fits optimised peak parameters were obtained in excellent agreement with the ideal values from the peak synthesis.

In cases of the more complex and less well-structured test spectra PVAc and PIB problems occurred in finding the global chi-squared minimum. This is due to the general feature of the Marquardt–Levenberg algorithm, which

converges globally, but does not necessarily find the absolute or global minimum. We have demonstrated two simple approaches to overcome these general problems:

1. increasing the number of iteration cycles in order to approach the global minimum from a range of slow convergence; and
2. distorting the parameter set by modification of individual parameters in order to exit the convergence radius of a local minimum towards the global minimum.

The latter examples demonstrated that the evaluation of the fit result cannot be performed on the basis of a chi-squared minimum alone. Statistically well-distributed residuals in case of noise are another significant indicator for a reasonable fit. Similarly, the numerical value of the so called Abbe criterion should be checked, because a significant deviation from the expectation value unity in case of

random noise will indicate shortcomings of the model, too.

References

1. Seah MP, Brown MT (1998) *J Electron Spectrosc Related Phenom* 95:71
2. Wagner CD, Naumkin AV, Kraut-Vass A, Allison JW, Powell CJ, Rumble Jr JR (2000) NIST X-ray photoelectron spectroscopy database, NIST standard reference database 20, Version 3.1, Gaithersburg, <http://srdata.nist.gov/xps>
3. Beamson G, Briggs D (1992) *High resolution XPS of organic polymers*. Wiley, Chichester
4. Moulder JF, Stickle WF, Sobol PE, Bomben KD (1995) *Handbook of X-ray photoelectron spectroscopy*. Physical Electronics, Eden Prairie
5. Hesse R, Chassé T, Szargan R (1999) *Fresenius J Anal Chem* 365:48