Practical Error Analysis of Fit Parameters with UNIFIT 2003 - How Good is the Accuracy of the Results of Peak Fits?

Hesse, R., Chassé, T.*, Streubel, P., Szargan, R.

Wilhelm-Ostwald-Institute for Physical and Theoretical Chemistry, University of Leipzig, D-04103 Leipzig
* Institute for Physical and Theoretical Chemistry, University of Tübingen, D-72076 Tübingen
Website: www.uni-leipzig.de/~unifit Contact: rhesse@rz.uni-leipzig.de

1. Motivation

Results from X-ray photoelectron spectra are affected by a rather large number of uncertainties involved in the process of measuring and the treatment of the spectra.

In this paper we present the theoretical background and illuminate examples of the evaluation of the peak fit parameter uncertainties. Finally we give recommendations that might be helpful to minimise these errors.

2. Theoretical base

The program code UNIFIT offers two ways to create a model function to describe an XPS core level spectrum M(i): product and convolution of Gauss-Lorentz function [1, 2]. Each component of the model function is defined with 5 parameters p_k (Product: intensity, G/L-ratio, energy, FWHM, asymmetry, Convolution: intensity, FWHM-Gaussian line, energy, FWHM-Lorentzian line, asymmetry).

With the approximation that $\chi^2(p_k)$ is close to a parabolic form around the minimum, the program UNIFIT 2003 uses the following definition from [3] for the calculation of the fit parameter error Δp_k :

1. We change one parameter p_k by an amount Δp_k ,
2. we optimise all other parameters $p_{j\neq k}$ to minimize χ^2, then the new value of χ^2 will be 1 greater than the old value.

$$\chi^2(p_k + \Delta p_k) = \chi^2(p_k) + 1 \qquad (1)$$

with

$$\chi^{2*} \cong 1 \qquad (2)$$

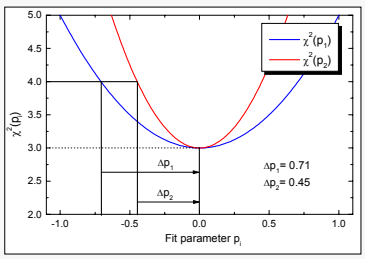

Fig. 1. Two functions $\chi^2(p_1)$ and $\chi^2(p_2)$ in one dimension and the corresponding errors Δp_1 and Δp_2; values of the x- and y-axis arbitrary

By expanding the left side of (1) to second order in a Taylor's series as a function of the parameter p_k , we get

$$\chi^2(p_k + \Delta p_k) = \chi^2(p_k) + \left.\frac{\partial \chi^2}{\partial p_k}\right|_{p_{k0}} \Delta p_k + \frac{1}{2}\left.\frac{\partial^2 \chi^2}{\partial p_k^2}\right|_{p_{k0}} (\Delta p_k)^2$$

The optimum values for the parameters p_k are those for which the function χ^2 is a minimum in parameter space, i.e. for which the derivatives with respect to the parameters are 0. Together with the right side of (1) we obtain Eq. (4),

$$\chi^2(p_k) + 1 = \chi^2(p_k) + \frac{1}{2}\left.\frac{\partial^2 \chi^2}{\partial p_k^2}\right|_{p_{k0}} (\Delta p_k)^2 \qquad (4)$$

and at p_{k0}

$$\Delta p_k = \sqrt{\frac{2}{\left(\frac{\partial^2 \chi^2}{\partial p_k^2}\right)}} \qquad (5)$$

Δp_k can be calculated iteratively as well as with the curvature matrix **H**. Is

$$B = H^{-1} \qquad (6)$$

and b_{kk} the diagonal terms of **B**, the result is.

$$\Delta p_k = \sqrt{2 \cdot b_{kk}} \qquad (7)$$

3. Analysis of test spectra

The following parameters were used for 7 test spectra fitted with the same starting conditions:

Model function: Product of Gauss-Lorentz Lines
Number of components: 2
Intensity of both components: 10000 Counts
G/L-ratio: 0.5
FWHM: 2 eV
Asymmetry: 0, fixed
Background: Constant, 1000 Counts
Line positions of component 1: 14.3 eV - 14.9 eV
Line positions of component 2: 15.7eV - 15.1 eV
The spectra were modified by a statistical noise.

Result:
The uncertainty of the fit parameters increases with decreasing distance between both peaks. The errors of the fit parameters exceed 100% if the distance between the lines is less than 20% of FWHM of the peaks!

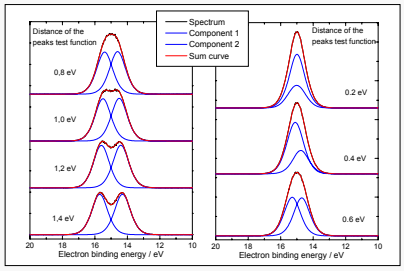

Fig.2. Peak fits for test spectra; the distance of the peaks is reduced stepwise from 1.4 eV to 0.2 eV

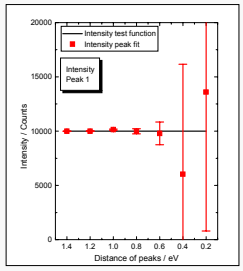

Fig.3. Illustration of the calculated intensities of the 1st peak with error bars

4. Influence of certain knowledge

In order to illustrate the influence of reducing the number of free fitting parameters an S 2p-spectrum of two doublets was fitted with 19 free parameters (Test B) and with frozen known values of the doublets (Test A):

Background: Polynomial 1st order + Tougaard, calculated during peak fitting
Asymmetry: 0, fixed

Test A: Intensity ratio doublet-lines: 1:2, fixed
Peak distance doublet-line: 1.2 eV, fixed
G/L-ratio doublet lines: 1, fixed
Number of free fit parameters: 13

Test B: All fit parameters variable
Number of free fit parameters: 19

Tab.1. Comparison of results of two different fits with (Test A) and without (Test B) frozen doublet parameters

	Test A		Test B	
	Energy/eV	Area/cps*eV	Energy/eV	Area/cps*eV
Peak 1	163.88 ± 0.01	18456 ± 6 %	163.70 ± 0.03	4161 ± 20 %
Peak 2	165.08 ± 0.01	9758 ± 8 %	164.41 ± 0.52	24537 ± 100 %
Peak 3	168.08 ± 0.03	3390 ± 21 %	167.94 ± 0.16	2262 ± 109 %
Peak 4	169.28 ± 0.03	1681 ± 32 %	169.16 ± 2.05	2665 ± 290 %

Result:
Reducing the free fitting parameters by applying certain values increases the reliability and accuracy of the results!

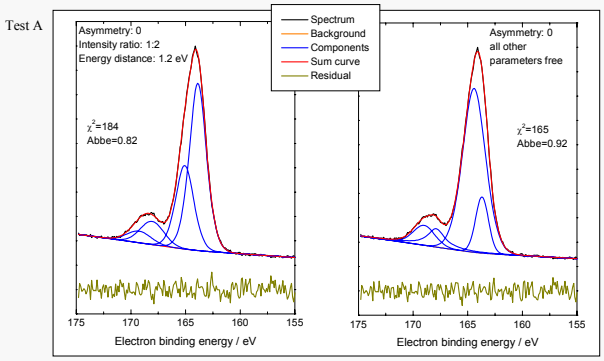

Fig. 4. Comparison of two different fits of a S 2p-spectrum with (Test A) and without (Test B) frozen doublet parameters

5. Influence of different noise levels

The noise level (usually $\sqrt{I}$, I = intensity) strongly influences the reliability and accuracy of the fitting results. In order to test the consequences of the noise level on the uncertainties three different artificial test spectra were generated and fitted using the same start parameters (Fig. 5.). An error analysis of the parameters was carried out (Fig. 6.).
Model: Product, 2 components
G/L-ratio: 0.5
FWHM: 1.2 eV
Distance between lines: 0.8 eV
Asymmetry: 0, fixed
Background: 10% of the intensity
Intensity of each component:
Spectrum 1: 100 Counts
Spectrum 2: 1000 Counts
Spectrum 3: 10000 Counts

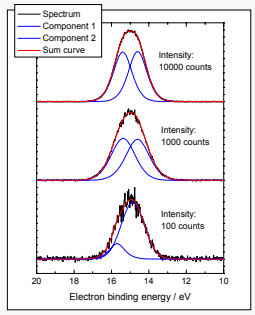

Fig. 5. Illustration of a peak fit of three test spectra with different intensities of the two single peaks generating the spectrum (100 counts, 1000 counts, 10000 counts)

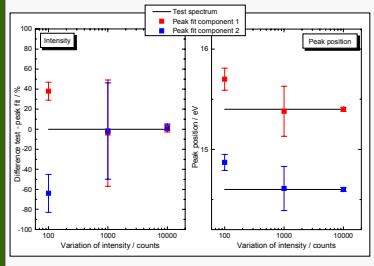

Fig. 6. Presentation of the calculated intensities and energies by means of peak fit, the estimated errors are displayed as error bars, left side: intensity, right side: energy

Result:
In case of a high noise level of the spectrum the peak fit can give false fit parameters and the error analysis may estimate incorrect uncertainties, too!

6. Some helpful recommendations to minimise uncertainties from peak fitting procedures

In order to minimise the uncertainties of the fit parameters the following recommendations should be considered:

1. recording of spectra with low noise and high peak to background ratio,
2. suitable choice of model function,
3. reducing the number of free parameters by applying existing information on the number and relations between the components of the fit model,
4. running the fit procedure including the background,
5. checking the values χ^2, χ^{2*} and the Abbe-criterion,
6. checking the distribution of normalized residuals for random behaviours,
7. calculation and check of parameter uncertainties,
8. consideration of series of spectra recorded under similar conditions rather than an individual spectrum.

Literature: [1] R. Hesse, T. Chassé, R. Szargan, Fresenius J. Analyt. Chem. 365 (1999) 48
[2] R. Hesse, T. Chassé, R. Szargan, Anal. Bioanal. Chem. 375 (2003) 856
[3] P.B. Bevington, Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill Book 1969

Important equations

Chi-square χ^2 :

$$\chi^2(\vec{p}) = \sum_{i=1}^{N} \frac{[M(i) - S(i,\vec{p})]^2}{M(i)}$$

Reduced chi-square χ^{2*}:

$$\chi^{2*}(\vec{p}) = \frac{\chi^2(\vec{p})}{N - P}$$

Curvature matrix **H**:

$$H = \begin{pmatrix} \frac{\partial^2 \chi^2}{\partial p_1^2} & \frac{\partial^2 \chi^2}{\partial p_1 \partial p_2} & \cdots & \frac{\partial^2 \chi^2}{\partial p_1 \partial p_P} \\ \frac{\partial^2 \chi^2}{\partial p_2 \partial p_1} & \frac{\partial^2 \chi^2}{\partial p_2^2} & \cdots & \frac{\partial^2 \chi^2}{\partial p_2 p_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 \chi^2}{\partial p_P \partial p_1} & \frac{\partial^2 \chi^2}{\partial p_n \partial p_2} & \cdots & \frac{\partial^2 \chi^2}{\partial p_P^2} \end{pmatrix}$$

Error of the peak areas A:

$$\Delta A(\vec{p}) = A(\vec{p} + \Delta\vec{p}) - A(\vec{p})$$

M(i) = Measuring spectrum
p_k = Fit parameter
S(i,p) = Model function
N = Number of measurement points
P = Number of fit parameters

Printed by URZ (University of Leipzig)