X-RAY SPECTROMETRY
X-Ray Spectrom. 2002; **31**: 173–177
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/xrs.529

Parameter refinement in the analysis of X-ray irradiated samples†

A. Carreras,[1]* R. Bonetto,[2]‡ G. Stutz,[1] J. Trincavelli[1]‡ and G. Castellano[1]

[1] Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba. Ciudad Universitaria, 5000, Córdoba, Argentina
[2] Centro de Investigación y Desarrollo de Procesos Catalíticos, CONICET–UNLP, Calle 47 No. 257–Cc 59, 1900, La Plata, Argentina

Received 11 January 2001; Accepted 29 June 2001

Although the refinement of parameters is a well known method in the scope of X-ray diffraction, it also appears as a powerful tool in other spectroscopic techniques. This work presents a method for the refinement of different atomic and experimental parameters in X-ray fluorescence (XRF). It consists of minimizing the differences between an experimental X-ray spectrum and a function proposed to account for the characteristic peaks and background spectrum from the corresponding sample, as well as for detection artifacts. The algorithm starts from some initial values for the different parameters involved, and a numerical iterative procedure produces improved values for them. After the general aspects of the method for refining XRF parameters are presented, a simple application is given for spectra measured in metallic samples with a monochromatic beam from a synchrotron source. In this case, the optimization algorithm is used for the determination of relative transition probabilities for the K shell. Copyright © 2002 John Wiley & Sons, Ltd.

INTRODUCTION

In X-ray diffraction (XRD), the method of parameter optimization using whole spectra is a widely used technique. Since the first time it was implemented by Rietveld,[1–3] it has yet to be extended to other spectroscopical techniques, with the exception of the POEMA code[4] for electron probe microanalysis (EPMA). On the basis of the good performance achieved for EPMA, in this work the method is implemented for X-ray fluorescence (XRF).

Although the method may be used in different spectroscopic systems, the description given here is restricted to XRF spectra acquired with an energy-dispersive system. The strategy followed is based on least-squares fitting of a selected region of the entire observed spectrum. In order to minimize the differences between the experimental and the calculated spectra, an iterative procedure is carried out. For this purpose, the expressions used for the predicted spectrum are based on fundamental parameters for characteristic lines.[5] An algorithm for the background continuum may be added, although in the examples shown here the corresponding intensity was low enough to be considered as constant. Finally, detection artifacts are conveniently taken into account.

If the fitted spectrum has N channels, the quantity to be minimized can be written as:

$$\chi^2 = \frac{1}{N-P}\sum_i \frac{(\tilde{I}_i - I_i)^2}{I_i} \quad (1)$$

where the summation runs over all the N data points, P is the number of parameters adjusted, and I_i and $\tilde{I}_i$ denote respectively the experimental and calculated intensities for the energy E_i of channel i. Thus, the models chosen for the prediction of $\tilde{I}_i$ will determine the values for χ^2 through the parameters involved. Usually, these are complicated functions, which generally implies a non-linear least-squares fitting. Since the risk of falling in local minima is not negligible, the initial guess for the parameters must be quite close to the correct values; alternatively, it is recommendable to begin with different estimates and check that the minimum achieved is always the same. The procedure proposed here must be regarded as a method for the refinement or optimization of parameters, and not for determining them. Consequently, it may become a fundamental tool when associated with a quantification routine.

In the next sections, theoretical support is given for the prediction of whole spectra in XRF, as well as a description of the method proposed. Some applications are then presented, including the characterization of detector properties and the refinement of atomic transition rates for K lines.

DESCRIPTION OF THE PROCEDURE

Two main issues are involved in the construction of the algorithm. Firstly, a full theoretical description for whole

*Correspondence to: A. Carreras, Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Ciudad Universitaria, 5000, Córdoba, Argentina.
E-mail: alejo@quechua.fis.uncor.edu
†Presented at SARX-2000, 7th Latin-American Conference on Analysis by X-ray Techniques, São Pedro, Brazil, 19–24 November 2000.
‡Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina.
Contract/grant sponsor: Consejo Nacional de Investigaciones Científicas y Técnicas (Argentina).
Contract/grant sponsor: Secretaría de Ciencia y Técnica (Universidad Nacional de Córdoba).
Contract/grant sponsor: Agencia Córdoba Ciencia SE.

spectra acquired in XRF must be given. Secondly, it is necessary to implement a numerical procedure to minimize the differences between experimental and predicted spectra.

Intensity predictions

Since all the spectra analysed in this work were measured for a monochromatic incident beam on sample and detector in vacuum, the background $B(E_i)$ below the characteristic peak regions was negligible. Therefore, no particular prediction for it is necessary, a constant value consequently being used.

When a sample of mass thickness t is irradiated with photons of energy E_0, the expressions arising from fundamental parameters[5] for the characteristic intensity $P_{j,q}$ of the line q from element j, is

$$P_{j,q} = \alpha C_j \frac{N_A \tau_j(E_0)}{A_j} \omega_j f_{j,q} \varepsilon_{j,q} \frac{\Delta\Omega}{4\pi} \times \frac{1 - \exp\{-[\mu(E_0)\,\mathrm{cosec}\,\psi_1 + \mu(E_{j,q})\,\mathrm{cosec}\,\psi_2]t\}}{\mu(E_0)\,\mathrm{cosec}\,\psi_1 + \mu(E_{j,q})\,\mathrm{cosec}\,\psi_2} \quad (2)$$

where α is a constant proportional to the number of incident photons, C_j is the mass concentration of element j, A_j its atomic weight, N_A is Avogadro's number, $\tau_j(E_0)$ is the photoelectric cross-section of element j for E_0, ω_j is the fluorescence yield for the atomic shell considered, $f_{j,q}$ is the transition rate related to the observed line q, $\mu(E_0)$ and $\mu(E_{j,q})$ are the mass absorption coefficients of the sample for the incident and the characteristic energy respectively, ψ_1 and ψ_2 are the incident and take-off angles, $\varepsilon_{j,q}$ is the detector efficiency for the corresponding energy and $\Delta\Omega$ is the solid angle subtended by the detector. This expression does not account for multiple scattering effects, as they are negligible in the samples analysed. In the applications presented in this work, the fluorescence yield coefficients were taken from Hubbell[6] and the mass absorption coefficients from Heinrich,[7] whereas transition rates have been optimized in most cases.

Although the theoretical prediction relies on the expressions given for the background and characteristic intensities, different X-ray detection effects must also be taken into account. The detector used in this work for spectrum acquisition was a silicon detector. For this kind of detector, the basic detection process involves a proportional conversion of photon energy into an electrical signal, which is shaped and amplified, and then passed to a multichannel analyser. A relationship between the channel number in which photons are registered and the corresponding energy must be supplied. A simple linear calibration is usually enough, involving two parameters: the 'gain' and the 'zero'.

For photons of energy E, the detector system response is a broadened peak whose distribution can be considered as Gaussian to a first approximation; its standard deviation σ being a function of photon energy:[8]

$$\sigma = (n^2 + \varepsilon F E)^{1/2} \quad (3)$$

where n is the uncertainty due to the electronic noise of the amplification process, F is the Fano factor and ε is the mean energy required for a single electron–hole pair formation (3.62 eV at 300 K).

Detector efficiency in Si detectors may strongly influence the results if it is not adequately taken into account. Typically, it is close to 100% between 3 and 10 keV. At lower energies, efficiency falls owing to absorption in the front window and layers; on the other hand, higher-energy photons have increasing probabilities of traversing the active silicon volume without being detected. The different thicknesses that must be known are typically an isolating beryllium window, a gold layer contact evaporated onto the front surface, a dead silicon layer, and the active detector thickness. For the examples shown in this work, energies were high enough to make the effect of absorption in the three front layers negligible. However, it is worth mentioning that correction for detection losses due to the finite active volume of Si were important.

Spectra collected with Si detectors may also show a spurious Si peak due to the photoelectric absorption of a photon within the dead Si layer of the detector. The consequent Si K photon may enter the active region and be registered, whereas Auger and photoelectrons are much more likely to be absorbed in the dead layer.[9] As a result, a photon of only 1.739 keV, corresponding to the Si K peak is registered instead of the one actually emitted by the sample. For the spectral energy regions considered in this work, these peaks were always excluded.

Finally, some of the charge carriers produced by a photon arriving at the detector may be 'trapped' before being collected. Thus, the output sent to the amplifier corresponds to an energy lower than the original one. This effect is manifested in asymmetrical peaks with low-energy tails, departing from the assumed Gaussian shape. Since the highest concentration of traps occurs in a transient region close to the detector surface, between the active volume and the dead layer, peaks appear to be more asymmetric for soft X-ray lines. Therefore, a modification to the Gaussian function is necessary in order to account for this effect, which is the most important in the examples presented in this work. This modification was included by means of the Hypermet[10] function

$$H_{j,q}(E_i) = A[G_{j,q}(E_i) + S_{j,q}(E_i) + T_{j,q}(E_i)]$$

in which A is a normalization factor, $G_{j,q}(E_i)$ is a Gaussian function centred at the characteristic energy $E_{j,q}$:

$$G_{j,q}(E_i) = \frac{1}{\sqrt{2\pi}\sigma_{j,q}} \exp\left[-\frac{(E_i - E_{j,q})^2}{2\sigma_{j,q}^2}\right]$$

$S_{j,q}(E_i)$ is the step function of height $s_{j,q}$ convoluted by the Gaussian:

$$S_{j,q}(E_i) = s_{j,q}\ \mathrm{erfc}\left(\frac{E_i - E_{j,q}}{\sqrt{2}\sigma_{j,q}}\right)$$

and $T_{j,q}(E_i)$ is an exponential tail of width $\beta_{j,q}$ and height $t_{j,q}$ convoluted by the Gaussian:

$$T_{j,q}(E_i) = t_{j,q}\, e^{(E_i - E_{j,q})/\beta_{j,q}}\ \mathrm{erfc}\left(\frac{E_i - E_{j,q}}{\sqrt{2}\sigma_{j,q}} + \frac{\sigma_{j,q}}{\sqrt{2}\beta_{j,q}}\right)$$

In these functions, the parameters $s_{j,q}$, $t_{j,q}$ and $\beta_{j,q}$ are not known *a priori* and must be refined if peak asymmetries must be taken into account.

There are two further artifacts that have not been considered in this first stage, since they would not influence the results. First, there is a finite probability that Si K photons produced in the detector following absorption of an incident photon will escape from the detector. In this event, the energy of the escape photon is not deposited in the detector, and the height of the consequent pulse is correspondingly reduced, i.e. an 'escape peak' occurs 1.739 keV below the main peak. Second, two pulses may reach the main amplifier within a very short time interval, resulting in the appearance of a spurious peak in the spectrum, corresponding to the sum of the energy of the two original pulses ('sum peak').

With all these spectral characteristics taken into account, the next step is to proceed with the minimization of χ^2 in Eqn (1).

Parameter optimization

Some of the magnitudes involved in the previous expressions may not be precisely known *a priori*; the challenge, therefore, is to find the set of parameters that best fits the general shape of the proposed function to the experimental spectrum:

$$\tilde{I}_i = B(E_i) + \sum_{j,q} P_{j,q} H_{j,q}(E_i) + P_{\mathrm{Si}} G_{\mathrm{Si}}(E_i) \qquad (4)$$

where $H_{j,q}$ is the modified Gaussian function associated with the peak intensity $P_{j,q}$ and P_{Si} is the internal fluorescent Si peak, spread by means of a Gaussian distribution G_{Si}. The parameters that may be optimized are: the scaling factor α of Eqn (2), and a possible scale factor for the background spectrum; the zero and gain of the detection chain; the peak-width factors n and F of Eqn (3); the transition rates $f_{j,q}$; the fluorescence yields ω_j and the mass concentrations C_j of Eqn (2); the parameters involved in the function $H_{j,q}$ for each peak; the four thicknesses associated with the detector efficiency; the amplitude of the P_{Si} peak; the transition energies for the involved decays; etc.

The algebraic expression for Eqn (1) involves the intensities $\tilde{I}_i$ of Eqn (4) in a complex way, and the minimization of χ^2 must be performed numerically. Although no perfect minimization procedure exists, certain characteristics of a given routine make it suitable for a particular application. Among the different possibilities (e.g. see Ref. 11), the downhill simplex algorithm[12] was chosen because, besides being a robust routine, it requires only function evaluations, not derivatives. This fact is particularly important, since it is often necessary to deal with functions whose computed derivatives do not accurately point the way to the minimum, usually because of truncation error in the method of derivative evaluation.

The algorithm is given a starting guess, i.e. a P-vector of independent variables, as the first point to try. Around this vector, a set of P additional points is generated, which build the $P + 1$ vertices of a P-dimensional figure called 'simplex'. The function to be minimized is evaluated in each of these vertices. Then, a sequence of reflections and contractions of the simplex is generated; through the corresponding function evaluation in the new P-points, the P-dimensional space is scanned until a minimum is found.

As in every extremization routine, care must be taken about the possibility of falling in local minima instead of the desired global extremum. A good way to overcome this problem is to start from a reasonable initial guess for the parameters to optimize; in addition, it is frequently a good idea to restart the minimization routine at the point where it claims to have found the minimum, reinitializing certain ancillary scale factors.[11] Another solution for confirming that the minimum found is the global one, is trying to get it by starting from widely varying initial values for the parameters.

A recommendable strategy for performing the minimization procedure is to carry it out by choosing only one or two parameters at a time; once their values have been achieved, they are set fixed and a new reduced group of parameters is allowed to vary. When all the chosen parameters are refined, the procedure may be restarted with the values obtained as initial guesses, varying all of them simultaneously.

Usually, a visual examination of the intermediate results helps in making an appropriate decision for the strategy to adopt. For this purpose, it is useful to plot the predicted and experimental spectra, as well as their differences, after the criterion of convergence is fulfilled.

Once convergence is achieved, an estimate for the uncertainties in the parameters concerned is assessed. For this purpose, the selected region of the experimental spectrum is regarded as a vector $\boldsymbol{y}$ whose components are the number of counts at each channel. The optimized parameters can also be thought as a vector $\boldsymbol{x}$, and the function relating both vectors, as a matrix M ($\boldsymbol{x}$). It can be seen[13] that the uncertainties of the parameters x_i, arising from the variance–covariance matrix V_x, can be related to the variance–covariance matrix V_y for the experimental spectrum by means of:

$$V_x = [A^{\mathrm{T}}(V_y)^{-1}A]^{-1}$$

where $A_{ij} = \partial M_i/\partial x_j$. These derivatives are calculated numerically in order to yield the matrix V_x, whose diagonal elements are the searched variances for the parameters x_i.

The optimization method described above was implemented in the computer program PRAXIS (an acronym for parameter refinement in the analysis of X-ray irradiated samples). The applications shown in the next section have been carried out by means of this program.

APPLICATIONS OF THE METHOD

Several spectra of standard metallic thin samples were measured in the XRD-beamline of the Laboratório Nacional de Luz Síncrotron (in Campinas, Brazil). Samples were positioned in a 45°–45° geometry and X-ray detection was carried out by means of an Amptek XR100 Si detector of 300 μm thickness, 7 mm^2 front area and a 25 μm Be window. The associated electronic chain resulted in a full width at half maximum (FWHM) of 287 eV at Mn Kα. The white synchrotron radiation beam was monochromatized by using an Si(111) double-crystal monochromator. Incident beams of

XRS

energies ranging from 6 to 12.5 keV were used; samples and detector were mounted in a chamber evacuated to 0.1 mbar, in order to reduce absorption in the air. In addition, a negligible background radiation was achieved because of the very low scattering in the incident beam, as well as in the fluorescence beam paths. The sample thicknesses were chosen so that the contribution of multiple scattering events was not important.

These experimental data were then processed with the program developed. For illustrating the methodology, an input data file was written for a pure vanadium sample 7.5 μm thick irradiated with an 8 keV beam. The gross initial estimates for the different parameters are given in Table 1. With this crude initial guess, the value obtained for χ^2 was 1300. Figure 1 shows the experimental data, the initial predicted spectrum (plot a), and some of the subsequent refined spectra. The first step of the process was to improve the calibration parameters, obtaining a χ^2 value of 406 (plot b). After the second step, which consisted in the refinement of the scaling factor α, a χ^2 value of 75.1 was obtained (plot c). The third and fourth steps were devoted to the improvement of n^2 and α respectively, yielding $\chi^2 = 13.7$ (plot d).

After a number of further optimization steps, a final set of data was achieved, also shown in Table 1. As can be seen, although the Fano factor obtained is in agreement with values reported in the literature,[14] its uncertainty is too large because of the high influence of n^2 in the FWHM of both peaks. If precision in this factor is required, a spectrum containing quite separate peaks must be used. Despite this uncertainty, the detector response is well described, since the value of n^2 compensates deviations in F. On the other hand, even when a poor initial guess has been given for f_α, this example evidences the 'refinement' of this parameter, since the final value obtained for this transition rate is in agreement with those appearing in the literature (see below). In this case, the uncertainty is quite small, which allows one to ensure that the value obtained is clearly distinguishable from the initial one.

Table 1. Example of the refinement process in a vanadium spectrum

Parameter	Initial	Final
α	5000	$18\,880 \pm 130$
Zero (keV)	0	$-0.196\,76 \pm 0.000\,98$
Gain (keV/channel)	0.020	$0.022\,156\,3 \pm 0.000\,004\,2$
n^2 (keV^2)	0.043	0.0874 ± 0.0011
F	0.11	0.122 ± 0.058
f_α	0.9	0.8795 ± 0.0063
s	0.0005	$0.000\,47 \pm 0.000\,68$
t	0.13	0.123 ± 0.036
β (keV)	0.25	0.223 ± 0.031

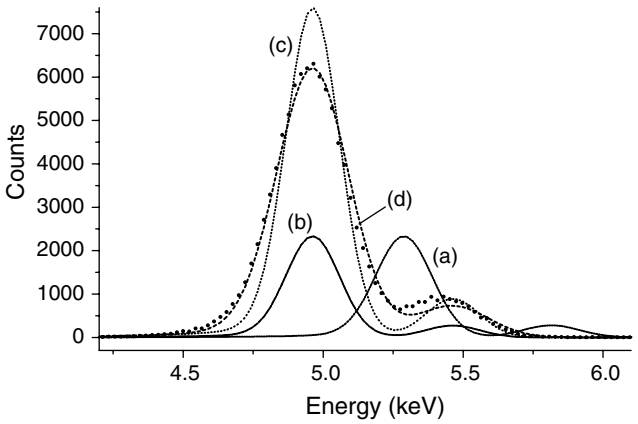

Figure 1. Vanadium spectrum predicted with PRAXIS at different stages. Dots: experimental; (a) initial guess; (b) after first step; (c) after second step; (d) after fourth step.

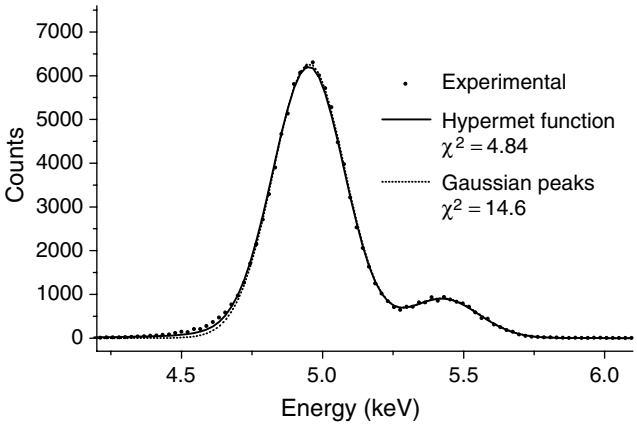

Figure 2. Vanadium spectrum predicted with PRAXIS after the refinement of several parameters.

The plot corresponding to this final set of values is shown in Fig. 2. As can be seen, the whole predicted spectrum region shows a very good agreement with experimental data. Figure 2 also shows the importance of including the asymmetry correction for incomplete charge collection: the best value for χ^2 obtained with purely Gaussian peaks is 14.6, which is reduced to 4.84 when shapes are corrected.

The process was repeated starting from different initial values, obtaining the same set of final results. The good performance of this application of the method implies that the models used for the different processes involved are correct.

Transition rates

Special attention has frequently been paid to inner-shell transition probabilities, because reliable experimental values can be used as a straight test for theoretical atomic models. In particular, separate relativistic Hartree–Fock solutions for atoms in their initial and final states have been used for calculating radiative decays of K or L vacancy states.[15,16] In addition an adequate knowledge of the transition rates allows one to improve the analyses performed by different spectroscopic techniques based on X-ray emission analysis, since peak overlaps between e.g. Kα and Kβ lines of neighbouring elements are frequently a problem for the analyst. Also, standardless algorithms must include accurate values for line ratios, since there is no comparison with standards, and it is not possible to cancel these factors out (e.g. see Ref. 17).

The spectra selected for this example were first fitted to find adequate values for scale factors and detector response, as explained above. Once the proper values for

Table 2. Transition rates for K decays determined by PRAXIS compared with values from the literature

Element	This work	Scofield[15]	Khan and Karimi[18]	Bé *et al.*[19]
Ti	0.879 ± 0.006	0.881	0.883	0.880
V	0.880 ± 0.006	0.880	0.883	0.878
Cr	0.875 ± 0.006	0.882	0.882	0.877
Mn	0.878 ± 0.005	0.878	0.883	0.878
Fe	0.874 ± 0.005	0.878	0.882	0.876
Co	0.876 ± 0.006	—	0.882	0.878

these parameters had been achieved, they were fixed and the line fractions refined instead. In this particular case, it is convenient to perform the refinement along with the peak scale factor α of Eqn (2), since changes in the intensity ratios may influence its value.

The values obtained for transition rates are compared with theoretical[15] and experimental[18,19] data from the literature in Table 2. As can be seen, a good agreement is achieved, following the general trend of data presented by other authors.

CONCLUSIONS

The implementation of a method for parameter optimization has been carried out for XRF, allowing the refinement of different magnitudes of interest in the frame of atomic physics as well as analytical and physical chemistry. Some examples have been shown here, involving spectra measured with monochromatic incident beams. However, the extension to conventional XRF (i.e. irradiation with polychromatic spectra generated in an X-ray tube) spectra is also possible: in this case, the incident spectral description is necessary, and work is being done for the optimization of the parameters involved in such a description. In addition, the scope of the proposed method includes other spectroscopic techniques, such as EPMA, particle-induced X-ray emission, etc., for which the corresponding theoretical description should be supplied.

The capabilities of this kind of procedure may be very useful in the knowledge of several basic magnitudes and atomic parameters. The examples for K-shell transition rate refinement given here show a good agreement with experimental and theoretical data from the literature. In view of these results, work is being done in the study of ratios between $K\alpha_1$, $K\alpha_2$, $K\beta_1$ and $K\beta_2$ lines, as well as L-shell transition rates. Further work will be carried out next in the optimization of Coster–Kronig yields, fluorescence yields, mass attenuation coefficients for low photon energies, etc. Alternatively, the effect of chemical bonds on transition rates and energies can also be faced by means of the procedure of parameter optimization proposed. An additional application may consist in the implementation of the refinement of compositions in multielemental samples, as a complement to the approach given by empirical algorithms or calibration methods.

Acknowledgements

The authors acknowledge financial support from the Consejo Nacional de Investigaciones Científicas y Técnicas (Argentina), the Secretaría de Ciencia y Técnica (Universidad Nacional de Córdoba) and the Agencia Córdoba Ciencia SE. This work was partially performed at the LNLS (Brazil) under proposal XRD 614/00.

REFERENCES

1. Rietveld H. *Acta Crystallogr.* 1966; **20**: 508.
2. Rietveld H. *Acta Crystallogr.* 1967; **22**: 151.
3. Rietveld H. *J. Appl. Crystallogr.* 1969; **2**: 65.
4. Bonetto R, Castellano G, Trincavelli J. *X-Ray Spectrom* 2001; **30**: 313.
5. Sherman J. *Spectrochim. Acta* 1955; **7**: 283.
6. Hubbell J. *NISTIR 89-4144*. National Institute of Standards and Technology: Gaithesburg MD, 1989.
7. Heinrich K. In *X-Ray Optics and Microanalysis*. Brown J, Packwood R (eds). University of Western Ontario: Ontario, 1987; 67.
8. Goldstein J, Newbury D, Etchlin P, Joy D, Romig Jr A, Lyman C, Fiori C, Lifshin E. *Scanning Electron Microscopy and X-Ray Microanalysis*, 2nd edn. Plenum Press: New York, 1992.
9. Reed S. *Electron Probe Microanalysis*, 2nd edn. Cambridge University Press: Cambridge, 1993.
10. Phillips G, Marlow K. *Nucl. Instrum. Methods* 1976; **137**: 525.
11. Press W, Flannery B, Teukolsky S, Vetterling W. *Numerical Recipes*. Cambridge University Press: Cambridge, 1989.
12. Nelder J, Mead R. *Comput. J.* 1965; **7**: 308.
13. Young R. *The Rietveld Method*. International Union of Crystallography, Oxford University Press: Oxford, 1993.
14. Knoll G. *Radiation Detection and Measurement*, 2nd edn. John Wiley & Sons: New York, 1989.
15. Scofield J. *Phys. Rev. A* 1974; **9**: 1041.
16. Scofield J. *Phys. Rev. A* 1974; **10**: 1507.
17. Trincavelli J, Van Grieken R. *X-Ray Spectrom.* 1994; **23**: 254.
18. Khan MdR, Karimi M. *X-Ray Spectrom.* 1980; **9**: 32.
19. Bé M-M, Lépy M-M, Plagnard J, Duchemin B. *Appl. Radiat. Isot.* 1998; **49**: 1367.