Resolving double peaks in high-resolution spectra by spectrum convolution

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS D: APPLIED PHYSICS
J. Phys. D: Appl. Phys. **37** (2004) 932–937 PII: S0022-3727(04)66291-7

Resolving double peaks in high-resolution spectra by spectrum convolution

A Likar[1,2], **T Vidmar**[2] **and M Lipoglavšek**[2]

[1] Faculty of Mathematics and Physics, University of Ljubljana, Slovenia
[2] Jožef Stefan Institute, Ljubljana, Slovenia

Received 18 July 2003
Published 24 February 2004
Online at stacks.iop.org/JPhysD/37/932 (DOI: 10.1088/0022-3727/37/6/019)

Abstract
It is shown that resolving doublets from high-resolution gamma-ray spectra using spectrum convolution offers several advantages over the usual peak search techniques that work with the original, non-convoluted spectrum. The spectral lines are approximated by Gaussian shapes with known widths and the convolution function is chosen as a projector which eliminates any linear background from the final results. The proposed method is more efficient compared with the standard method in frequently encountered cases where doublets are embedded within other peaks of the spectrum. Analytical expressions for the uncertainties of peak areas are given.

1. Introduction

In a previous work [1], we advocated the use of spectrum convolution for the determination of peak positions, areas and widths. In this contribution, we show that such an approach is also surprisingly efficient when doublets have to be resolved. With standard techniques the original spectrum is used as the data source. The shape of the background below a peak or a doublet is then modelled with appropriate functions with free parameters, which have to be determined essentially from the part of the spectrum in close vicinity of the peaks. The shapes of the peaks are added to the background shapes with yet another, entirely different set of coefficients which has to be determined in parallel with those fixing the background. The set of parameters describing the background is of no direct interest, but has to be deduced in order to extract the parameters specifying the peaks. Simultaneous determination of many parameters result in their increased uncertainties. In [1] we have shown that proper projection operators, acting upon the spectrum, can successfully remove the background, thus reducing the number of the parameters to be adjusted and decreasing the uncertainty of the essential parameters of the peak shape. The peak shape parameters can be determined one at a time by applying different projectors to the original spectrum.

The resolving of doublets is quite a delicate task, especially in the presence of other peaks close to the doublet. A limited portion of the spectrum is available for determining the background parameters, limiting the resolving power of any method employed. It is our aim to demonstrate that an approach that uses spectrum convolution is more efficient in this regard than the approach which fits the background. We analyse the simplest and thoroughly analytically solvable limiting case of small peaks on a large, constant background and give expressions for the covariance matrix of the area of a doublet for the standard least squares method and for the spectrum convolution method. A comparison then shows that the latter is favourable in the domain of the relevant parameters commonly encountered in practice.

2. The convolution approach

The convolution method [2–6] has been considered in the past only as a means to locate peaks in the spectrum. In such an approach, peak localization and peak evaluation are done separately. In our method of doublet resolution we extract peak areas, as well as their position, from the convoluted spectrum alone. The measured spectrum $p(x)$ is transformed by

$$I(y) = \int_{y-\Delta/2}^{y+\Delta/2} p(x)T(x-y)\,\mathrm{d}x \qquad (1)$$

The interval of the integration is finite, $[y-\Delta/2, y+\Delta/2]$. A proper choice of the function $T(x-y)$ is the key to a successful implementation of the method. Suppose an ideal spectrum without statistical fluctuations, which we denote by $p_0(x)$, exhibits a doublet on a background $\bar{n}(x)$

$$p_0(x) = A_1 z(x, -a, \sigma) + A_2 z(x, a, \sigma) + \bar{n}(x) \qquad (2)$$

Here, $z(x, a, \sigma)$ represents the shape of the total absorption peak with the scalable width σ and the amplitude A

proportional to the intensity of the doublet component. We assume that the shape is a normalized Gaussian one given by

$$z(x, a, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right) \tag{3}$$

We denote the peak position by a and its width by σ, which we assume is known in advance, be it from calibration or from the analysis of well defined single peaks.

The function $T(x - y)$ is chosen so as to filter out the background $\bar{n}(x)$. This is achieved by the requirement that for any y we should have

$$\int_{-\infty}^{\infty} \bar{n}(x) T(x - y)\, \mathrm{d}x = 0 \tag{4}$$

The experimental spectrum $p(x)$ can be written as

$$p(x) = p_0(x) + w(x) \tag{5}$$

where $w(x)$ represents the statistical fluctuation of the number of counts in the channel x and can be treated as a random variable with zero mean

$$\langle w(x) \rangle = 0 \tag{6}$$

and variance

$$\langle w^2(x) \rangle = \bar{n}(x) \tag{7}$$

We limit ourselves to the case of weak peaks on a large background. In the case of a constant background $\bar{n}(x) = \bar{n}$ this allows for an analytical treatment of the problem on the one hand and represents the least favourable case for doublet deconvolution on the other. There is no correlation between the fluctuations in different channels of the measured spectrum, so that

$$\langle w(x) w(x') \rangle = 0$$

for all $x \neq x'$. The brackets $\langle\ \rangle$ denote the ensemble average of a quantity. The so-called covariance matrix $\boldsymbol{R}$ of the spectrum with the elements

$$R(x, x') = \langle w(x) w(x') \rangle \tag{8}$$

is therefore proportional to the unit matrix. This considerably simplifies the treatment of the deconvolution process. Another simplification is achieved by selecting a simple transformation function $T(x, y)$, such as

$$T(x, y) = \frac{(x-y)}{\sigma^2} \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) \tag{9}$$

being the first derivation of the Gaussian shape function. A constant background is projected out to zero by this transformation and a linear one to a constant. We shall show that the influence of both is eliminated from the final results. Furthermore, any background directly below a peak, which is asymmetric with respect to the peak position, is also eliminated. This makes it possible to extract only the doublet components describing the peaks in the least squares procedure. In the actual implementation of the transformation function $T(x, y)$ one may choose an optimal value for the width σ or an even more sophisticated shape. Here, in order to illustrate our point, the width is chosen to be the same as the width of the individual peaks in the doublet to keep the resulting expressions sufficiently simple.

3. The least squares method

As this is the case with the standard approach to peak area determination, we utilize the well established least squares technique to extract the relevant parameters with our method as well. The crucial difference, as already mentioned, is that in the convolution approach there is no need for the background parameters to be fitted.

The least squares method provides us with concise expressions for the most probable amplitudes A_1 and A_2 and their covariance matrix. The model of the data is in our case given by a doublet with the components given by equation (3). It is convenient to formulate the method in the matrix form. The observation matrix $\boldsymbol{H}$, sometimes called the model matrix, is connected to the measurement vector $\boldsymbol{z}$, which is the spectrum itself:

$$\boldsymbol{z} = \boldsymbol{H}\boldsymbol{A} + \boldsymbol{w} \tag{10}$$

Here, the state vector $\boldsymbol{A}$ is given by

$$\boldsymbol{A} = \begin{bmatrix} \bar{n} \\ A_1 \\ A_2 \end{bmatrix} \tag{11}$$

and the two column matrix $\boldsymbol{H}$ by

$$\boldsymbol{H} = [1 \quad z(x, -a, \sigma) \quad z(x, a, \sigma)] \tag{12}$$

The rows of the matrix $\boldsymbol{H}$ are denoted by the channel number x. The vectors $\boldsymbol{z}$ and $\boldsymbol{w}$ represent the portion of the spectrum of interest (centred so that $x \in [-\Delta/2, \Delta/2]$) and the corresponding fluctuations, respectively.

The most probable value $\hat{\boldsymbol{A}}$ can be calculated from [7]

$$\hat{\boldsymbol{A}} = \boldsymbol{P}\boldsymbol{H}^{\mathrm{t}}\boldsymbol{R}^{-1}\boldsymbol{z} \tag{13}$$

where $\boldsymbol{P}$ is the covariance matrix of the vector $\hat{\boldsymbol{A}}$, given by

$$\boldsymbol{P} = (\boldsymbol{H}^{\mathrm{t}}\boldsymbol{R}^{-1}\boldsymbol{H})^{-1} \tag{14}$$

The superscript 't' denotes the transposition of a matrix. For the original, untransformed spectrum the final expressions covariance matrix $\boldsymbol{P}$ can be calculated in a straightforward manner, since the covariance matrix $\boldsymbol{R} = \bar{n}\boldsymbol{I}$ is proportional to the unity matrix $\boldsymbol{I}$. Matrix products can be calculated by integration. We give here only the expression for the uncertainty of the peak area P_A, derived in this way, which we need for comparison with the convolution method:

$$P_A = \frac{\alpha^2 - \beta\Delta}{(\beta - \gamma)[2\alpha^2 - \Delta(\beta + \gamma)]} \tag{15}$$

Here, the auxiliary quantities are defined as

$$\begin{aligned} \alpha &= \frac{1}{2}\left[\mathrm{erf}\left(\frac{\sqrt{2}a}{2\sigma} + \frac{\sqrt{2}\Delta}{4\sigma}\right) - \mathrm{erf}\left(\frac{\sqrt{2}a}{2\sigma} - \frac{\sqrt{2}\Delta}{4\sigma}\right)\right] \\ \beta &= \frac{1}{4\sigma\sqrt{\pi}}\left[\mathrm{erf}\left(\frac{a}{\sigma} + \frac{\Delta}{2\sigma}\right) - \mathrm{erf}\left(\frac{a}{\sigma} - \frac{\Delta}{2\sigma}\right)\right] \\ \gamma &= \frac{1}{2\sigma\sqrt{\pi}} \exp\left(-\frac{a^2}{\sigma^2}\right)\mathrm{erf}\left(\frac{\Delta}{2\sigma}\right) \end{aligned} \tag{16}$$

In the limit of very large Δ, we obtain the well known covariance matrix for the amplitudes of a doublet:

$$\boldsymbol{P} = \frac{2\bar{n}\sigma\sqrt{\pi}}{1-\exp\left(-2a^2/\sigma^2\right)} \begin{bmatrix} 1 & -\exp\left(-\dfrac{a^2}{\sigma^2}\right) \\ -\exp\left(-\dfrac{a^2}{\sigma^2}\right) & 1 \end{bmatrix} \tag{17}$$

This limiting covariance matrix is the optimal one and therefore represents the benchmark for what can be achieved by other methods.

4. The case of a convoluted spectrum

The convoluted spectrum is the transformed original one

$$\boldsymbol{u} = T\boldsymbol{z} \tag{18}$$

Since the background has been projected out, the new observation matrix is

$$\boldsymbol{TH} = [Tz(x,-a,\sigma) \quad Tz(x,a,\sigma)] \tag{19}$$

and the state vector $\boldsymbol{A}$ assumes the form

$$\boldsymbol{A} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \tag{20}$$

The columns of $\boldsymbol{TH}$ are orthogonal to a constant residuum of the transformed background of the spectrum. The transformation therefore not only eliminates the constant background from the spectrum but also removes any influence of the linear background on the peak amplitudes.

It is tempting to use expressions (13) and (14) to formulate the least squares method in a completely analogous manner as in the case of the treatment of the non-convoluted spectrum by putting

$$\hat{\boldsymbol{A}} = \boldsymbol{P}(\boldsymbol{TH})^{\mathrm{t}}(\boldsymbol{TRT}^{\mathrm{t}})^{-1}\boldsymbol{u} \tag{21}$$

where $\boldsymbol{P}$ is the covariance matrix of the vector $\hat{\boldsymbol{A}}$, given by

$$\boldsymbol{P} = ((\boldsymbol{TH})^{\mathrm{t}}(\boldsymbol{TRT}^{\mathrm{t}})^{-1}\boldsymbol{TH})^{-1} \tag{22}$$

Such a proposition, however, is not viable since the transformation T is a projection which does not have an inverse. Even sophisticated methods of inverting $\boldsymbol{TRT}^{\mathrm{t}}$, such as singular value decomposition, do not help. The inverse can only be restored by using the information on the background which has been projected out. This is, however, contrary to the original aim to work with the transformed spectrum only. The simplest proposition we can follow is to put

$$\boldsymbol{TRT}^{\mathrm{t}} \to \boldsymbol{I} \tag{23}$$

and use equations (21) and (22) to obtain the vector $\boldsymbol{A}$. It is clear that this simplification introduces additional noise into equation (21), but it does make analytical treatment of the method possible. Furthermore, in many practically occurring situations it enables more accurate determination of the peak amplitudes as extracted from the original spectrum. The covariance matrix of vector $\boldsymbol{A}$ is now not given by equations (22) and (23). A straightforward analysis shows that the correct covariance matrix $\boldsymbol{P}$ in case proposition (23) is accepted is given by

$$\boldsymbol{P} \to (\boldsymbol{G}^{\mathrm{t}}\boldsymbol{G})^{-1}\boldsymbol{G}^{\mathrm{t}}(\boldsymbol{TRT}^{\mathrm{t}})\boldsymbol{G}(\boldsymbol{G}^{\mathrm{t}}\boldsymbol{G})^{-1} \tag{24}$$

Here, $\boldsymbol{R}$ is the covariance matrix of the original, non-convoluted spectrum and $\boldsymbol{G} = \boldsymbol{TH}$. In our case $\boldsymbol{P}$ can be computed analytically. The diagonal elements can be written with the help of the following rather complicated expression:

$$\begin{aligned} P_{AA} = \bar{n}\frac{16\sigma^3\sqrt{3\pi}\,\mathrm{e}^{(a^2/\sigma^2)}}{243}\Bigg[& \frac{27\sigma^6\,\mathrm{e}^{(a^2/\sigma^2)}}{(\sigma^4\,\mathrm{e}^{(a^2/\sigma^2)}-(a^2-\sigma^2)^2)^2} \\ & + \frac{2\,\mathrm{e}^{(a^2/6\sigma^2)}(a^2-\sigma^2)(4a^4-36a^2\sigma^2+27\sigma^4)}{(\sigma^4\,\mathrm{e}^{(a^2/\sigma^2)}-(a^2-\sigma^2)^2)^2} \\ & + \frac{27\sigma^2(a^2-\sigma^2)^2}{(\sigma^4\,\mathrm{e}^{(a^2/\sigma^2)}-(a^2-\sigma^2)^2)^2}\Bigg] \end{aligned} \tag{25}$$

5. Comparison of variances

We are now in a position to compare the uncertainties of amplitudes obtained by treating the original spectrum and the transformed one. We assume that the portion of the spectrum available for the least squares treatment is the same in both cases. In order for the uncertainty of the amplitudes for the transformed spectrum to be close to the ones given by equation (25), we only need to have 3σ channels available to the left of the lower peak and the same amount available to the right of the upper peak of the doublet. This is intuitively clear if we consider the form of the transformed doublet, shown in figure 1. The main swing of the transformed peaks is accomplished within the interval $\pm 3\sigma$ around the centre of the peak. We can therefore use equations (15) and (25) for the comparison. The ratio of the two variances as a function of available number of channels Δ for different separations $2a$ of the peaks in the doublet is shown in figure 2. We can see that the treatment of the transformed spectrum is favourable compared to the treatment of the original one if $\Delta < 2a + 10\sigma$ and the relative separation of the doublet components $2a/\sigma$ is

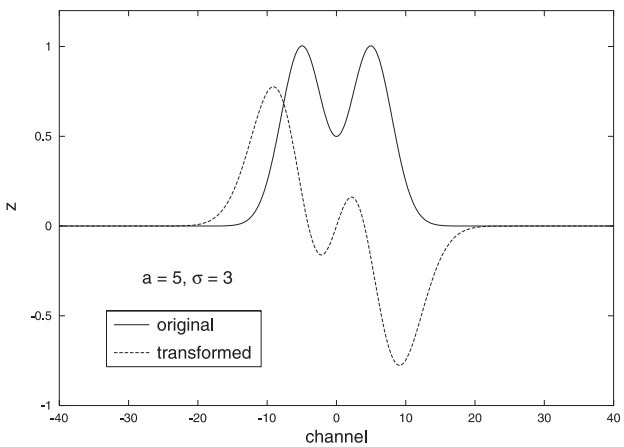

Figure 1. Original (——) and transformed (- - - -) portion of a spectrum containing a doublet with the individual peak widths of $\sigma = 3$ channels, which are typical for high resolution spectra, and the separation of peaks of $2a = 10$ channels. The transformation $T(x, y)$ has been carried out in accordance with equation (9).

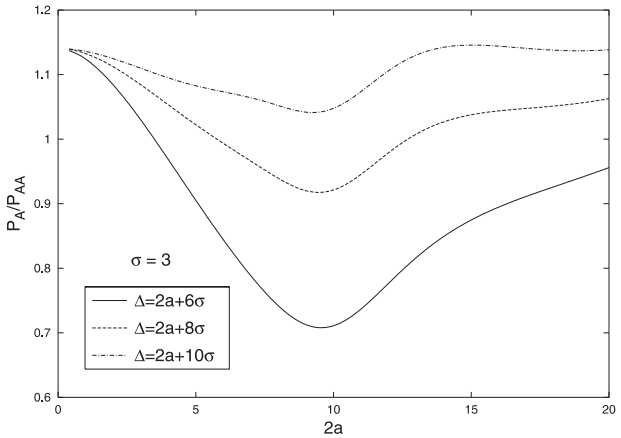

Figure 2. The ratio P_A/P_{AA} of variances given by equations (15) and (25) as a function of the separation $2a$ of the peaks in the doublet. The ratio is given for three different spans of the region of interest Δ, as indicated in the figure. The width of individual peaks in the doublet is $\sigma = 3$ channels. The ratio otherwise scales as a/σ. The noise reduction due to the transformation of the spectrum is quite significant in the case of $\Delta = 2a + 6\sigma$ (——) when the peak separation is in the range $1 < 2a/\sigma < 8$, and diminishes for larger values of Δ (- - - - and —·—). The limiting value of the ratio for $\Delta \to \infty$ and $a \to \infty$ is $8/9\sqrt{3} \approx 1.5396$, which represents the largest possible noise increase factor for an individual transformed peak.

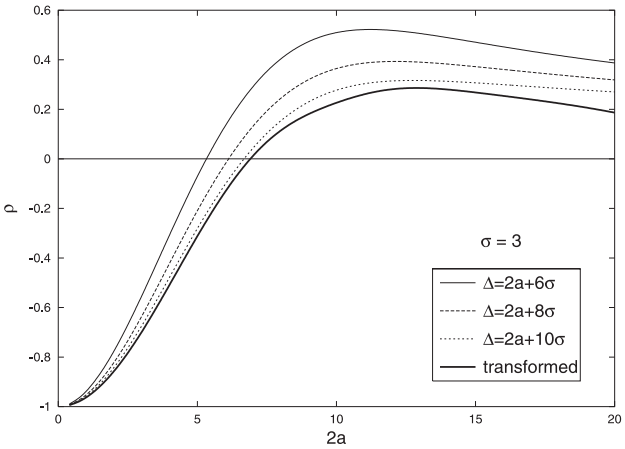

Figure 3. The correlation coefficient ϱ between the amplitudes of the doublet components as a function of the separation $2a$ of the peaks in the doublet. The solid, short-dashed and long-dashed lines correspond to the treatment of the original spectrum and are given for three different spans of the region of interest Δ, respectively. The heavy line corresponds to the treatment of the transformed spectrum.

in the region $1 < 2a/\sigma < 8$. For the separation $2a < \sigma$ the treatment of the original spectrum does indeed give slightly better results, but the uncertainties are so high in this region that the difference is of no practical importance. The rule of thumb is therefore the following: the original spectrum should be treated when there is a $\pm 5\sigma$ margin or more around the doublet which is not influenced by other peaks. The real margin tends to be larger—when treating original spectrum the approximation of a strictly constant background on both sides of the doublet is usually not viable. In figure 3 the results for the correlation coefficient ϱ between the amplitudes of the doublet components are given. Small correlations are always desirable. Generally speaking, we see that the transformed spectrum decouples the components of the doublet better.

It is easy to use the transformation method in cases where a doublet lies on a part of the background that deviates significantly from a linear shape. The transformation we use in such cases is T^2. This corresponds to a double differentiation of the original spectrum. We do not present any details of the analysis of this transformation, since it involves prohibitively complicated expressions to be reported here. However, it has been done in the same manner as the presented one. The qualitative results are quite similar to the ones presented for the case of the transformation T. For a restricted number of channels available for the analysis of the background the transformed spectrum offers better information for doublet decomposition than the original one.

It can be argued that in the transformed spectrum the peaks are broadened compared with the original one, which should diminish the number of channels available for background estimate around the doublet. However, doublets located in vicinity of other peaks can be straightforwardly treated by enlarging the observation matrix $\boldsymbol{TH}$ and adding a new column representing the interfering peak. The covariance matrix is then only slightly affected in case the interference is weak. A similar procedure is less accurate when treating the original spectrum since in addition to the column representing the interfering peak the columns due to background should be added to the observation matrix and these are usually poorly known.

6. Implementation of the convolution method

The analytical results reported in the previous sections were first thoroughly checked on a number of simplified simulated spectra in order to obtain statistically significant results. The spectra were constructed as Gaussian peaks on a constant background, which is sufficient for checking the derived expressions, since statistical fluctuations of the number of counts in individual channels were adequately simulated. However, to fully demonstrate the advantages of the convolution approach for resolving multiplets, we utilized our implementation of the method on selected regions of two complex spectra. The first of the two was obtained using the Monte Carlo package GEANT [8] and the second was a spectrum measured in our laboratory. The regions of the two spectra that we selected for analysis contained a background, resulting from incomplete interactions of photons belonging to peaks of higher energies, and four merged peaks (figures 4 and 5).

The spectrum generated with the Monte Carlo method is in all major respects a veristic substitute for measured spectra since a detailed HPGe detector model, complete with dead layers, the central core and the crystal rounding, was used to create it, including a surrounding lead shield. The parameters of the model were obtained in an earlier study [9] by adapting them to measured efficiency values for point sources positioned at a number of space points above the detector. The crucial advantage of Monte Carlo generated spectra lies in the fact that the shapes, widths, areas and positions of the peaks in such spectra are exactly known. Any peak search method can thus be accurately tested by running it on such spectra.

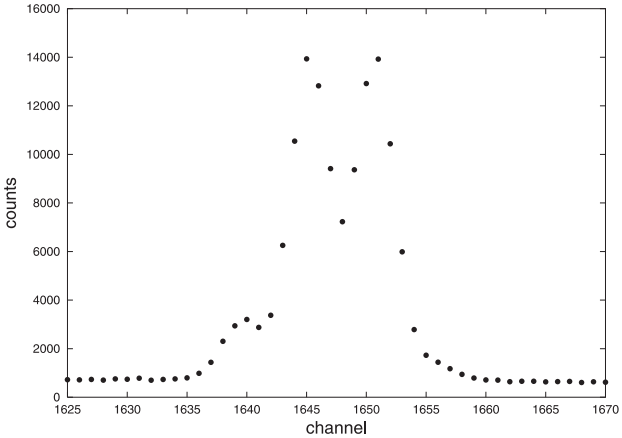

Figure 4. The portion of the Monte Carlo spectrum containing the quadruplet considered for analysis with our program and two other codes (table 1).

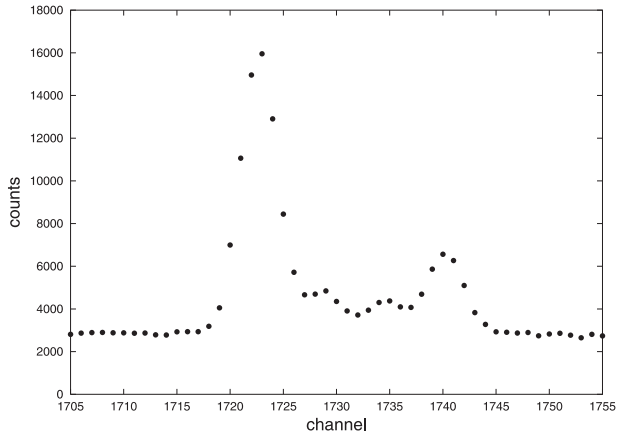

Figure 5. The portion of the measured spectrum containing the quadruplet considered for analysis with our program and two other codes (table 2).

Table 1. Results of the quadruplet decomposition for the Monte Carlo generated spectrum (figure 4), using methods A, B and C, as explained in the text.

Source	Area	|Δ area|
MC (ref.)	10 049	—
A	10 051	2
B	9 769	280
C	10 036	13
MC (ref.)	54 782	—
A	54 624	158
B	54 397	385
C	54 708	74
MC (ref.)	54 697	—
A	54 628	69
B	54 614	83
C	54 756	59
MC (ref.)	2 785	—
A	2 733	52
B	2 679	106
C	2 663	122

Table 2. Results of the quadruplet decomposition for the measured spectrum (figure 5), using methods A, B and C, as explained in the text.

Source	Area	|Δ area|
A (ref.)	57 473	—
B	57 765	292
C	57 630	157
A (ref.)	8 354	—
B	8 566	212
C	8 441	87
A (ref.)	6 464	—
B	6 725	261
C	6 530	66
A (ref.)	16 295	—
B	16 586	291
C	16 313	18

We confronted our implementation [1] of the convolution peak search method with two other peak search codes. The first was a classical least squares system [10], which demands human intervention to operate properly and proved to be very reliable. In tables 1 and 2 its results are denoted by (A). A parabolically shaped background and Gaussian peaks were assumed to fit the data with this package. The second method [11] is a well known and widely used commercial system (B), which was supplied with the appropriate information on the Gaussian shape of the peaks and the energy dependence of the full-width at half-maximum (FWHM) parameter for the case of the Monte Carlo spectrum. As it can be inferred from table 1, the method (A) and our method (C) are in close agreement with the true, known values for the Monte Carlo spectrum. The approach (B) is in this case less successful. For the measured spectrum (table 2), the values provided by method (A) were considered as the reference ones. We again see that the results of our approach (C) are much closer to those values than the results of approach (B). It should be mentioned that the uncertainties of the peak areas as reported by different programs always contain the statistical uncertainty of the peak area, which is not relevant in our case, since the peak areas of the Monte Carlo spectrum are known exactly and the areas of the measured spectrum obtained by manual analysis are assumed to be so. We have therefore not listed any uncertainties of the peak areas reported by different codes.

The values given in tables 1 and 2, however, only serve to illustrate the suitability of the convolution approach for the treatment of multiplets. The real test of the method comes from considering two different IAEA test sets of spectra and several other Monte Carlo spectra, which had been done with success in [1]. There, the program was confronted with numerous double and multiple peaks of different complexities and performed well both in terms of reporting the correct (relative and absolute) peak areas, positions and widths and in terms of the accuracy of the estimation of their uncertainties.

7. Conclusions

To our knowledge we have carried out the first thorough analysis of doublet resolution using the spectrum convolution approach, with which any realistic background is projected out to a high degree. When one is faced with spectra involving a dense concentration of peaks, the portion of the spectrum available for resolving the doublets may be severely limited. In such cases spectrum convolution provides

an attractive possibility by offering better doublet resolution than the traditional analysis of fitting the background along with the peaks. Even in very close vicinity of strong peaks the transformed spectrum is virtually free of background and additional peaks can be easily included into the least squares procedure by adding them to the observation matrix $\boldsymbol{G} = \boldsymbol{TH}$. This possibility is severely limited if the original spectrum is considered since even a slight variation of the background shape is important across a wide spectrum region.

References

[1] Likar A and Vidmar T 2003 A peak-search method based on spectrum convolution *J. Phys. D: Appl. Phys.* **36** 1903

[2] Mariscotti M A 1967 *Nucl. Instrum. Methods* **50** 309

[3] Routti J T and Prussin S G 1969 *Nucl. Instrum. Methods* **72** 124

[4] Black W W 1969 *Nucl. Instrum. Methods* **71** 317

[5] Robertson A, Prestwich W V and Keneth T J 1972 *Nucl. Instrum. Methods* **100** 317

[6] Hnatowicz V 1976 *Nucl. Instrum. Methods* **133** 137

[7] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes* 2nd edn (Cambridge: Cambridge University Press)

[8] Brun R, Bruyant F, Maire M, McPherson A and Zanarini P 1987 *GEANT CERN Report* DD/EE/84-1 (Geneva)

[9] Ashrafi S, Likar A and Vidmar T 1999 *Nucl. Instrum. Methods* A **438** 421

[10] Radford D C 1995 *Nucl. Instrum. Methods* A **361** 297

[11] Canberra Industries Inc. 1986 VAX/VMS *Hypermet Peak Search Program User's Manual* 07-0305,
Canberra Industries Inc. 1985 *Peak Search Program Algorithm Manual* 07-0064