Miroslav Morháč
Institute of Physics, Slovak Academy of Sciences, Bratislava, Slovakia

Sophisticated algorithms of analysis of spectroscopic data

ACAT2008, November 3-7, Erice, Italy

1.Introduction and motivation

•one of the most delicate problems of any spectrometric method is that related to the extraction of the correct information out of the spectra sections, where due to the limited resolution of the equipment, the peaks as the main carrier of spectrometric information are overlapping.

•conventional methods of peak searching based usually on spectrum convolution are inefficient and fail to separate overlapping peaks.

•the deconvolution methods can be successfully applied for the determination of positions and intensities of peaks and for the decomposition of multiplets.

•however before the application of deconvolution operation we need to remove the background from spectroscopic data.

•one of the basic problems in the analysis of the spectra is the separation of useful information contained in peaks from the useless information (background, noise).

•in order to process data from numerous analyses efficiently and reproducibly, the background approximation must be, as much as possible, free of user-adjustable parameters.

•baseline removal, as the first preprocessing step of spectrometric data, critically influences subsequent analysis steps.

•in the contribution we present a new algorithm to determine peak regions and separate them from peak-free regions.

•it allows to propose a new baseline estimation method based on sensitive non-linear iterative peak clipping with automatic local adjusting of width of clipping window.

•moreover automatic setting of peak regions can be used to confine intervals of fitting and to fit each region separately.

2. Background estimation

Goal: Separation of useful information (peaks) from useless information (background)

•very efficient method of background estimation, based on Statistics-sensitive Non-linear Iterative Peak-clipping algorithm (SNIP), has been developed in [1].

•SNIP method implicitly determines peak regions and peak-free regions.

•in [2] we extended the SNIP method for multidimensional spectra. In multidimensional spectra, the algorithm must be able to recognize not only continuous background but also to include all the combinations of coincidences of the background in some dimensions and the peaks in the other ones.

M.Morháč, J. Kliman, V. Matoušek, M. Veselský, I. Turzo: Background elimination methods for multidimensional coincidence gamma-ray spectra. NIM A 401 (1997) pp. 113-132.

•in [3] we proposed several improvements of the SNIP algorithm. Further we have derived a set of modifications of the algorithm that allow estimation of specific shapes of background and ridges as well.

M. Morháč, V. Matoušek: Peak clipping algorithms of background estimation in spectroscopic data, Applied Spectroscopy, Volume 62 (1), pp. 91-106, 2008.

•the SNIP algorithm is a multi-pass clipping loop which replaces each channel value y(i) with smaller value of y(i) and the mean [y(i-j)+y(i+j)]/2, where $j \in <1,m>$ and m is a given free parameter. To propose a suitable parameter one can use 2m+1=w where w is real width of an object (peak, doublet, multiplet) that should be preserved.

•choosing it too big can cause remaining background residua under peaks.

•on the contrary choosing it too small can cause that the baseline undergoes the real peaks and thus after subtraction decreases their net areas.

One-dimensional spectra

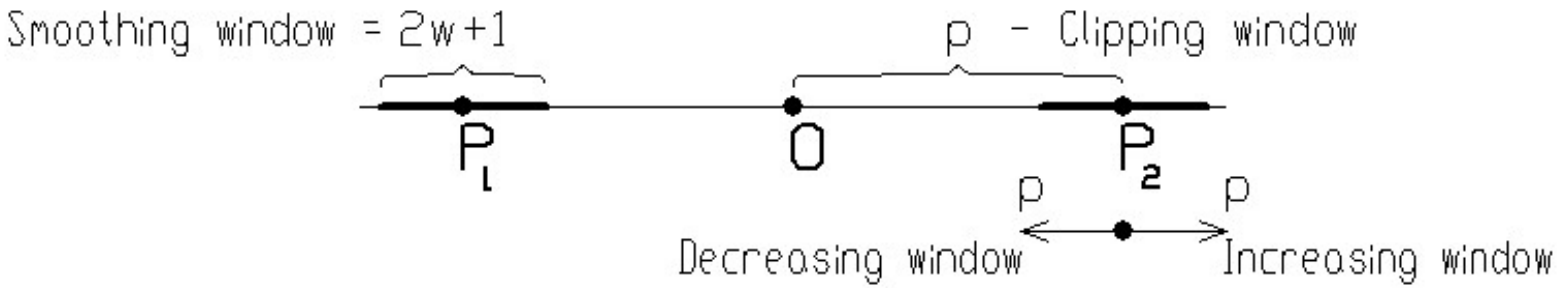

Fig. 1 Principle of background estimation in one-dimensional spectra with simultaneous smoothing

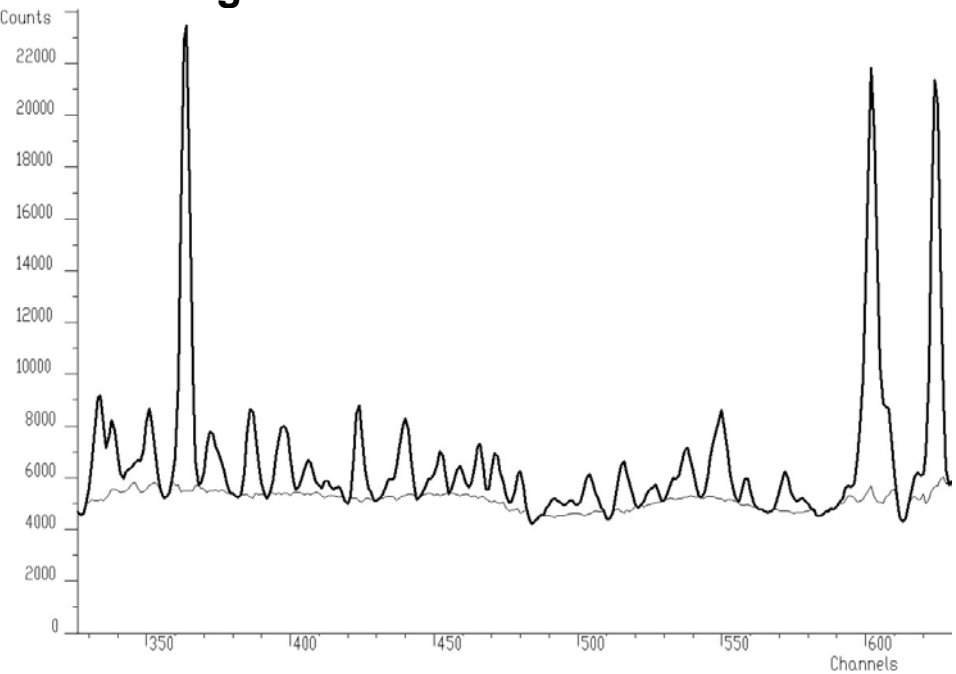

Fig. 2 An example of γ- ray spectrum with estimated background using increasing clipping window.

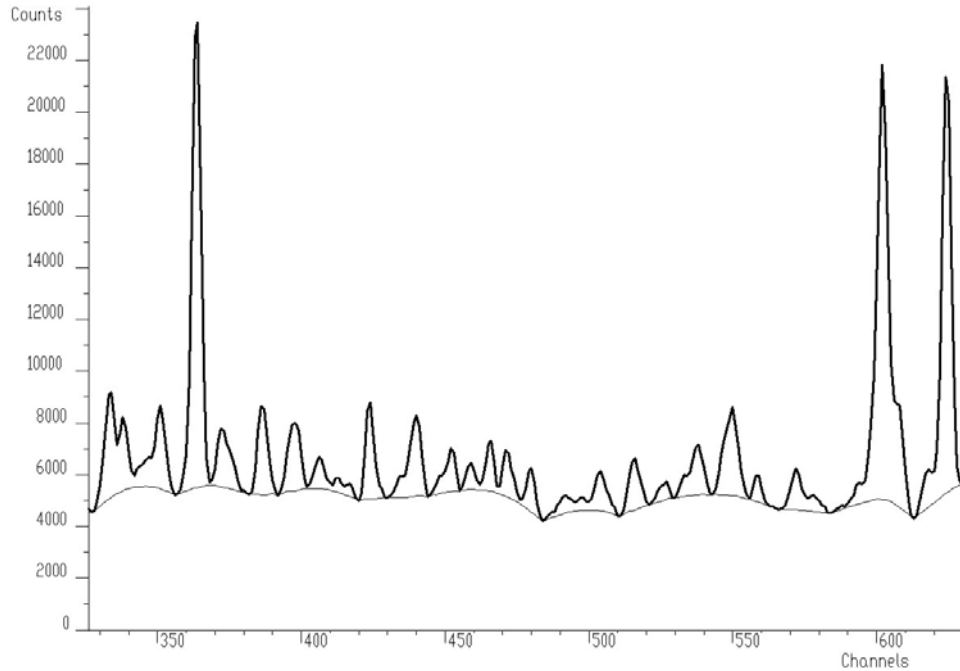

Fig. 3 An example of γ- ray spectrum with estimated background using decreasing clipping window.

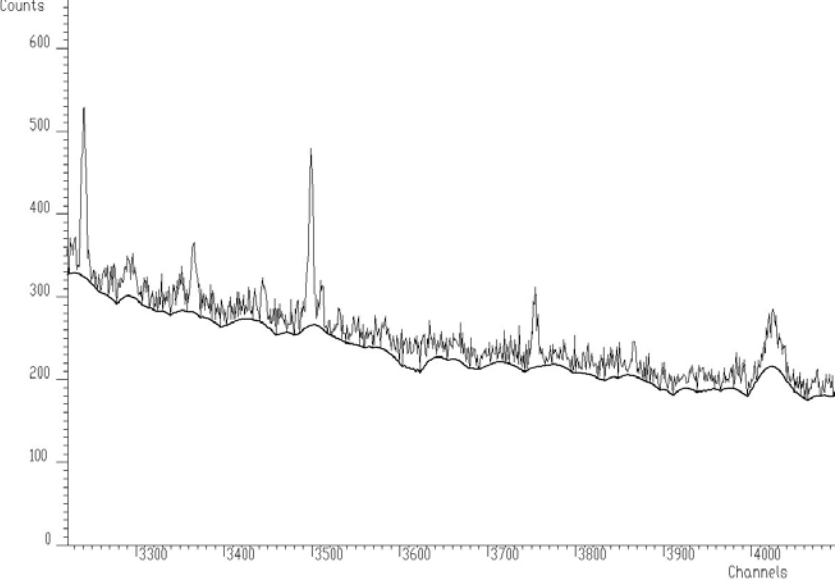

Fig. 4 Background estimated using SNIP algorithm without smoothing.

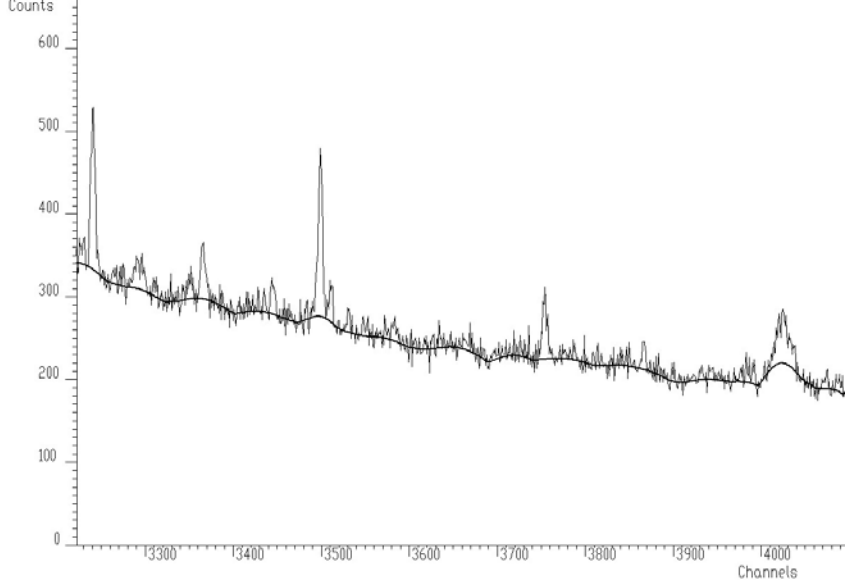

Fig. 5 Background estimated using SNIP algorithm with simultaneous smoothing.

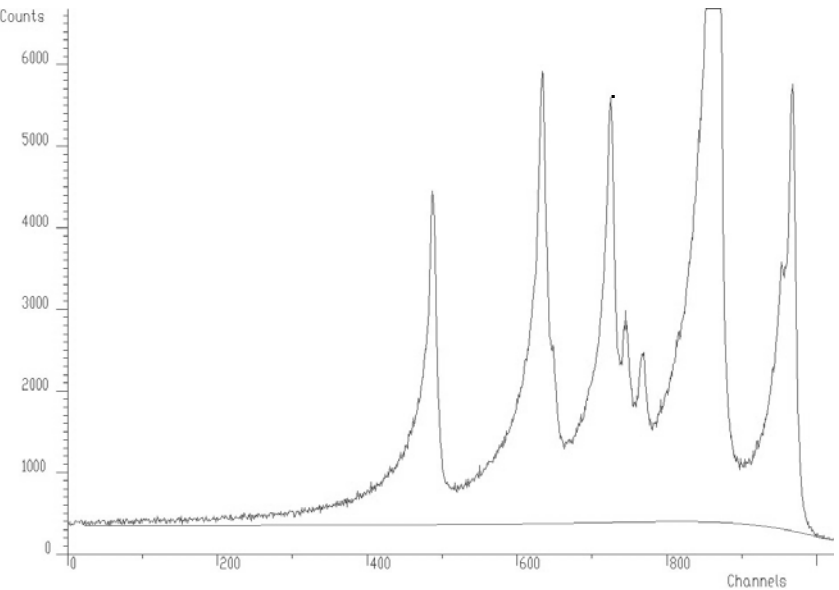

Fig. 6 Electron spectrum and estimated background employing non-symmetrical clipping window (80,20)

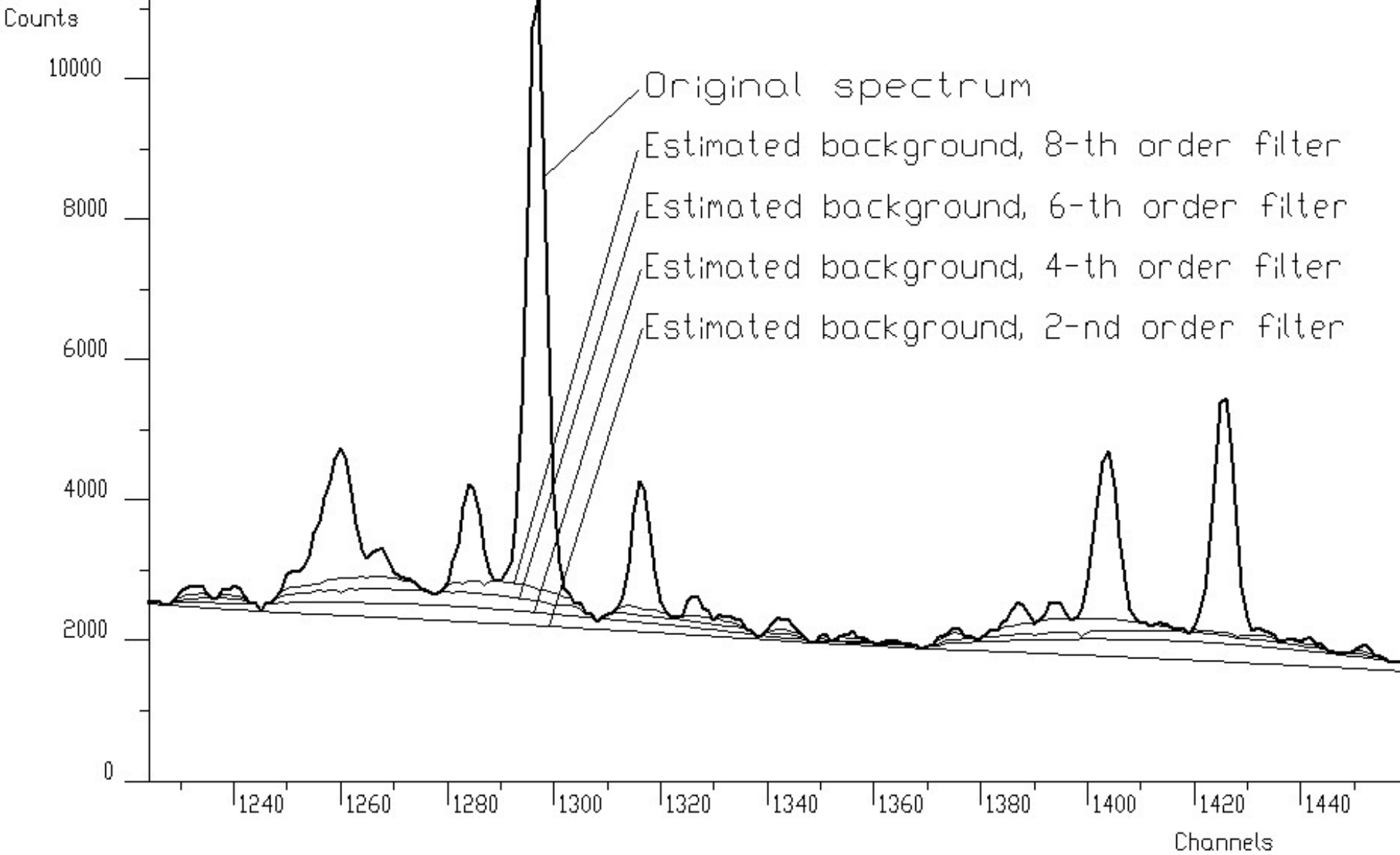

Fig. 7 Illustration of the influence of clipping filter order on estimated background

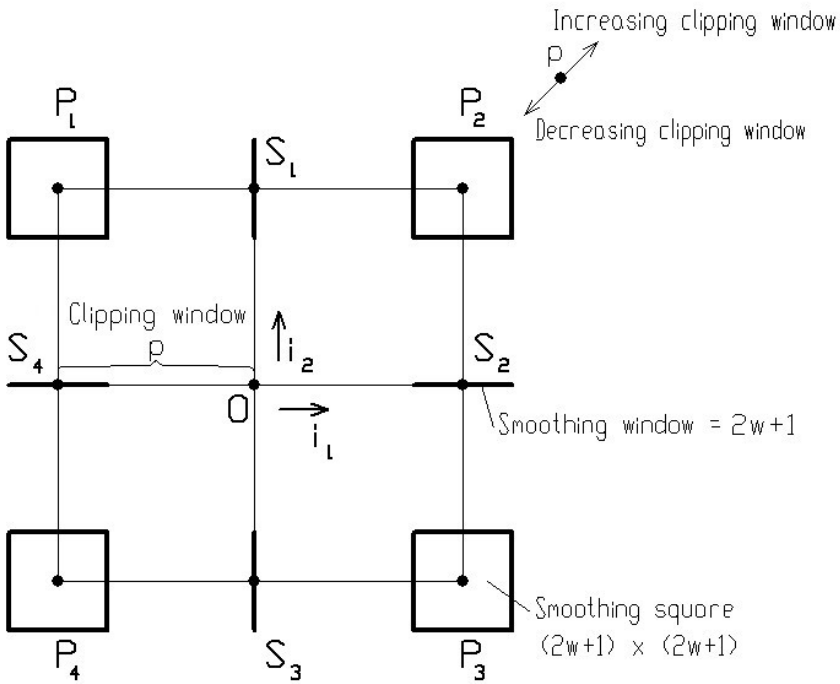

Fig. 8 Principle of background estimation in two-dimensional spectra with simultaneous smoothing.

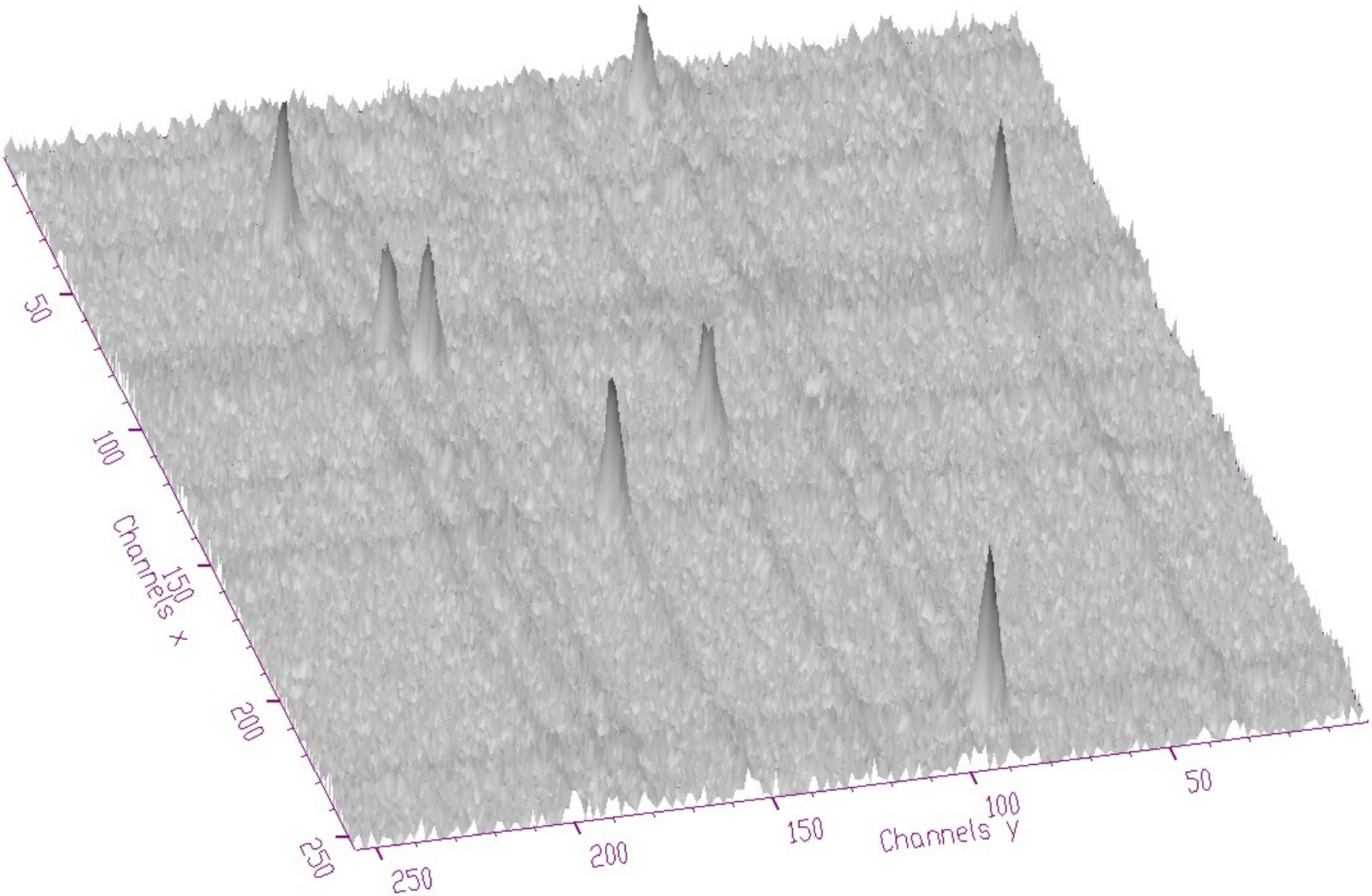

Fig. 9 Original synthetic spectrum with 8 peaks and high level of noise (30%)

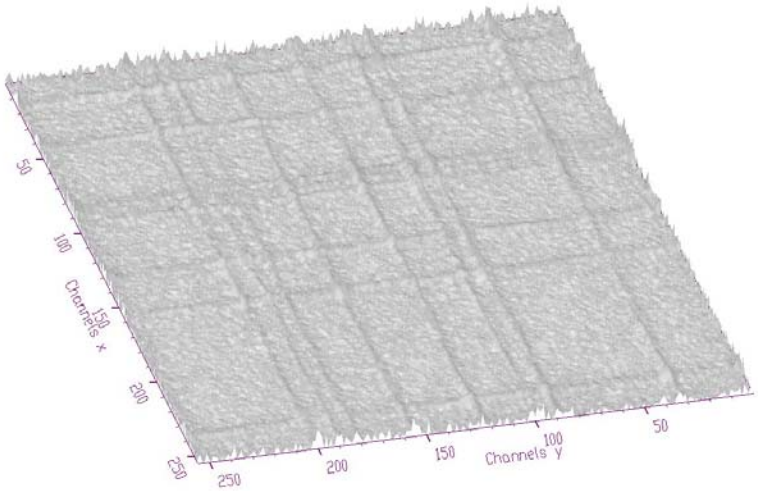

Fig. 10 Background estimated without smoothing

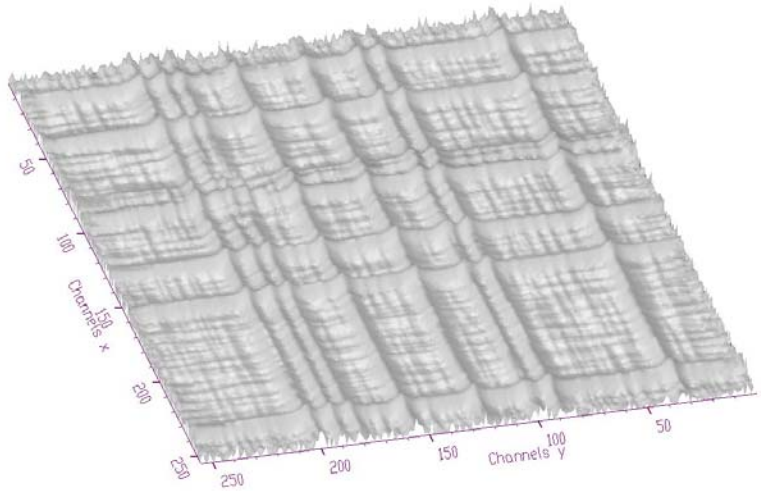

Fig. 11 Background estimated with smoothing

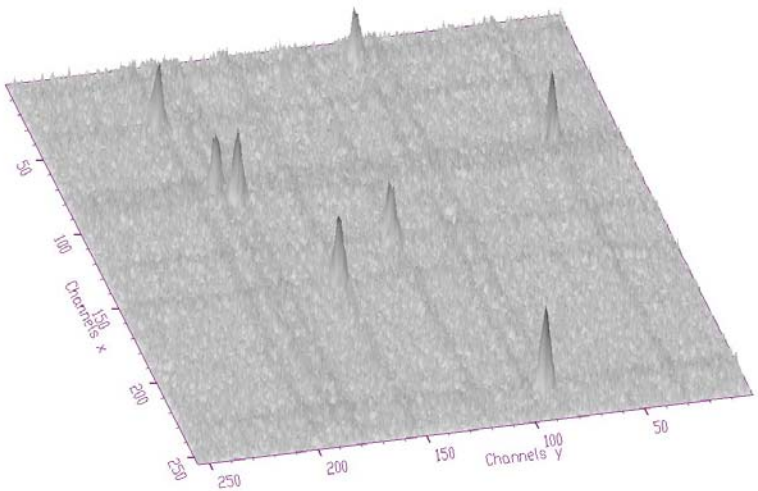

Fig. 12 Spectrum from Fig. 8 after subtraction of background from Fig. 9

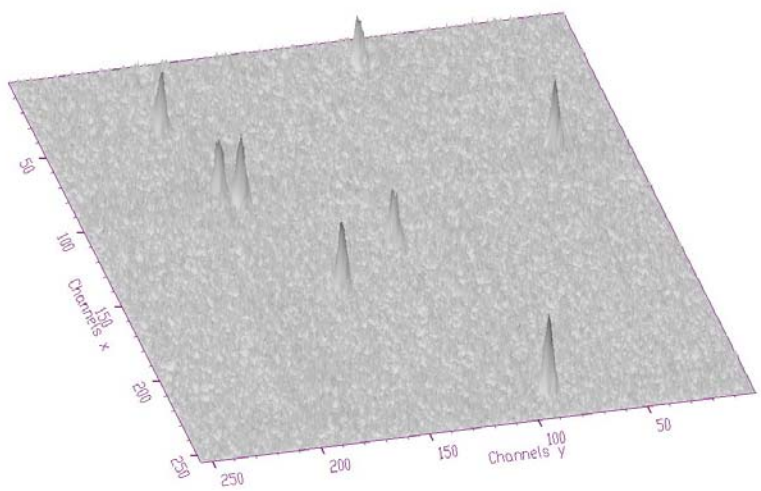

Fig. 13 Spectrum from Fig. 8 after subtraction of background from Fig. 10

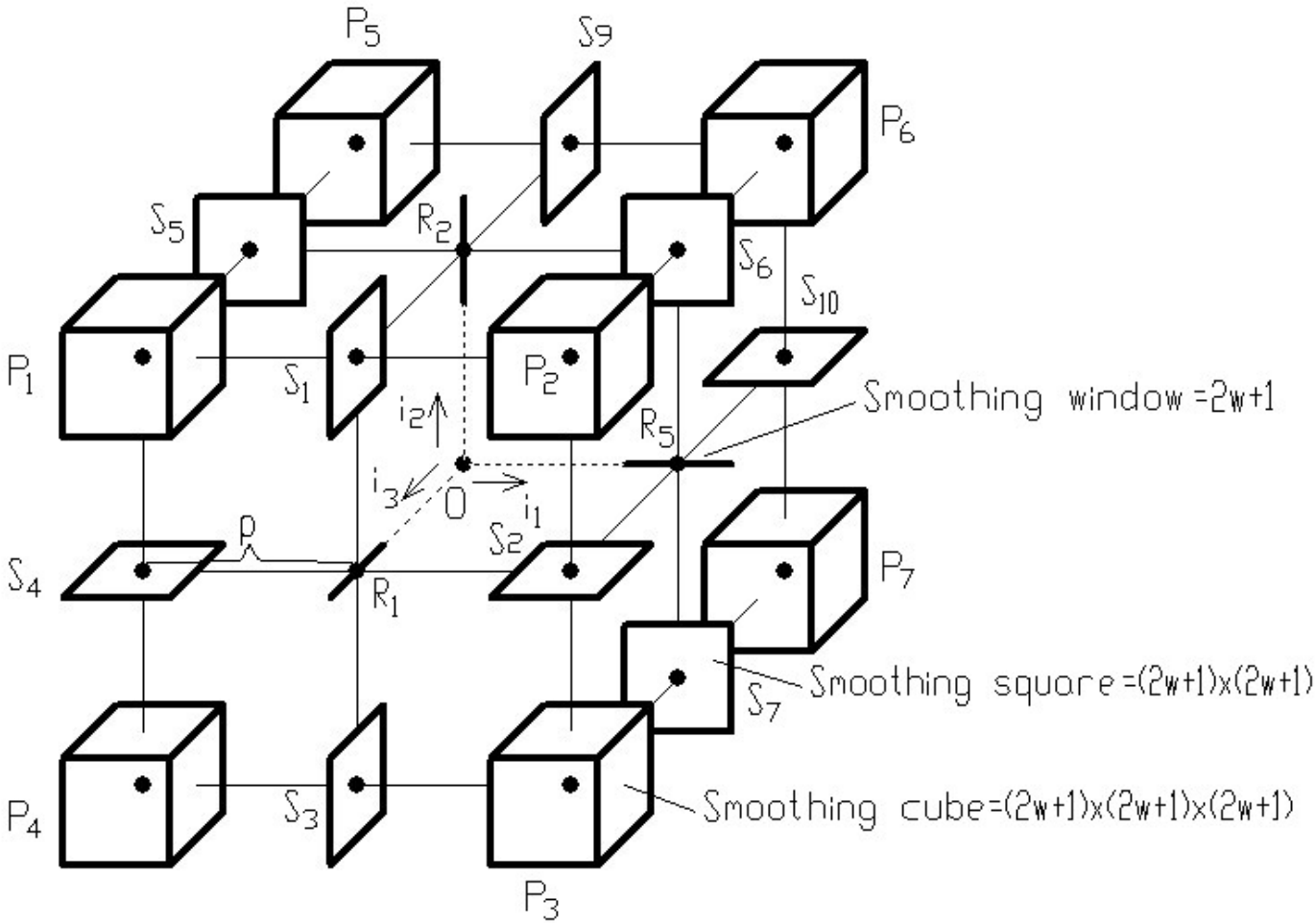

Fig. 14 Principle of background estimation in three-dimensional spectra with simultaneous smoothing.

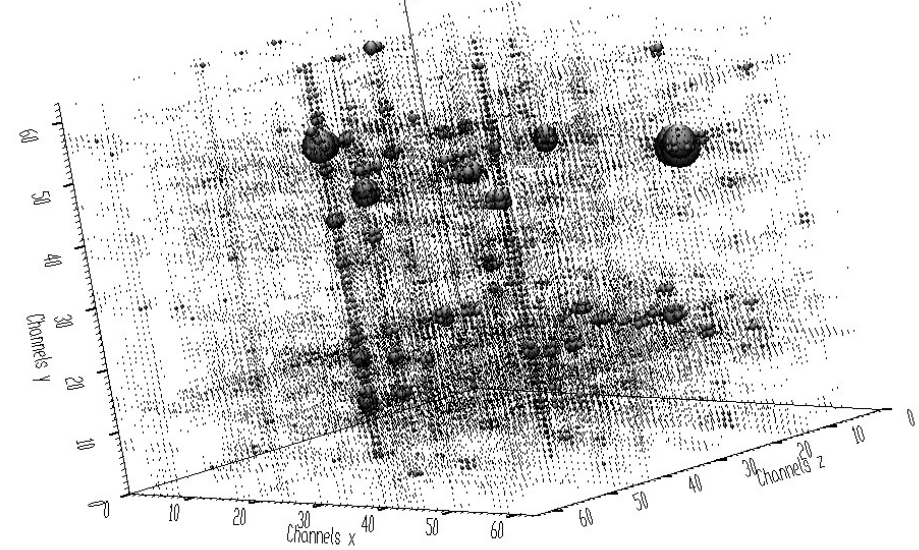

Fig. 15 Original 3-fold gamma-ray spectrum

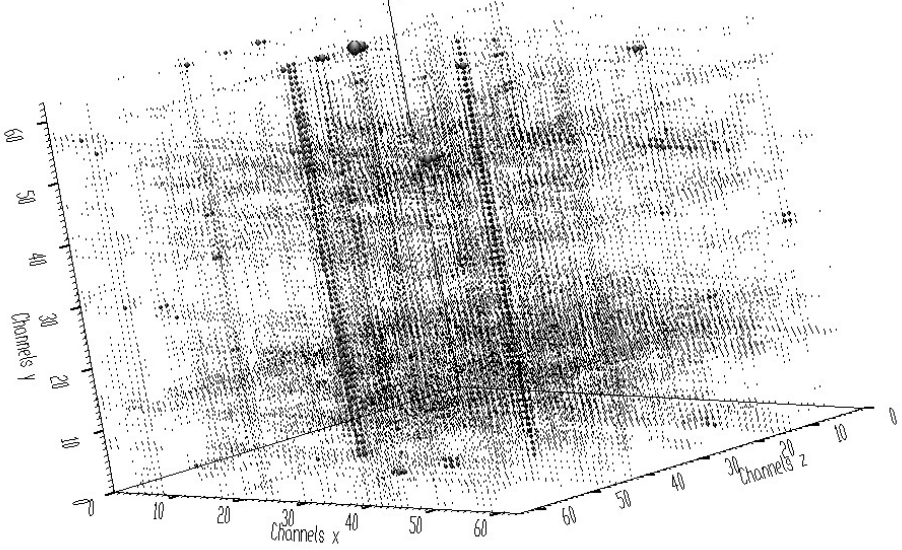

Fig. 16 Estimated background of the spectrum from Fig. 14

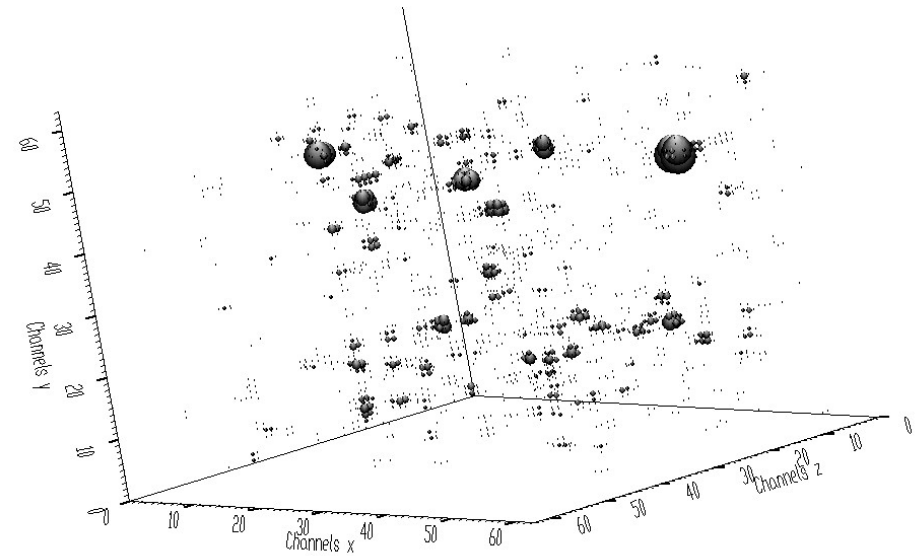

Fig. 17 Spectrum from Fig. 14 after subtraction of background

- the algorithm can be extended even to higher dimensions (4D, 5D)
- however there is a problem to determine the widths of objects (peaks) that should be preserved in the spectrum (not included into background) beforehand
- there can be different reasons that can deteriorate the estimate of the background

  -background can have very complicated shape

  -widths of peaks can differ (it is the case in high energy spectra)

  -due to enormous number of peaks (e.g. in gamma ray spectra) they can be overlapping and the cluster is much wider than the clipping window
- it motivated us to develop the algorithm for the estimation of peaks regions. Subsequently this can substantially improve the estimate of the background.

3. Estimation of peaks regions

•*Goal: Separation of peaks containing regions from peak-free regions*

•the estimation of widths of peaks regions is tightly connected with the localization and identification of peaks in the spectrum.

•it is assumed that peaks can be described by Gaussian function and that the background may be approximated by a linear function within short intervals

•sophisticated methods to locate the peaks and the intervals where they lie are based on the use of the convolution methods

•in [4-5] the authors propose a peak-search method based on two-pass convolution of the spectrum with the first derivative of the Gaussian. Besides of the identification of peaks the method can be utilized for the determination of peaks intervals and peaks areas.

•the method is based on convolution of the spectrum with the first derivative of an unnormalized Gaussian function (filter)

$$f'(t) = \left[\exp\left(-\frac{(t-x)^2}{2\delta^2} \right) \right]' = -\frac{t-x}{\delta^2} \exp\left(-\frac{(t-x)^2}{2\delta^2} \right) \qquad (1)$$

where δ is the width of the filter. The convolution with the Gaussian yields

$$c_1(x) = \int_{-\infty}^{\infty} G(t) f'(t) dt = \frac{\sqrt{2\pi}(x-a)\sigma\delta}{\left(\sigma^2+\delta^2\right)^{3/2}} \exp\left(-\frac{(x-a)^2}{2\left(\sigma^2+\delta^2\right)} \right) \qquad (2)$$

By repeating the operation of convolution of the filter (1) with (2) we get

$$c_2(x) = \frac{2\pi\delta^2\sigma\left((x-a)^2-\sigma^2-2\delta^2\right)}{\left(\sigma^2+2\delta^2\right)^{5/2}} \exp\left(-\frac{(x-a)^2}{2\left(\sigma^2+2\delta^2\right)} \right) \qquad (3)$$

•these two convolutions filter out the background and smooth the data by removing statistical disturbances. The function (3) has a negative minimum value at the position of peak. Using the crossing-over points of (3) with zero value one can determine

$$\sigma = \sqrt{(x_0 - a)^2 - 2\delta^2} \qquad (4)$$

where a is the position of the peak and x_0 is the appropriate zero crossing-over point.

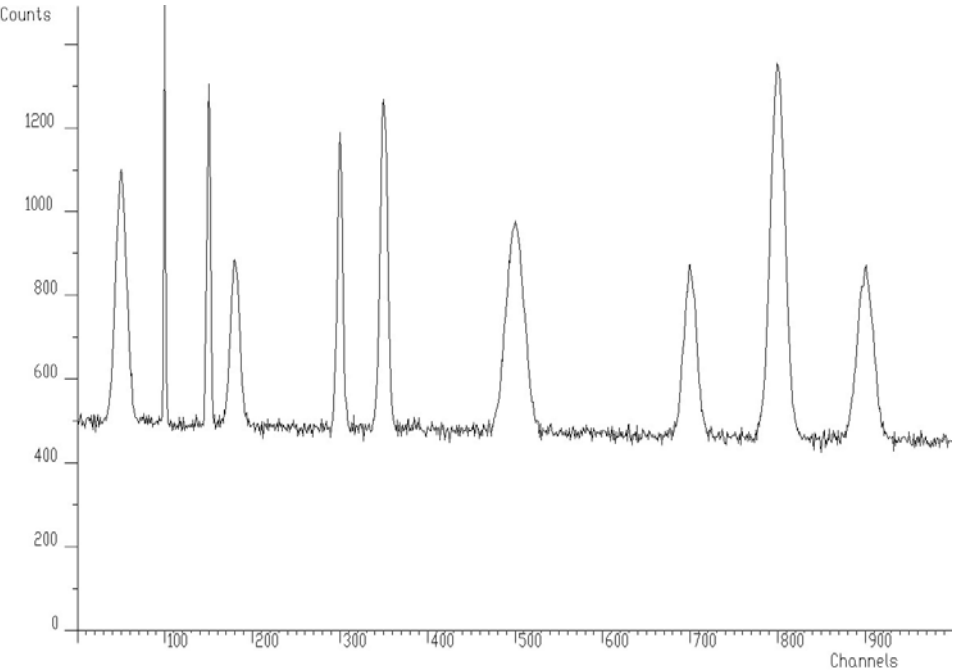

Fig. 18 A synthetic spectrum consisting of 10 peaks with σ from the range 1 - 10

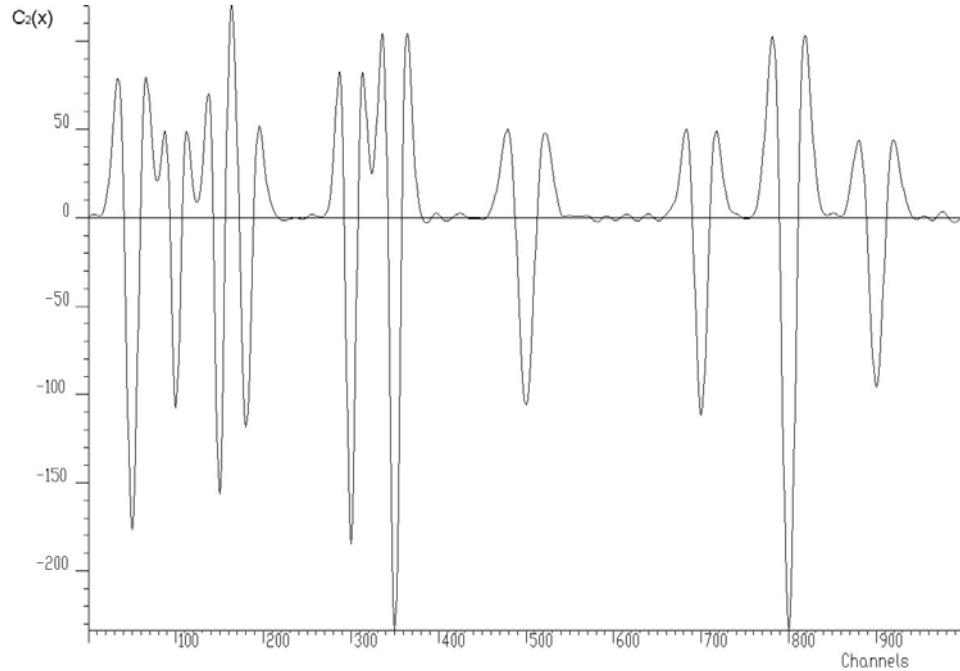

Fig. 19 The synthetic spectrum from Fig. 18 twice convolved with the first derivative of Gaussian function

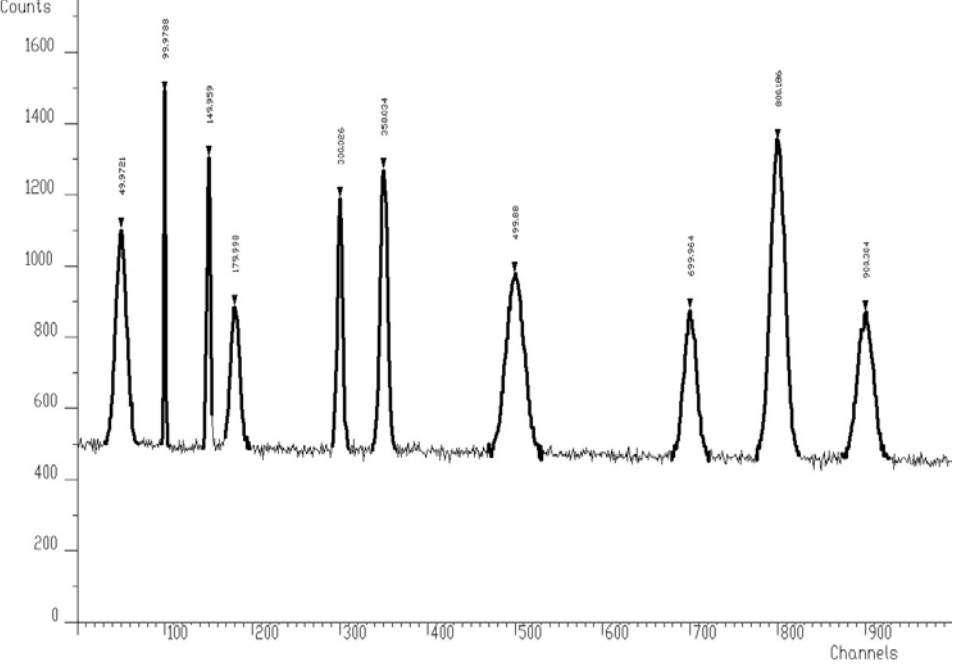

Fig. 20 The synthetic spectrum from Fig. 18 and determined peaks positions (denoted by markers) and their regions (drawn as thick lines)

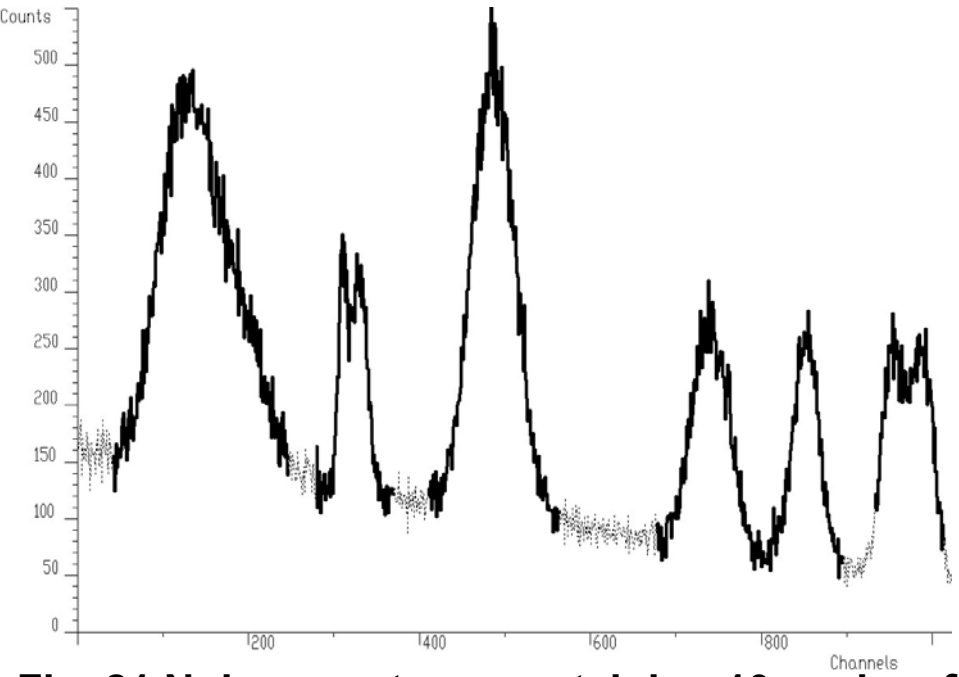

Fig. 21 Noisy spectrum containing 10 peaks of different widths with outlined peaks regions

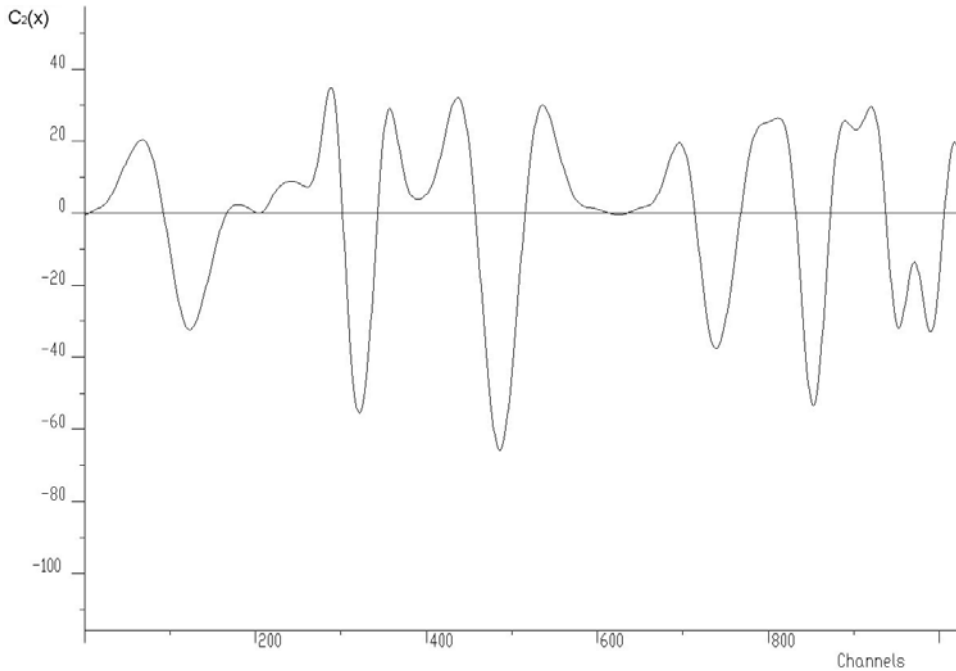

Fig. 22 Two-pass convolution of the spectrum from Fig. 21 with the first derivative of Gaussian function.

•In these two examples in several places the estimates of peaks regions are incorrect. We can conclude that this algorithm is not suitable for overlapping doublets or multiplets.
•therefore it is unavoidable to improve this algorithm. During the study of the examples given above we observed that the estimate of σ for appropriate peaks depends sensitively on chosen δ.
•when studying the function (3) it can be proved that it reaches its maximum value for δ=σ.
•further let us define inverted positive double pass first derivative of the Gaussian

$$g(x) = \begin{cases} -c_2(x) & \text{if} \quad c_2(x) < 0 \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

•the basic idea for the new proposed algorithm is to scan the whole range of possible σ, i.e., to generate series of functions g(x) to form a matrix M(x,δ) from them and subsequently to find local maximums.
•the positions of the maximums give both positions of peaks and their σ.
•it was found that σ on both sides estimated according to (4) should satisfy the condition

$$\sigma_l \geq \delta \quad \text{or} \quad \sigma_r \geq \delta \tag{6}$$

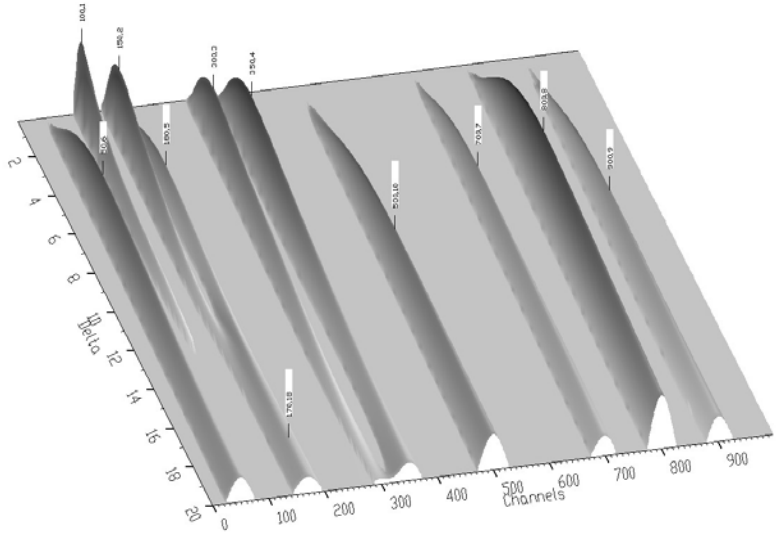

Fig. 22 Two-dimensional matrix M(x,δ).

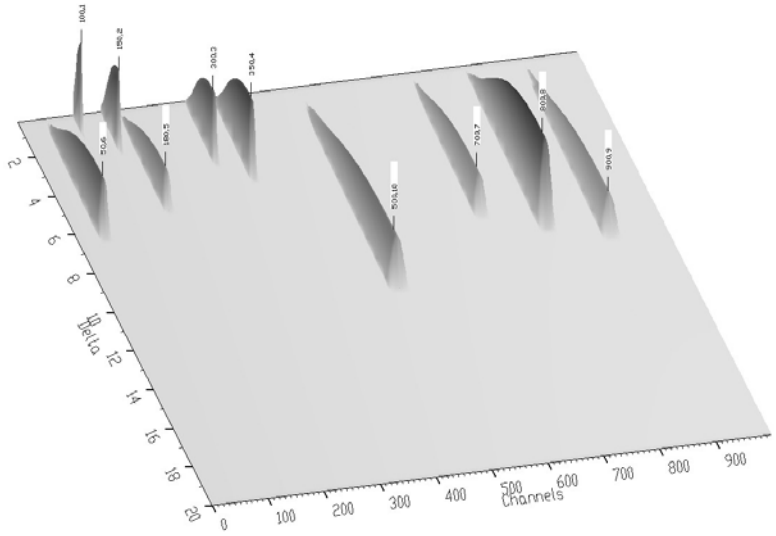

Fig. 23 The same matrix as in Fig. 22 after cutting-off ridges according to (6)

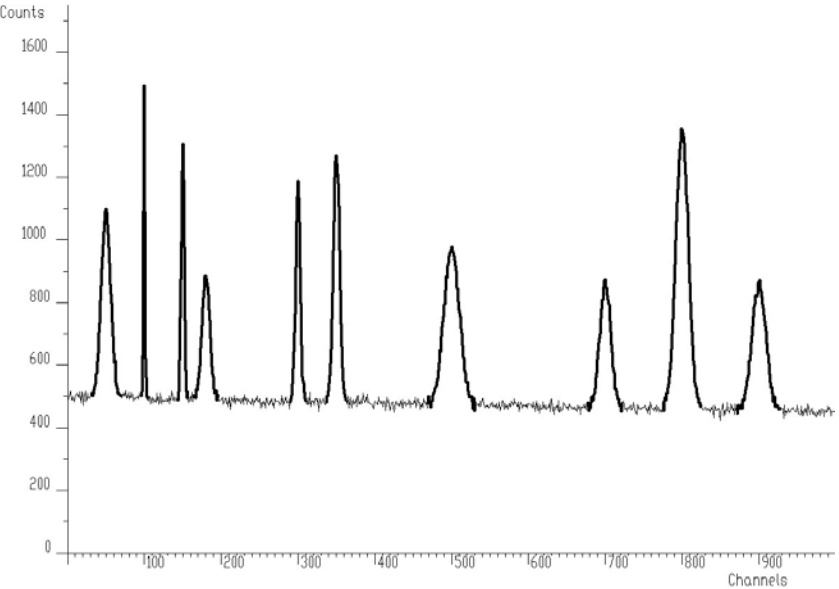

Fig. 24 Spectrum from Fig. 20 with regions estimated using a new proposed algorithm.

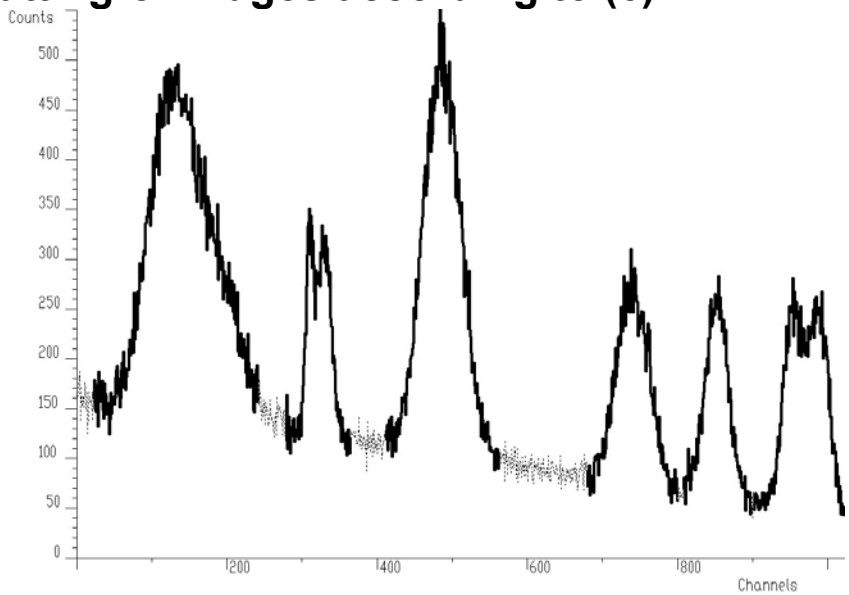

Fig. 25 Spectrum from Fig. 21 with regions estimated using a new proposed algorithm.

Background estimation algorithm with clipping window adaptive to peak regions widths

•let us assume we have determined peak regions and peak-free regions in the spectrum **y** using the algorithm derived in the previous section. Let us assume we have identified *k* regions. We construct vector **r**

$$\mathrm{r}=\left[0,0,\ldots,0,w_1,w_1,\ldots,w_1,0,\ldots,0,w_2,w_2,\ldots,w_2,0,\ldots,0,w_k,w_k,\ldots,w_k,0,0,\ldots,0\right]^{\mathrm{T}} \quad (7)$$

where zero values stand in the positions in peak-free regions and the value w_j represents width of *j*-th region and stands in its position. Let us define

$$m=\max\left\{w_j\right\}\ j\in\langle 1,k\rangle$$

•then the SNIP algorithm with clipping window adaptive to peak regions widths can be defined as follows:

$$y_p(i)=\begin{cases}\min\left\{y_{p-1}(i),\frac{1}{2}\left[y_{p-1}(i+p)+y_{p-1}(i-p)\right]\right\} & \text{if} \quad p\le r(i)\\ y_p(i) & \text{otherwise,}\end{cases} \qquad p\in\langle 1,m\rangle$$

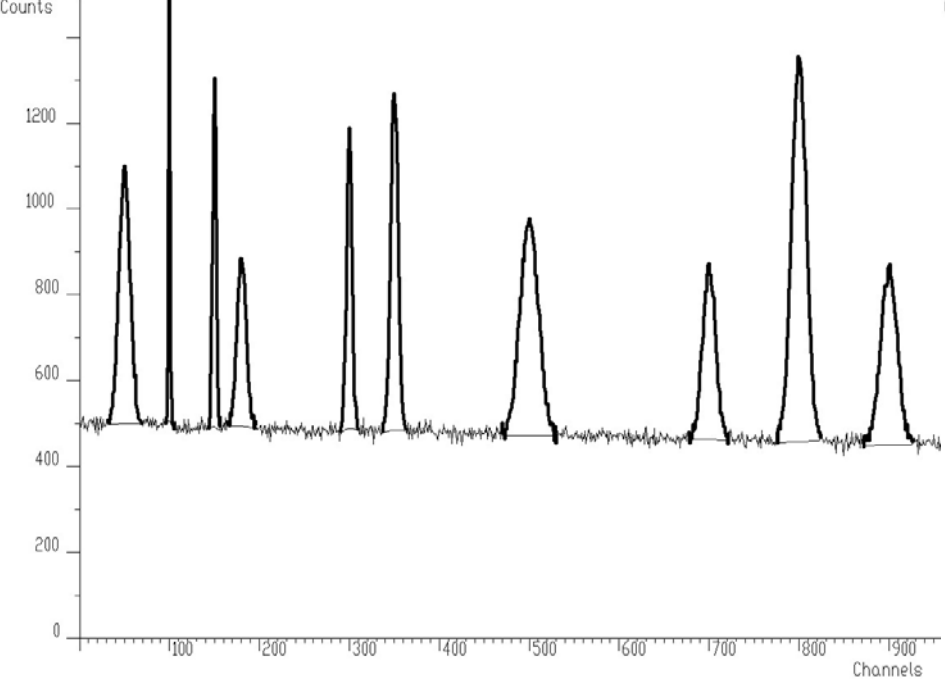

Fig. 26 Synthetic spectrum from Fig. 20 with background estimated using SNIP algorithm with decreasing clipping window adjusted to peak regions widths

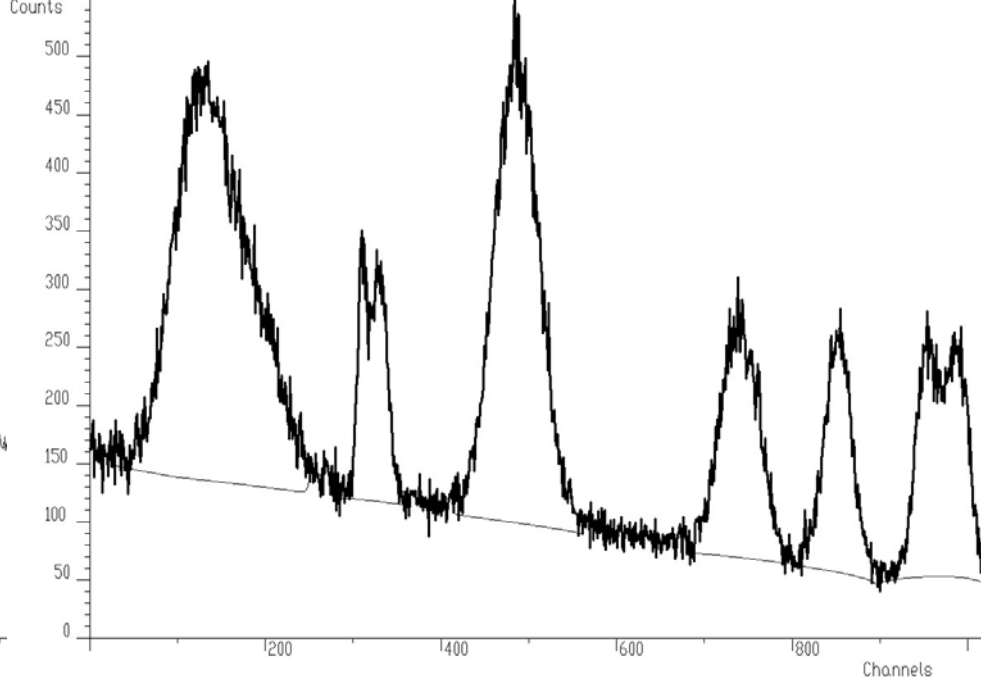

Fig. 27 Spectrum from Fig. 21 with background estimated using SNIP algorithm with simultaneous smoothing with decreasing clipping window adjusted to peak regions widths.

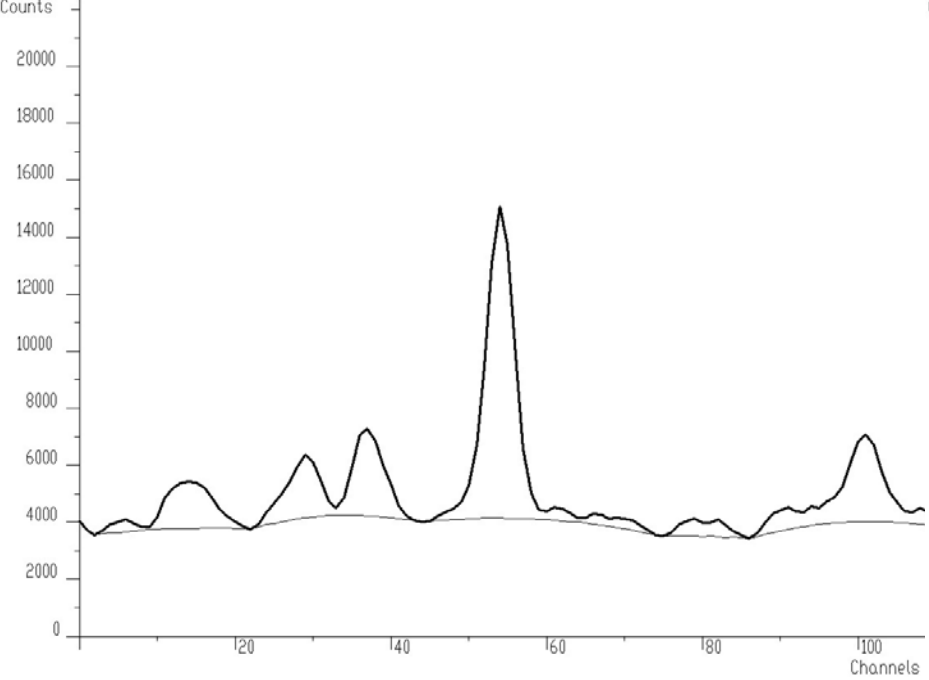

Fig. 28 Experimental γ-ray spectrum with estimated background using fixed width of clipping window given as user defined parameter

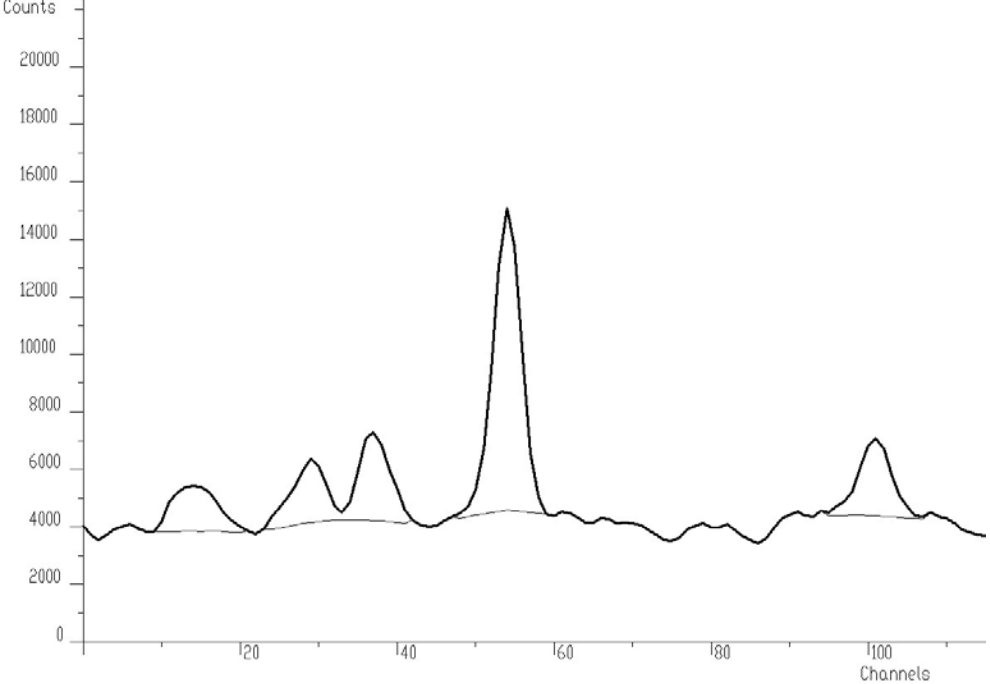

Fig. 29 Experimental γ-ray spectrum with estimated background using clipping window with width automatically adjustable to the widths of peak regions.

4. Deconvolution

Goal: Improvement of the resolution in spectra

•the accuracy and reliability of the analysis depend critically on the treatment in order to resolve strong peak overlaps.

•the peaks as the main carrier of spectrometric information are very frequently positioned close to each other.

•the extraction of the correct information out of the spectra sections, where due to the limited resolution of the equipment, signals coming from various sources are overlapping, is a very complicated problem.

•deconvolution and restoration are the names given to the endeavor to improve the resolution of an experimental measurement by mathematically removing the smearing effects of an imperfect instrument, using its resolution function.

•to devise reliable methods for doing this has been a long-term endeavor by many scientists and mathematicians.

•stationary discrete system that satisfies the superposition principle can be described by convolution sum

$$y(i) = \sum_{k=0}^{i} x(k)h(i-k) + n(i) = x(i) * h(i) + n(i) \qquad i = 0,1,\ldots,N-1 \qquad (8)$$

or in matrix form

$$\boldsymbol{y} = H\boldsymbol{x} + \boldsymbol{n}$$

Least Square Solution. To find least square solution of above given system of linear equations the functional $\| H\hat{\boldsymbol{x}} - \boldsymbol{y} \|^2$ should be minimized

least squares estimate of the solution is

$$\hat{\boldsymbol{x}} = (H^T H)^{-1} \boldsymbol{H}^T \boldsymbol{y} \qquad (8a)$$

Illustrative example

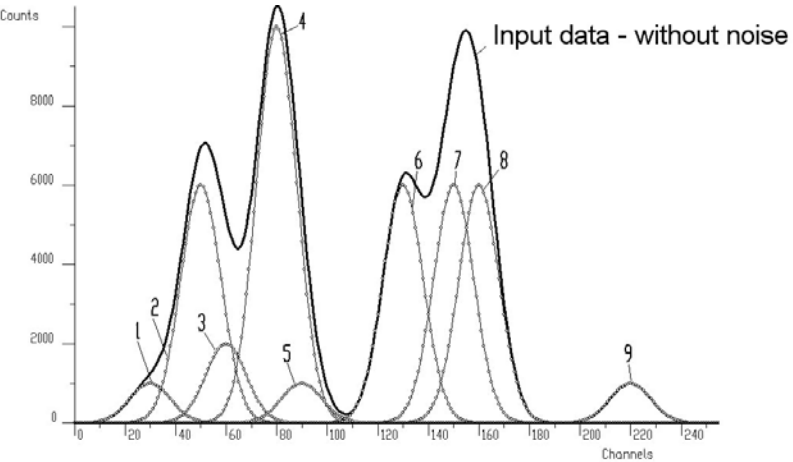

Fig. 30 Example of synthetic spectrum (without noise) composed of 9 Gaussians

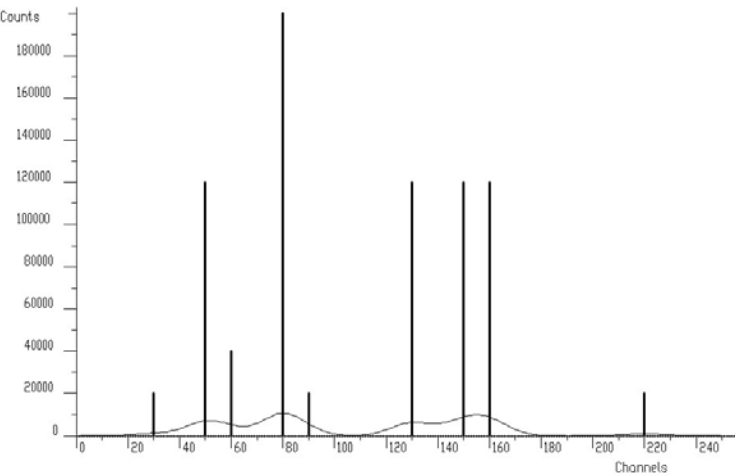

Fig. 32 Original (thin line) and deconvolved spectrum (bars)

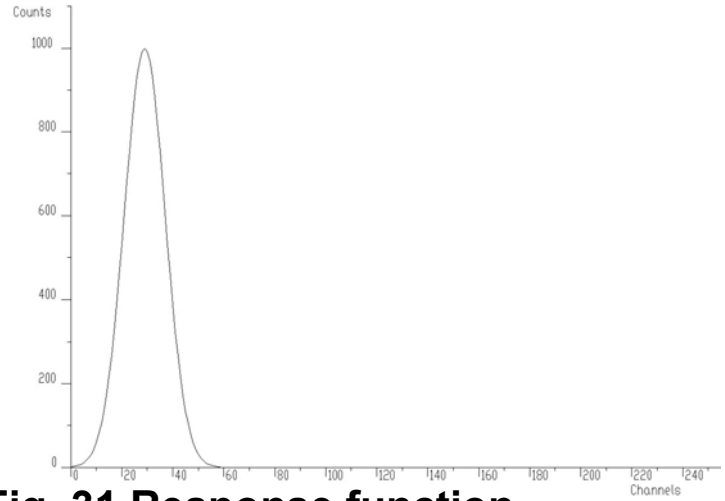

Fig. 31 Response function

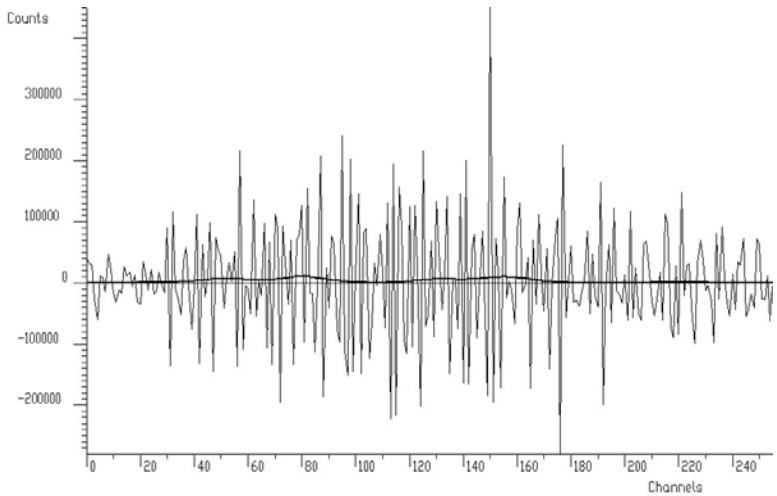

Fig. 33 Original spectrum from Fig. 31 + 1% of noise of the amplitude of small peaks (#1, #5 and #9 - thick line) and deconvolved spectrum (strong oscillations) using unconstrained solution (Eq. 3) of the Toeplitz system of linear equations (thin line).

Three types of regularization methods are very often used [6]:

•smoothing,
•constraints imposition (for example only non-negative data are accepted),
•choice of a prior information probability function - Bayesian statistical approach.

Tikhonov-Miller regularization the functional

$$|| H\hat{\boldsymbol{x}} - \boldsymbol{y} ||^2 + \alpha || Q\hat{\boldsymbol{x}} ||^2 \qquad (8b)$$

is minimized. The solution can be obtained by solving the equation

$\hat{\boldsymbol{x}} = \left(H^T H + \alpha Q^T Q\right)^{-1} H^T \boldsymbol{y}$ Q, α being the regularization matrix and parameter, respectively.

• zero-th order or Tikhonov regularization

$$\hat{\boldsymbol{x}} = \left(H^T H + \alpha\right)^{-1} H^T \boldsymbol{y}$$

Together with chi-square also the sum of squares of elements of the estimated vector is minimized.

Riley Algorithm. This algorithm is commonly called iterated Tikhonov regularization. To obtain smoother solution one may use algorithm of Tikhonov-Miller regularization with iterative refinement

$$\boldsymbol{x}^{(n+1)} = \boldsymbol{x}^{(n)} + \left(H^T H + \alpha Q^T Q\right)^{-1} \left(H^T \boldsymbol{y} - H^T H \boldsymbol{x}^{(n)}\right) \qquad \boldsymbol{x}^{(0)} = 0$$

•however, because of its iterative nature, the Riley algorithm lends itself to another type of regularization, so called Projections On Convex Sets – POCS. There exist many regularization methods based on this kind of regularization [7-9]. It means to set all negative elements to zero after each iteration.

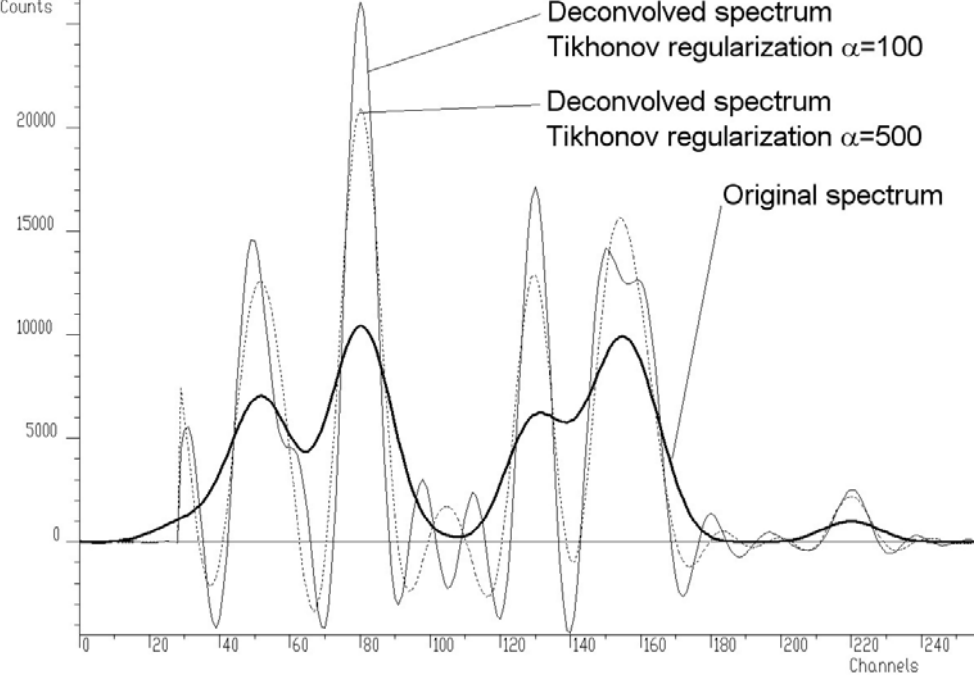

Fig. 34 Solution obtained using Tikhonov method of regularization.

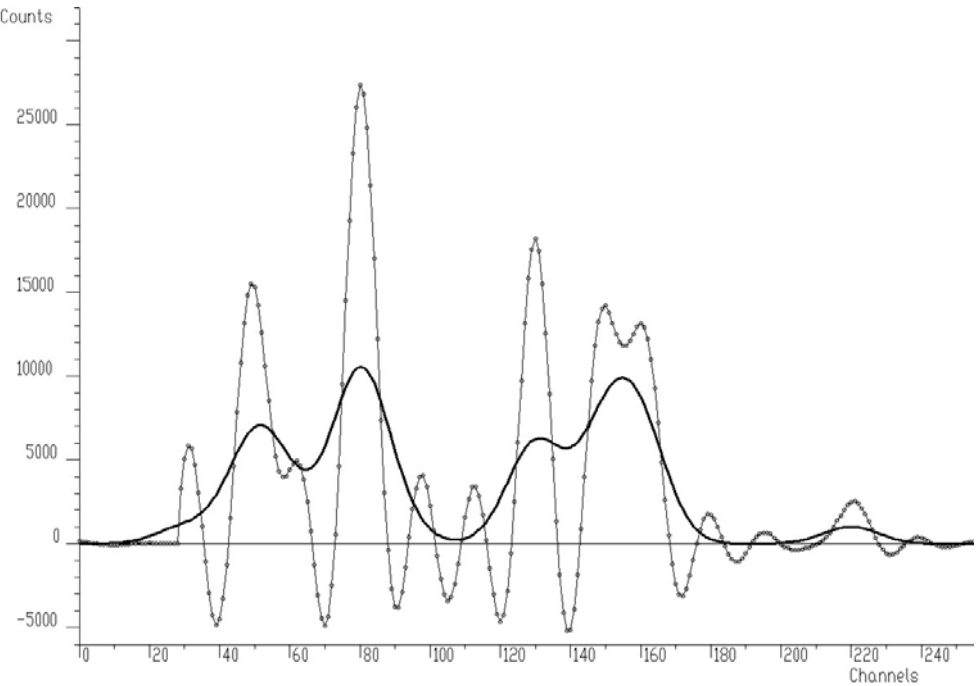

Fig. 35 Illustration of Riley algorithm of deconvolution - thick line is original spectrum, thin line represents spectrum after deconvolution.

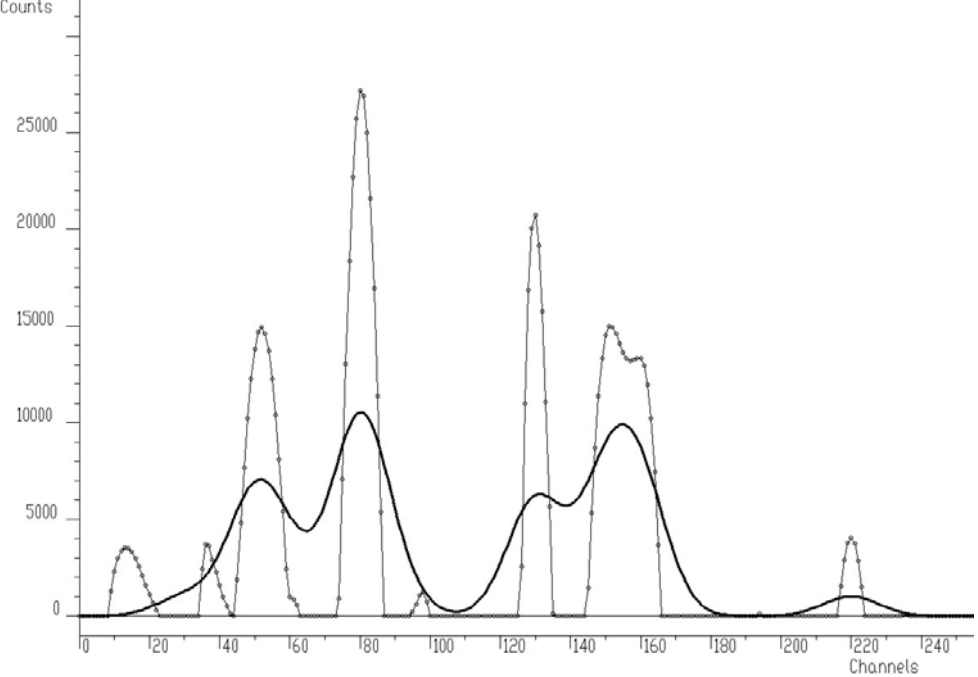

Fig. 36 Riley deconvolution with POCS regularization - thick line is original spectrum, thin line represents spectrum after deconvolution.

Van Cittert Algorithm. The basic form of Van Cittert [10] algorithm for discrete convolution system is

$$\boldsymbol{x}^{(n+1)} = \boldsymbol{x}^{(n)} + \mu\left(H^T HH^T \boldsymbol{y} - H^T HH^T H\boldsymbol{x}^{(n)}\right) = \boldsymbol{x}^{(n)} + \mu\left(\boldsymbol{y'} - A\boldsymbol{x}^{(n)}\right) \tag{9}$$

μ is the relaxation factor.

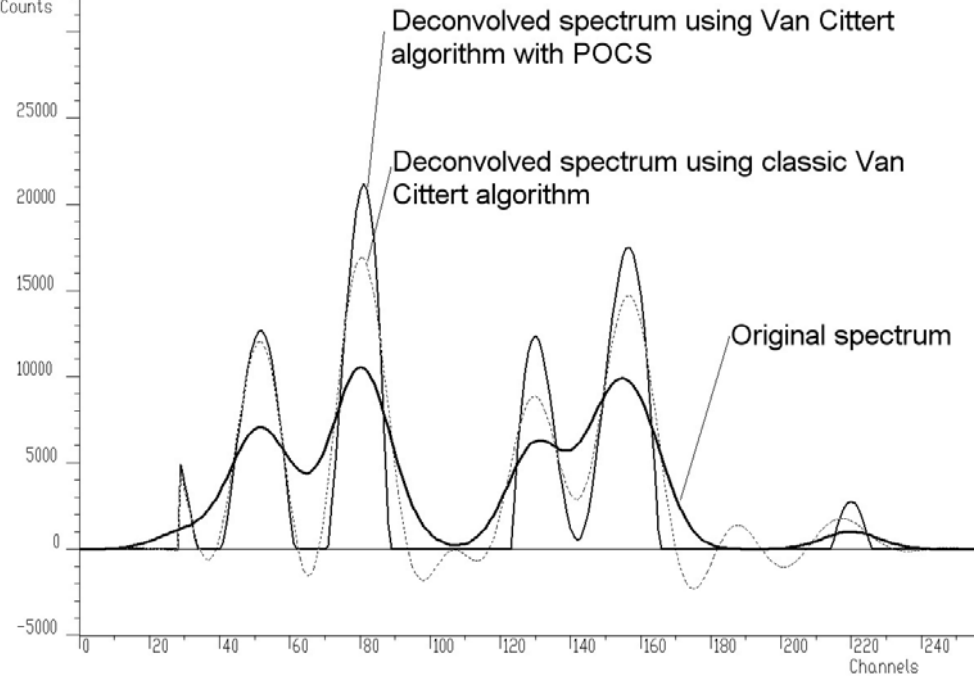

Fig. 37 Original spectrum (thick line), deconvolved spectrum using Van Cittert algorithm (without regularization) and deconvolved spectrum using Van Cittert algorithm and regularized via POCS method.

Gold Algorithm. Further, if we choose the local variable relaxation factor $\mu_i = \dfrac{x^{(n)}(i)}{\sum_{m=0}^{M-1} A_{im} x^{(n)}(m)}$

and we substitute it into Eq. 9 we get

$$x^{(n+1)}(i) = \frac{y'(i)}{\sum_{m=0}^{M-1} A_{im} x^{(n)}(m)} x^{(n)}(i) \qquad (10)$$

This is the Gold deconvolution algorithm [11- 13]. Its solution is always positive when the input data are positive, which makes the algorithm suitable for the use for naturally positive definite data, i.e., spectroscopic data.

M. Morháč, J. Kliman, V. Matoušek, M. Veselský and I. Turzo, Efficient one and two dimensional Gold deconvolution and its application to gamma-ray spectra decomposition, Nucl. Instrum. Methods Phys. Res., A 401 (2-3) (1997) pp. 385-408.

M. Morháč, Deconvolution methods and their applications in the analysis of gamma-ray spectra, Nucl. Instrum. Methods Phys. Res., Sect. A 559 (1) (2006), pp. 119-123.

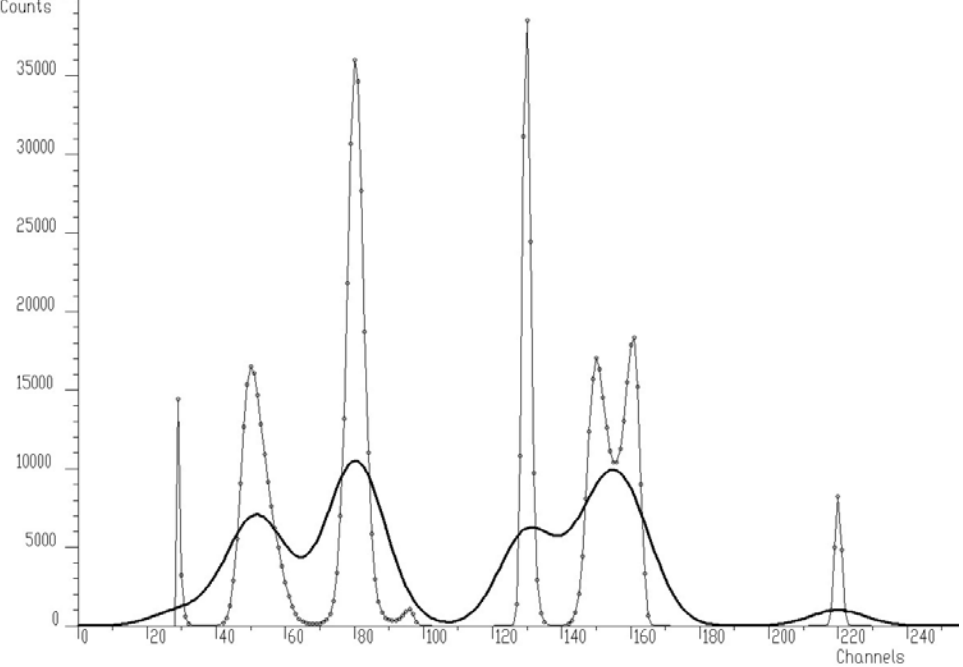

Fig. 39 Original spectrum (thick line) and deconvolved spectrum using Gold algorithm (thin line) after 10000 iterations. Channels are shown as small circles.

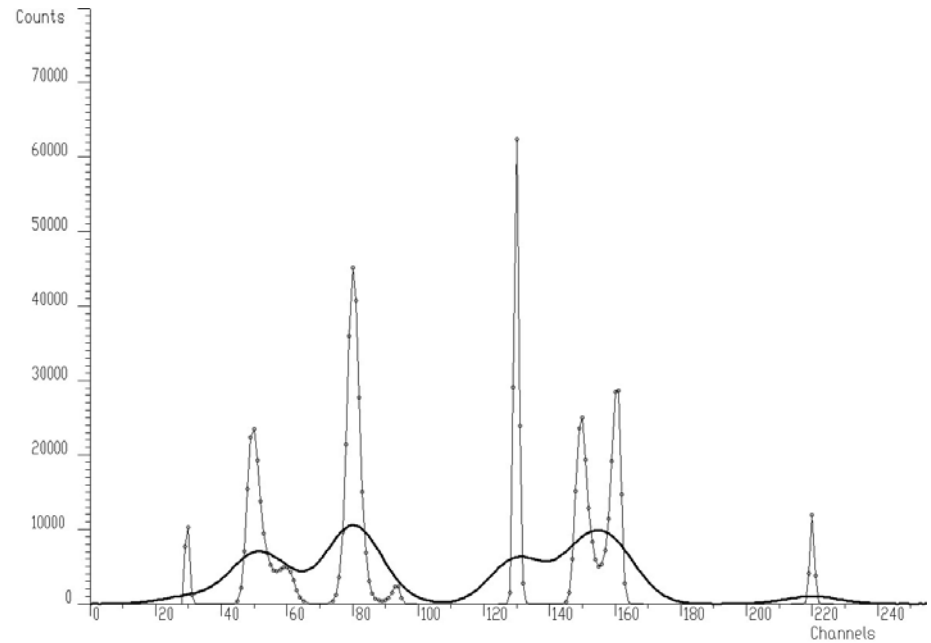

Fig. 40 Original spectrum (thick line) and deconvolved spectrum using Gold algorithm (thin line) after 50000 iterations.

Richardson-Lucy Algorithm. Richardson-Lucy like algorithms [14, 15] use a statistical model for data formation and are based on the Bayes formula [16]. The Bayesian approach consists of constructing the conditional probability density relationship

$$p(x \mid y) = \frac{p(y \mid x)p(x)}{p(y)}$$

The Bayes solution is found by maximizing the right part of the equation. The maximum likelihood (ML) solution maximizes the density p(y|x) over x. For discrete data the algorithm has the form

$$x^{(n+1)}(i) = x^{(n)}(i)\sum_{j=0}^{N-1} h(j,i)\frac{y(j)}{\sum_{k=0}^{M-1} h(j,k)x^{(n)}(k)} \qquad i \in \langle 0, M-1 \rangle \tag{11}$$

This iterative method forces the deconvoluted spectra to be non-negative. The Richardson-Lucy iteration converges to the maximum likelihood solution for Poisson statistics in the data. It is also sometimes called the expectation maximization (EM) method.

***Muller Algorithm*.** Setting the probability to have Gauss statistics, another deconvolution algorithm based on Bayes formula was derived by Muller (1997)

$$x^{(n+1)}(i) = x^{(n)}(i)\frac{\sum_{j=0}^{N-1} h(j,i)y(j)}{\sum_{j=0}^{N-1} h(j,i)\sum_{k=0}^{M-1} h(j,k)x^{(n)}(k)} \qquad i \in \langle 0, M-1 \rangle \tag{12}$$

This algorithm again applies the positivity constraint to the decovolved data.

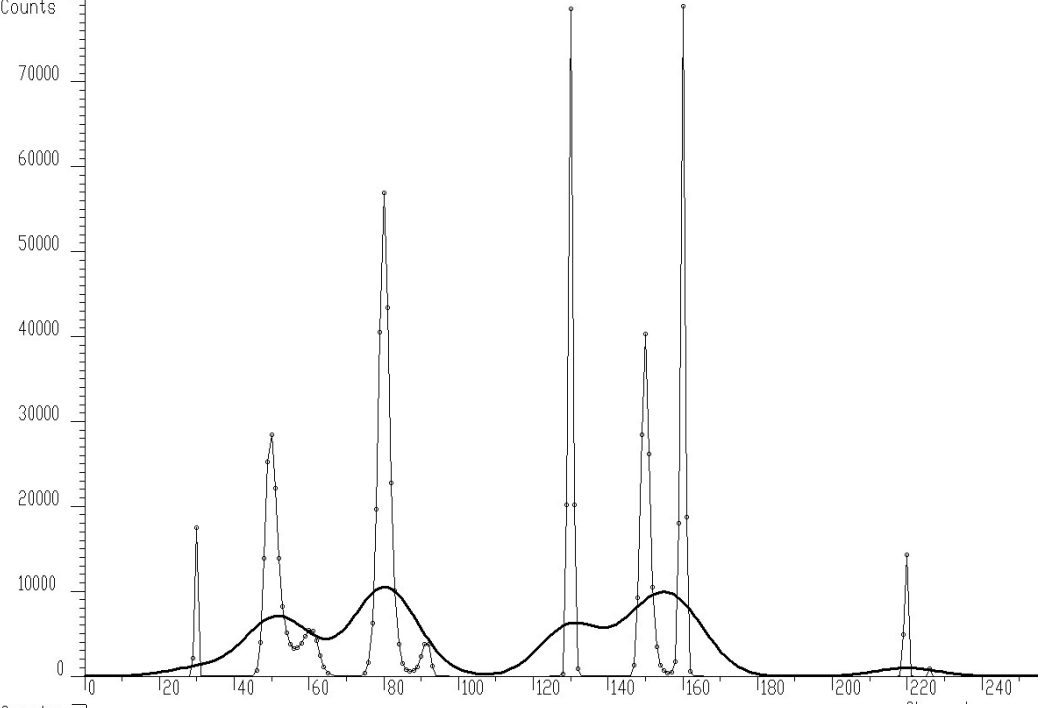

Fig. 41 Original spectrum (thick line) and deconvolved spectrum using Richardson-Lucy algorithm (thin line) after 50 000 iterations.

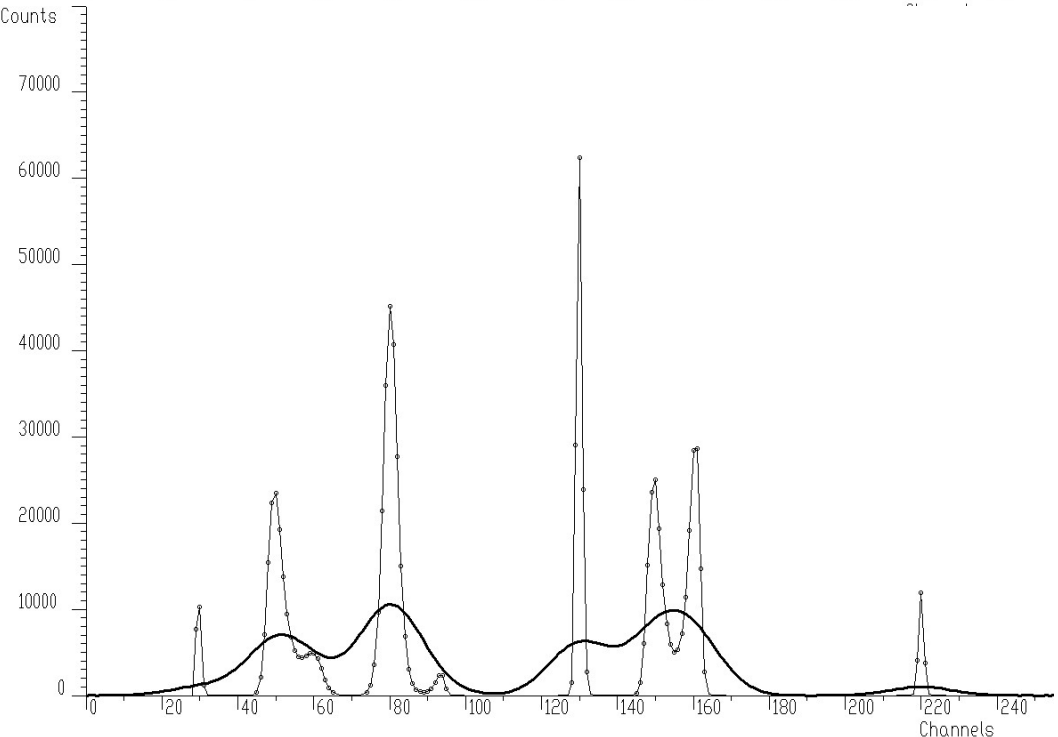

Fig. 42 Original spectrum (thick line) and deconvolved spectrum using Muller algorithm (thin line) after 50000 iterations.

Maximum A Posteriori Deconvolution Algorithm. The maximum a posteriori (MAP) solution maximizes over x the product $p(y \mid x)p(x)$ For discrete data the algorithm has the form

$$x^{(n+1)}(i) = x^{(n)}(i)\exp\left\{\sum_{j=0}^{N-1} h(j,i)\left[\frac{y(j)}{\sum_{k=0}^{M-1} h(j,k)x^{(n)}(k)} - 1\right]\right\} \tag{13}$$

Positivity of the solution is assured by the exponential function. Moreover the non-linearity permits superresolution [17].

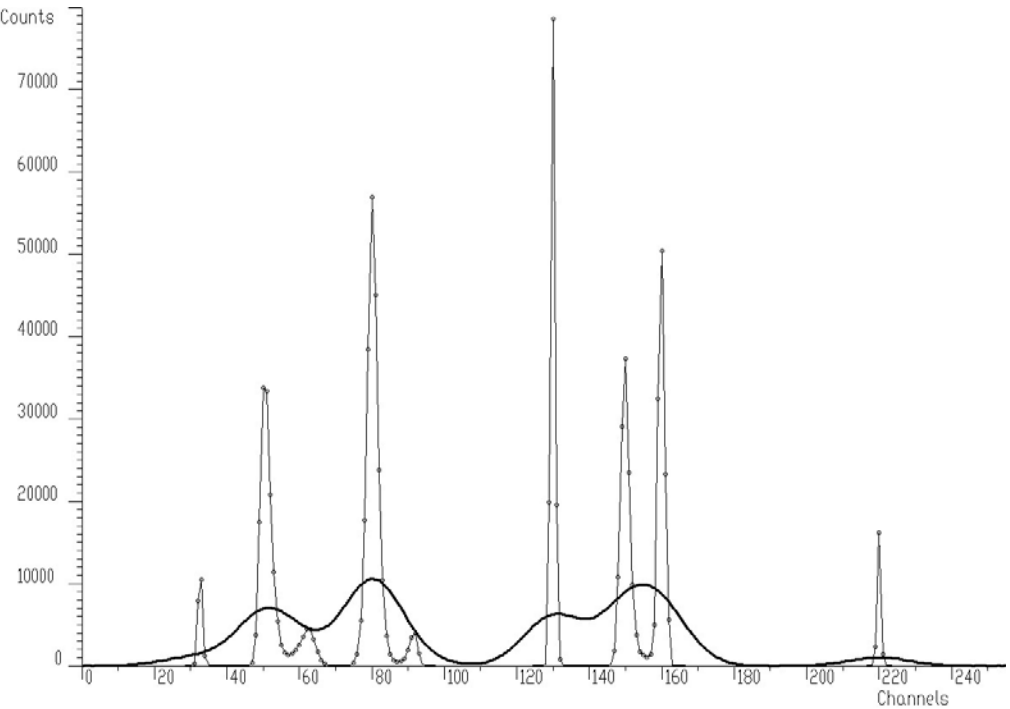

Fig. 43 Original spectrum (thick line) and deconvolved spectrum using MAP algorithm (thin line) after 50000 iterations.

Boosted Deconvolution Algorithm [13].

There exist many problems inherent to the solution of the problem of deconvolution and decomposition in general:
•computational complexity,
•influence of the errors due to the measured noise,
•ill conditionality of systems,
•regularization of the solution.

Iterative positive definite deconvolutions (Gold, Richardson-Lucy, Muller and MAP.) converge to stable states. It is useless to increase the number of iterations, the result obtained practically does not change. Instead of it we can stop iterations, apply a boosting operation and repeat this procedure. Boosting operation should decrease sigma of peaks. Then, the algorithm of boosted Gold, Richardson-Lucy or MAP deconvolution is as follows:

1. Set the initial solution $\boldsymbol{x}^{(0)} = [1,1,...,1]^T$
2. Set required number of repetitions R and iterations L
3. Set the number of repetitions r=1
4. According to either Eq. 10 (Gold), or Eq. 11 (Richardson-Lucy) or Eq. 12 (Muller) or Eq. 13 (MAP) for $n = 0,1,...,L-1$ find solution $\boldsymbol{x}^{(L)}$
5. If r=R stop the calculation, else
    a. apply boosting operation, i.e., set $x^{(0)}(i) = \left[x^{(L)}(i) \right]^p \quad i = 0,1,...,N-1$
    p is boosting coefficient >0, $p \approx 1.1-1.2$
    b. r=r+1
    c. continue in 4.

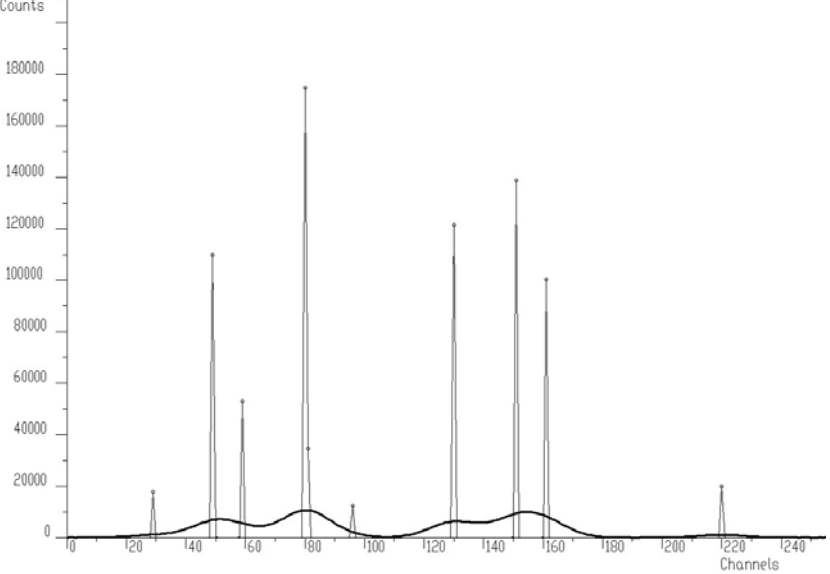

Fig. 44 Original spectrum (thick line) and deconvolved spectrum using boosted Gold algorithm (thin line) after 200 iterations and 50 repetitions (p=1.2).

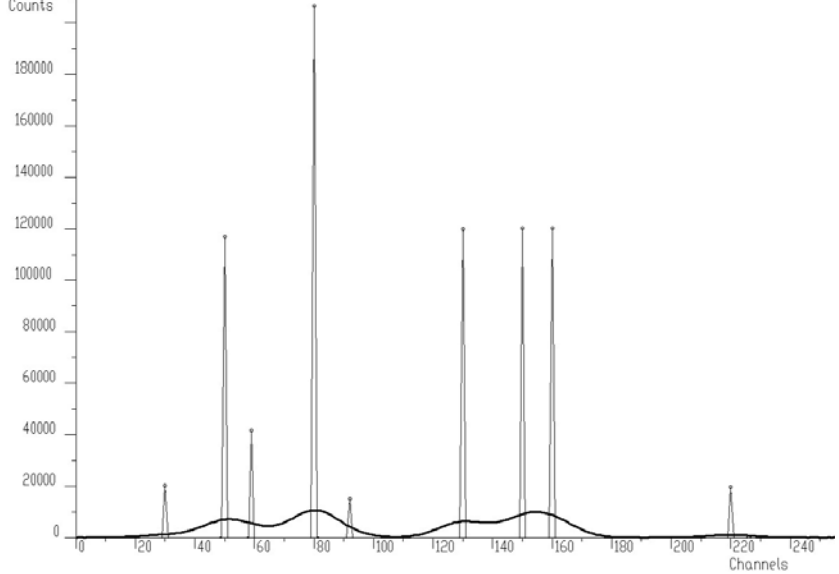

Fig. 45 Original spectrum (thick line) and deconvolved spectrum using boosted Richardson-Lucy algorithm (thin line) after 200 iterations and 50 repetitions

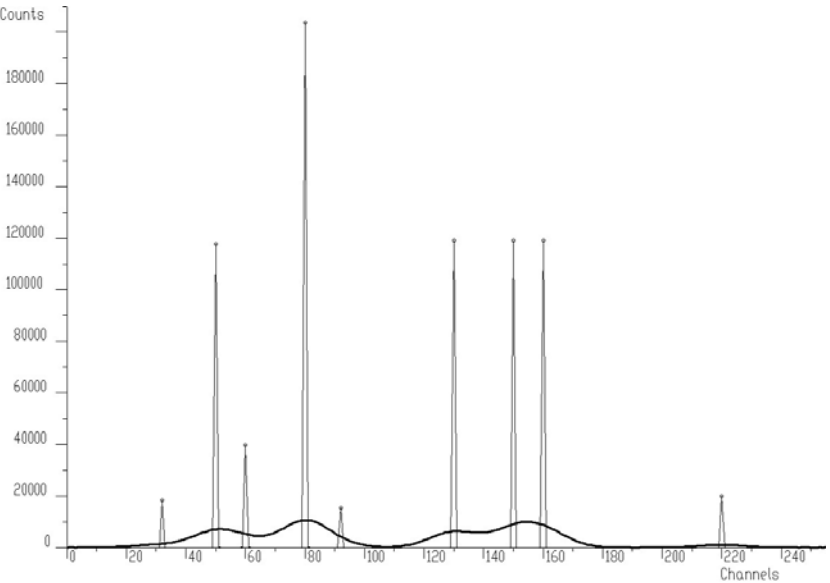

Fig. 46 Original spectrum (thick line) and deconvolved spectrum using boosted MAP algorithm (thin line) after 200 iterations and 50 repetitions

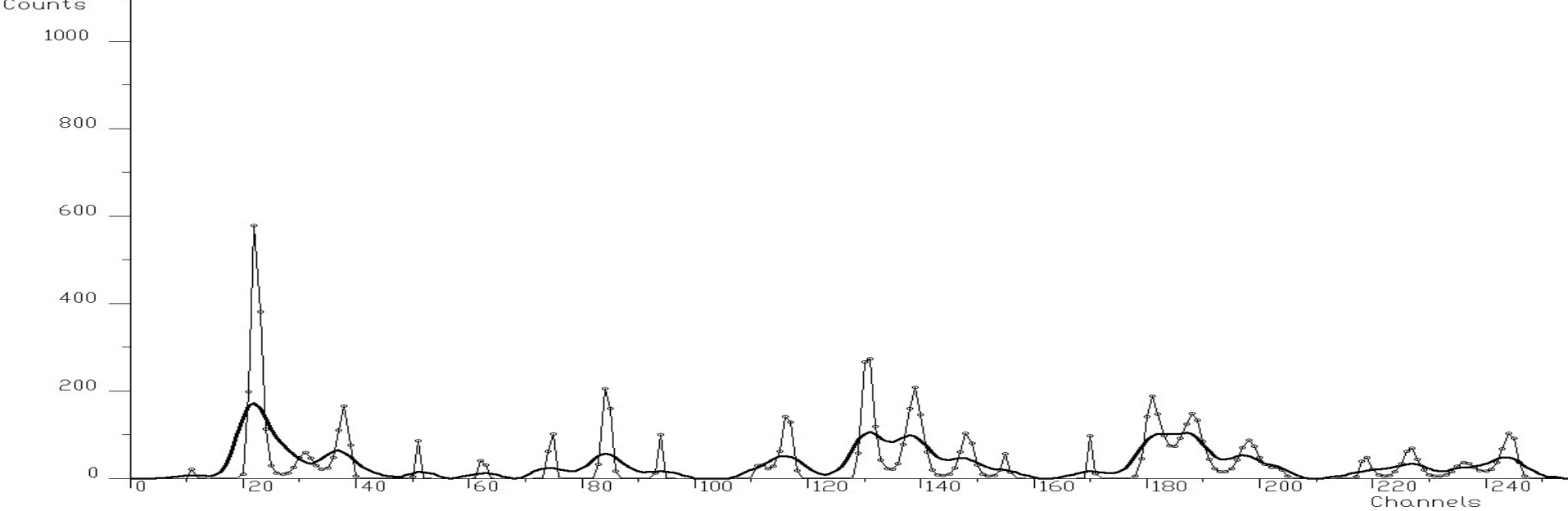

Fig. 47 Original gamma-ray spectrum (thick line) and deconvolved spectrum using classic Gold algorithm (thin line) after 10 000 iterations.

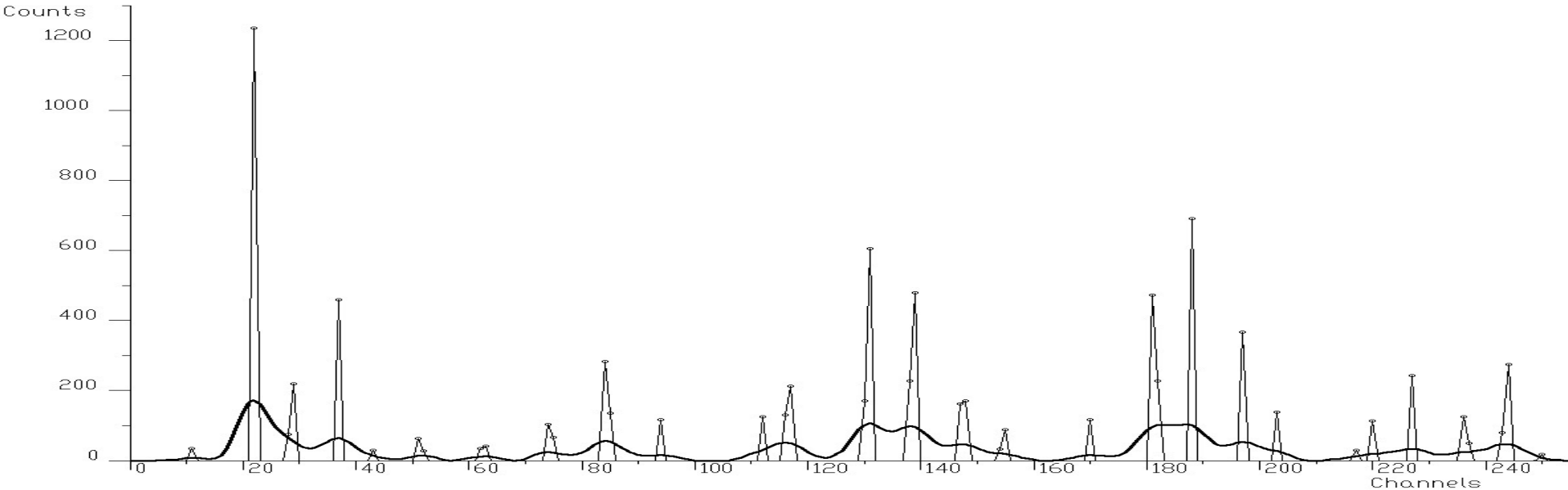

Fig. 48 Gamma-ray spectrum (thick line) and deconvolved spectrum using boosted Gold algorithm (thin line) after 200 iterations and 50 repetitions (p=1.2)

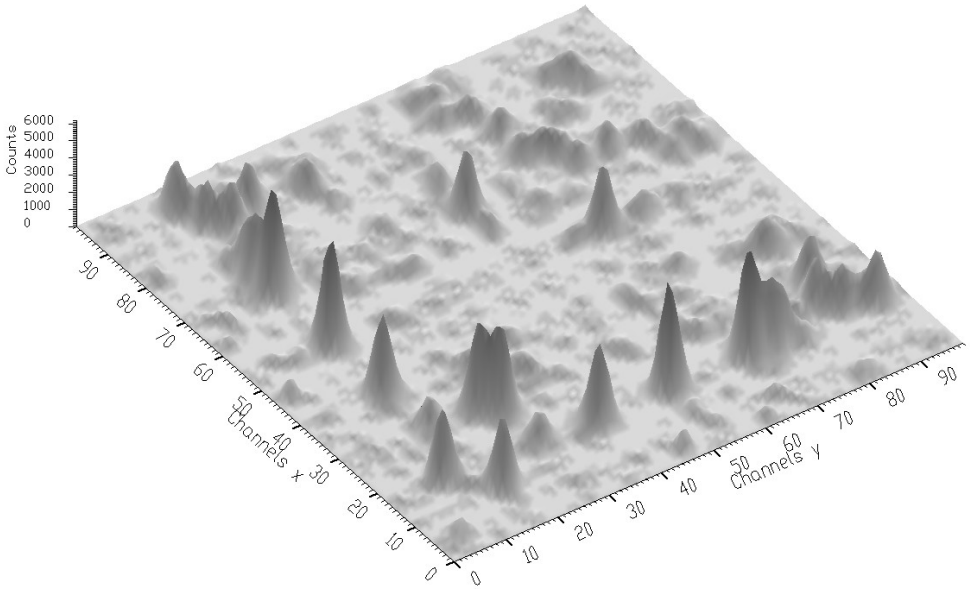

Fig. 49 Experimental gamma-gamma-ray spectrum (after background elimination).

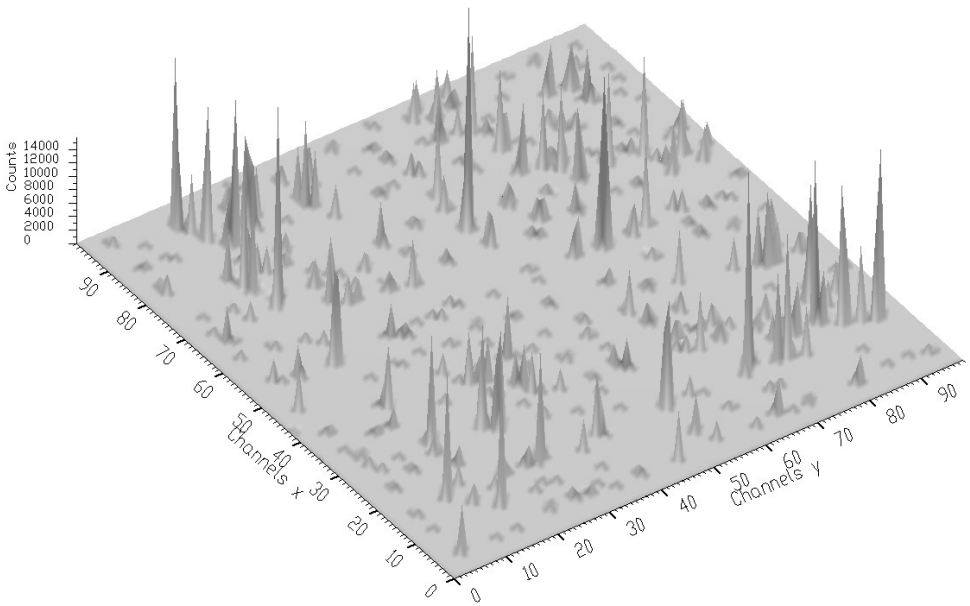

Fig. 50 Spectrum from Fig. 49 after boosted Gold deconvolution (50 iterations repeated 20 times).

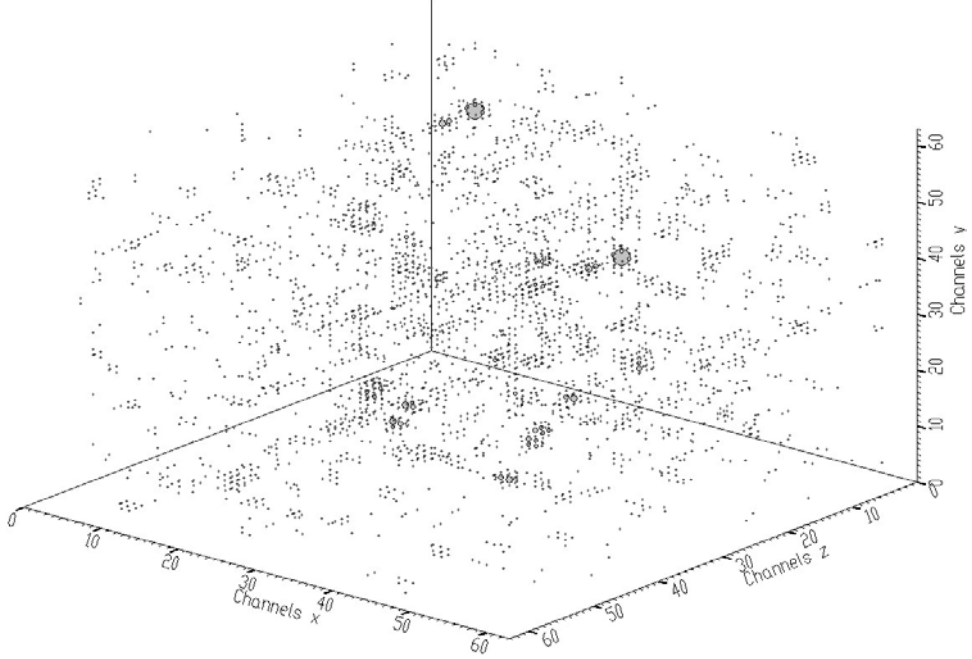

Fig. 51 Experimental gamma-gamma-gamma-ray spectrum (after background elimination).

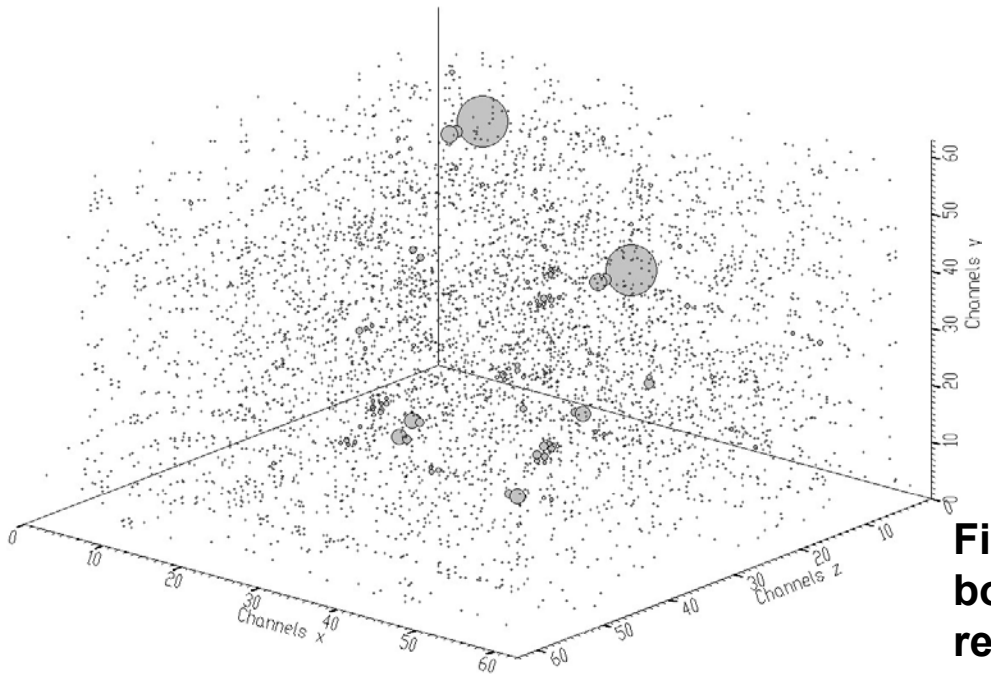

Fig. 52 Spectrum from Fig. 51 after boosted Gold deconvolution (50 iterations repeated 20 times).

Deconvolution algorithm based on Tikhonov regularization of squares of negative values

•in the classic Tikhonov algorithm of zero-th-order regularization we minimize the sum of squares of errors plus the sum of squares of contents of all channels.
•the weight (trade-off) between both components is expressed by the regularization coefficient α (see (8b))
•the second component (sum of squares of channel contents) means in fact the energy of the spectrum.
•by minimization of the energy of the spectrum we wish to suppress oscillations in the solution, both positive and negative.
•the basic idea of the proposed algorithm is that we are interested only in the minimization of squares of negative values.
•we do not care about channels with positive values. We do not regularize them.
•in the first iteration step we calculate the solution according to (8a) without regularization.
•from this solution we determine positions of channels with negative values.
•in the next iteration step we regularize only the channels with negative values. It means that .

$$Q_0^{(k)}(i,j) = \begin{cases} 1 \; if \; x^{(k)}(i) < 0 \;\; and \; i = j \\ 0 \;\; else \end{cases}$$

where k is the iteration step. The iterative deconvolution algorithm based on minimization of squares of negative values (MSNV) can be expressed as follows

$$\hat{\boldsymbol{x}}^{(k)} = \left(H^T H + \alpha Q_0^{(k)T} Q_0^{(k)}\right)^{-1} H^T \boldsymbol{y}^{(k)}$$
$$\boldsymbol{y}^{(k+1)} = H\hat{\boldsymbol{x}}^{(k)}$$

Morhac M., Matousek V.: Complete positive deconvolution of spectrometric data. Digital Signal Processing, Accepted for publication, 2008, doi 10.1016/j.dsp.2008.06.002.

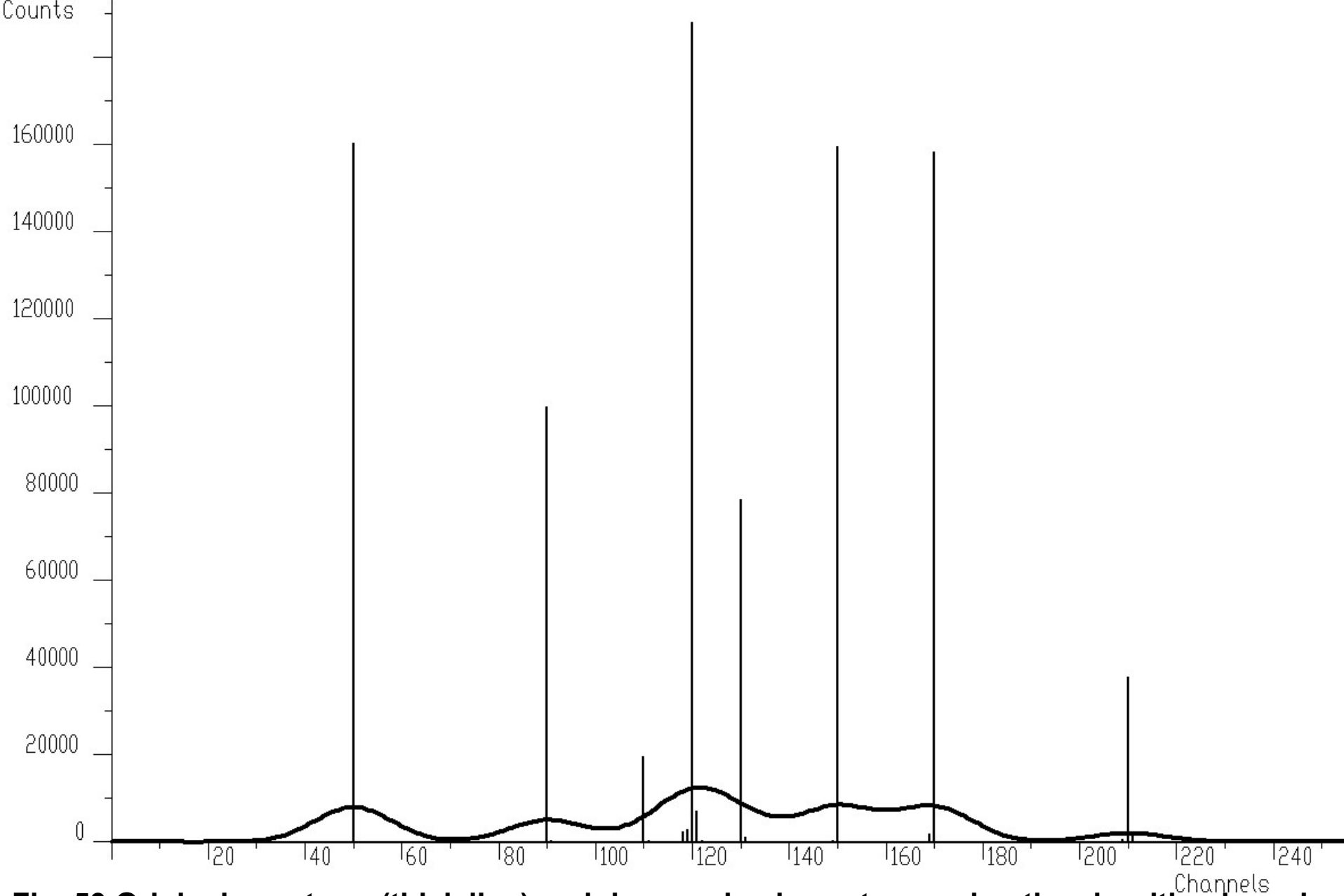

Fig. 53 Original spectrum (thick line) and deconvolved spectrum using the algorithm based on minimum of squares of negative values (bars) for α=1000 and 1000 iterations.

5. Identification of spectroscopic information carrier objects

Peak identification

•Goal: to identify automatically peaks positions in spectrum with the presence of the continuous background and statistical fluctuations - noise

•algorithm of identification of peaks in multidimensional spectra – originally developed in [18]

•peak searching based on smoothed second differences or on convolution with the second derivative of Gaussian

Morháč M., Kliman J., Matoušek V., Veselský M., Turzo I., Identification of peaks in multidimensional coincidence gamma-ray spectra, NIM A 443 (2000) 108.

•generalization for n-dimensional spectra

Peak searching algorithm based on Smoothed Second Differences

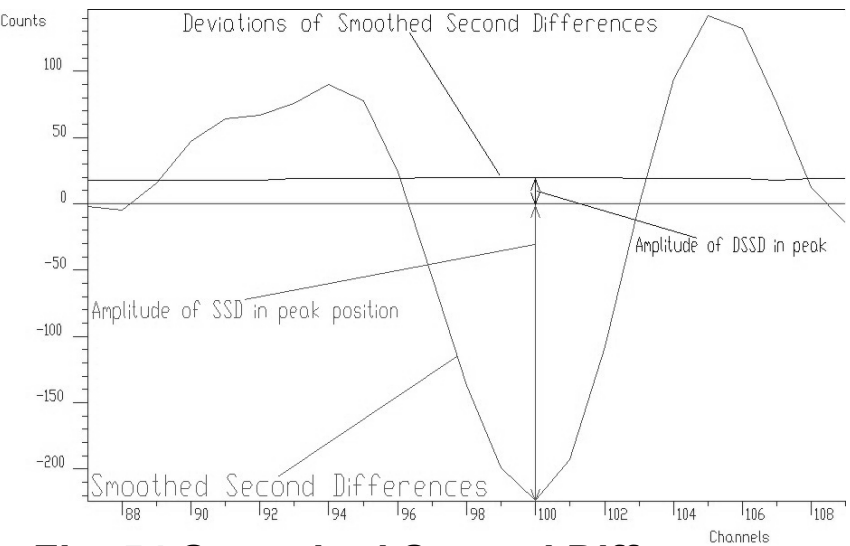

Fig. 54 Smoothed Second Differences (SSD) and its standard deviation in the vicinity of peak

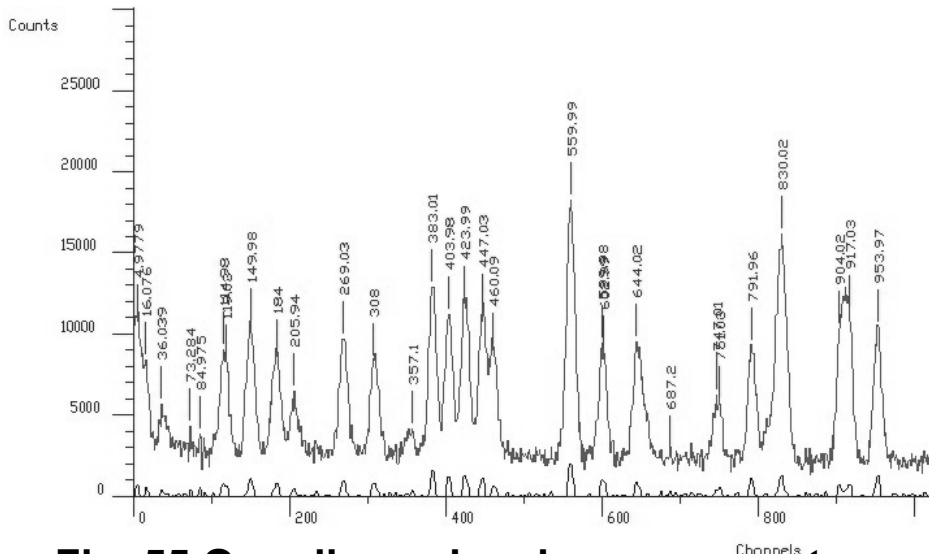

Fig. 55 One-dimensional γ -ray spectrum with found peaks and its inverted positive SSD spectrum

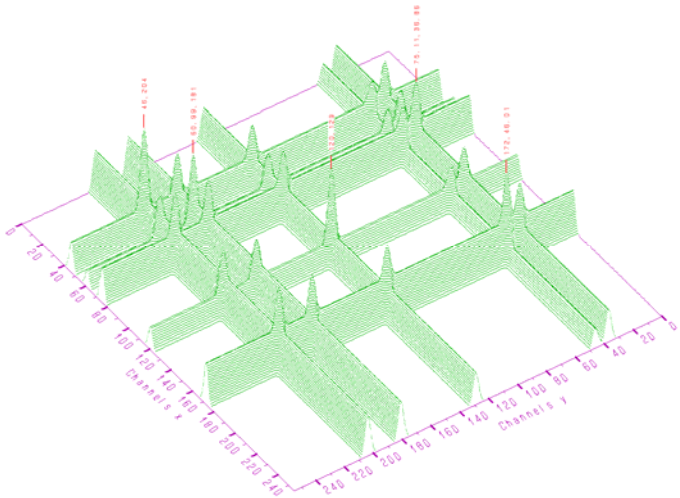

Fig. 56 Synthetic 2D spectrum with found peaks denoted by marks

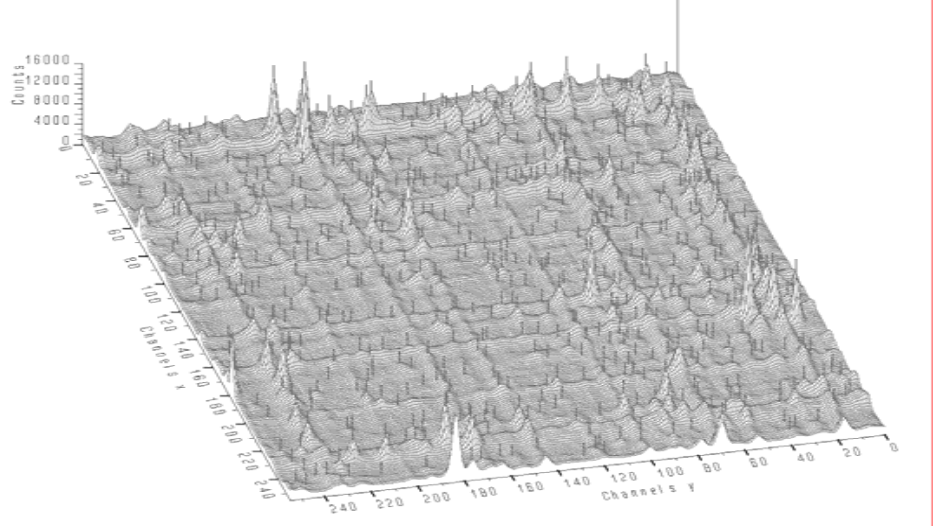

Fig. 57 Coincidence gamma-ray 2D spectrum with found peaks denoted by marks

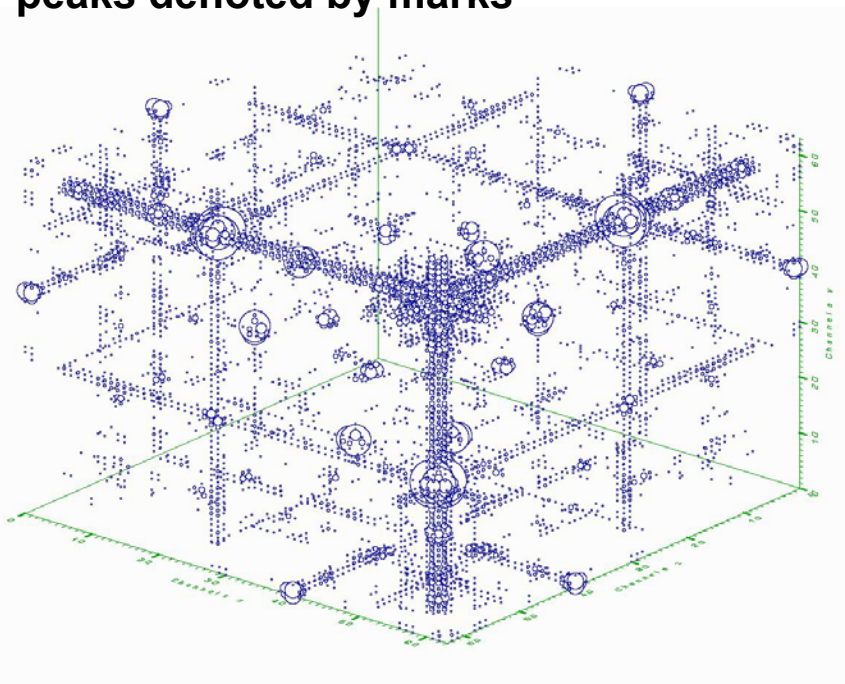

Fig. 58 Three-dimensional γ-γ-γ - coincidence spectrum

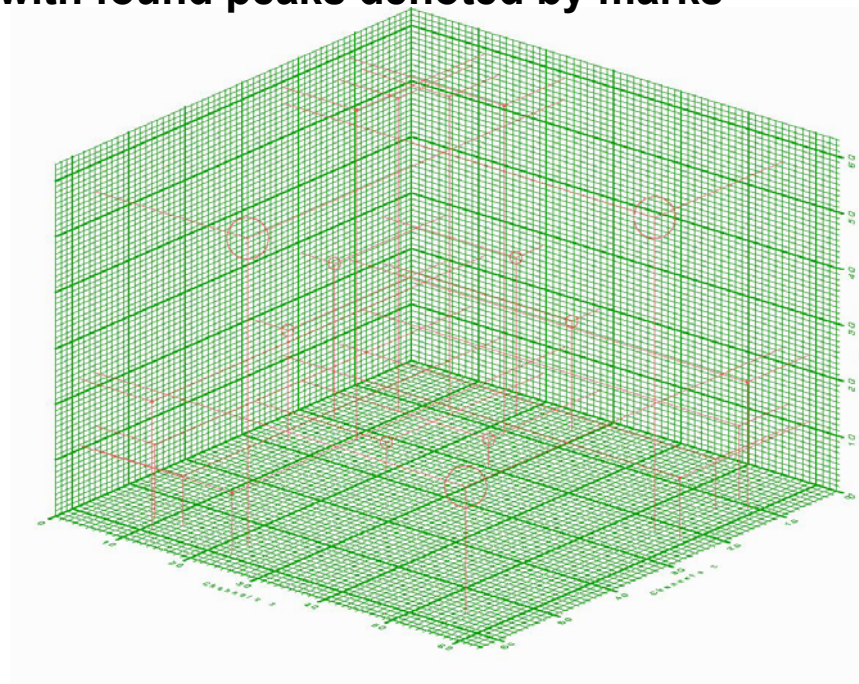

Fig. 59 Found peaks in three-dimensional γ-γ-γ denoted by markers

High resolution peak searching algorithm (based on Gold deconvolution)

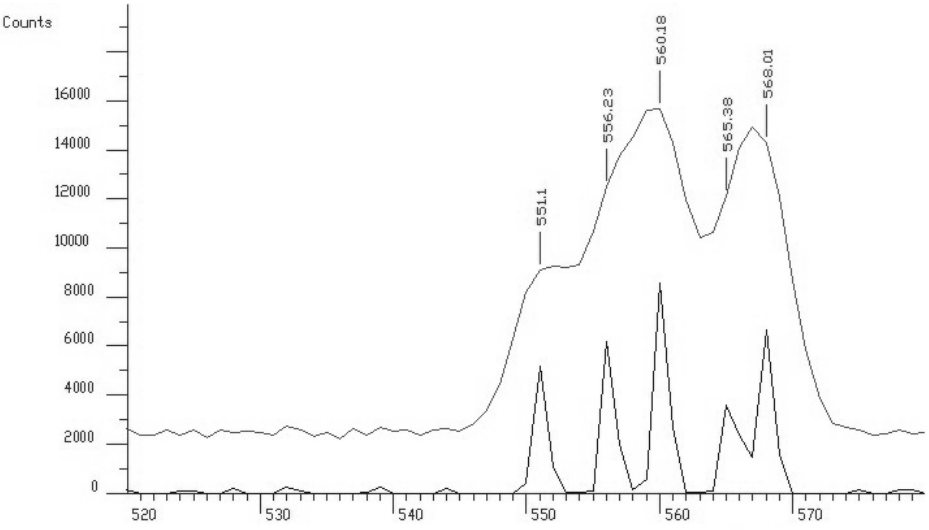

Fig. 60 An example of cluster of peaks with identified peaks and spectrum after background elimination and deconvolution

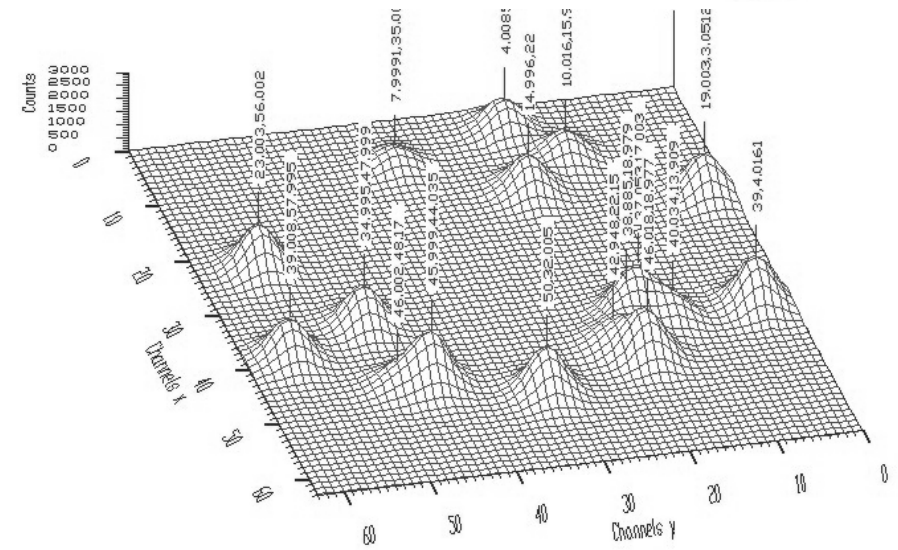

Fig. 61 Original two-dimensional spectrum with closely positioned peaks denoted by markers

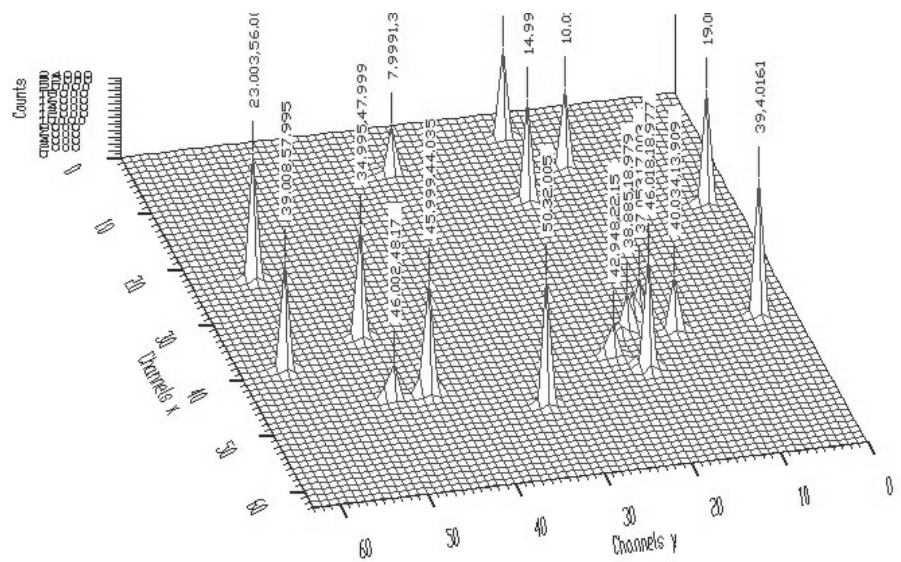

Fig. 62 Spectrum from Fig. 61 after deconvolution

Multidimensional peak searching algorithm for low-statistic spectra[19]

M. Morháč: Multidimensional peak searching algorithm for low-statistics nuclear spectra. NIM A, Vol. 581, pp. 821-830, 2007.

- multidimensional smoothing algorithm based on Markov chain method
- useful for low-statistics data, for peak searching

- principle of the algorithm of one-dimensional spectra smoothing based on Markov chains
- smoothed spectrum S is calculated according to the following formula

$$S(i) = S(i-1) \cdot P^{+}(i-1) / P^{-}(i) \qquad i \in \langle 1, N-1 \rangle$$

where

$$P^{\pm}(i) = A(i) \sum_{k=1}^{m} \exp\left[\frac{y(i \pm k) - y(i)}{\sqrt{y(i \pm k) - y(i)}} \right]$$

A(i) is determined from the condition

$$P^{+}(i) + P^{-}(i) = 1 \quad \text{and} \quad \sum_{i=0}^{N-1} S(i) = 1$$

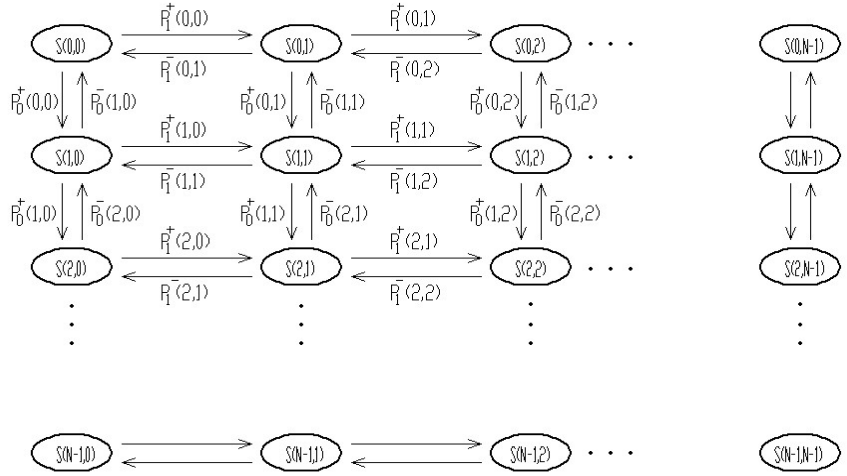

Fig. 63 Principle of the algorithm of two-dimensional spectra smoothing based on Markov chains

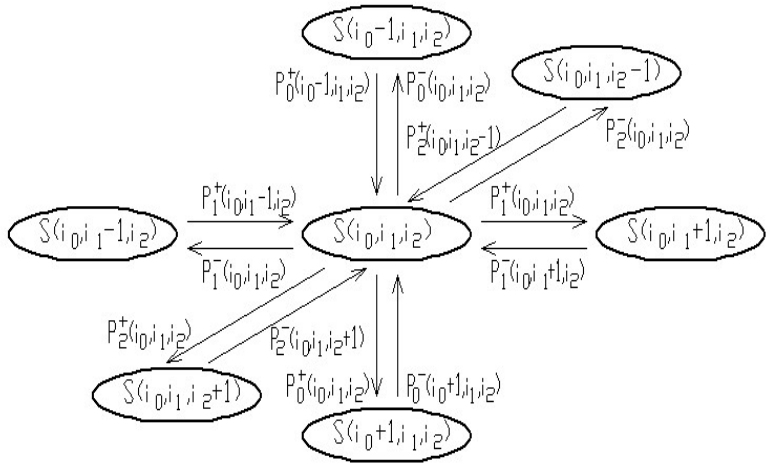

Fig. 64 One element of three-dimensional array of Markov chain of smoothed spectrum

$$S\left(i_0,i_1\right)=\left[S\left(i_0-1,i_1\right)P_0^{+}\left(i_0-1,i_1\right)+S\left(i_0,i_1-1\right)P_1^{+}\left(i,i-1\right)\right]\Big/\left[P_0^{-}\left(i_0,i_1\right)+P_1^{-}\left(i_0,i_1\right)\right]$$

$$P_0^{\pm}\left(i_0,i_1\right)=A\left(i_0,i_1\right)\sum_{k=1}^{m}\exp\left[\frac{y\left(i_0\pm k,i_1\right)-y\left(i_0,i_1\right)}{\sqrt{y\left(i_0\pm k,i_1\right)+y\left(i_0,i_1\right)}}\right]\qquad P_1^{\pm}\left(i_0,i_1\right)=A\left(i_0,i_1\right)\sum_{k=1}^{m}\exp\left[\frac{y\left(i_0,i_1\pm k\right)-y\left(i_0,i_1\right)}{\sqrt{y\left(i_0,i_1\pm k\right)+y\left(i_0,i_1\right)}}\right]$$

$$P_0^{+}\left(i_0,i_1\right)+P_0^{-}\left(i_0,i_1\right)+P_1^{+}\left(i_0,i_1\right)+P_1^{-}\left(i_0,i_1\right)=1\qquad i_0,i_1\in\left\langle 1,N-1\right\rangle$$

$$\sum_{i_0=0}^{N-1}\sum_{i_1=0}^{N-1}S\left(i_0,i_1\right)=1$$

•generalization for n-dimensional spectra

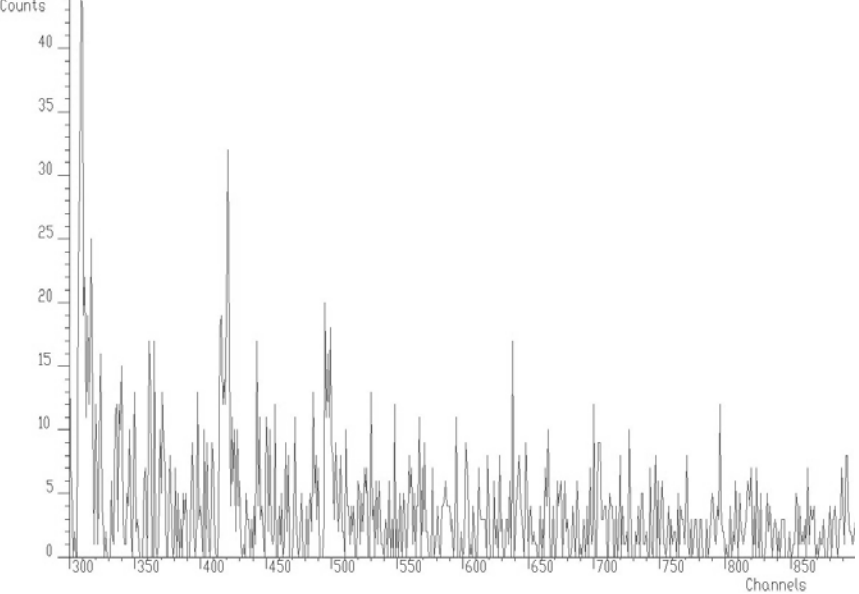

Fig. 65 One-dimensional noisy spectrum

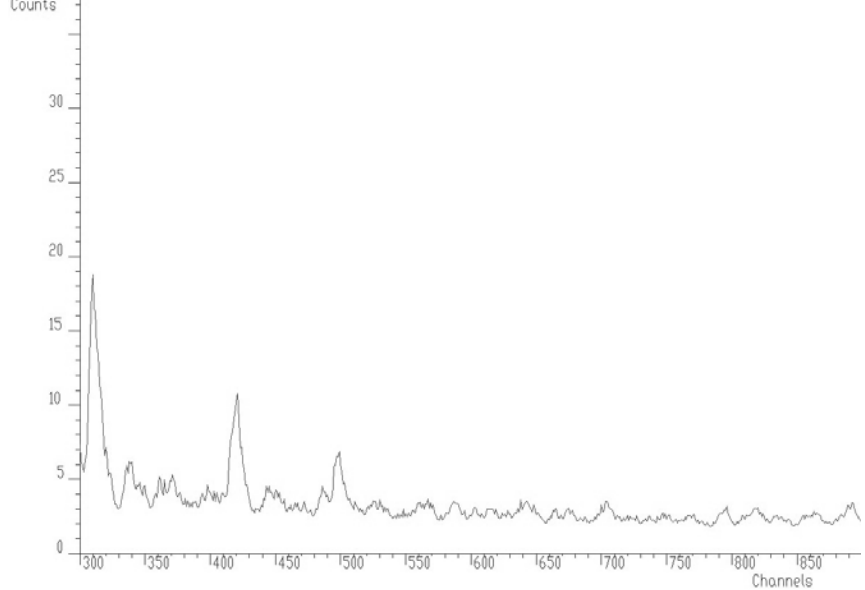

Fig. 66 One-dimensional Markov spectrum (m=7)

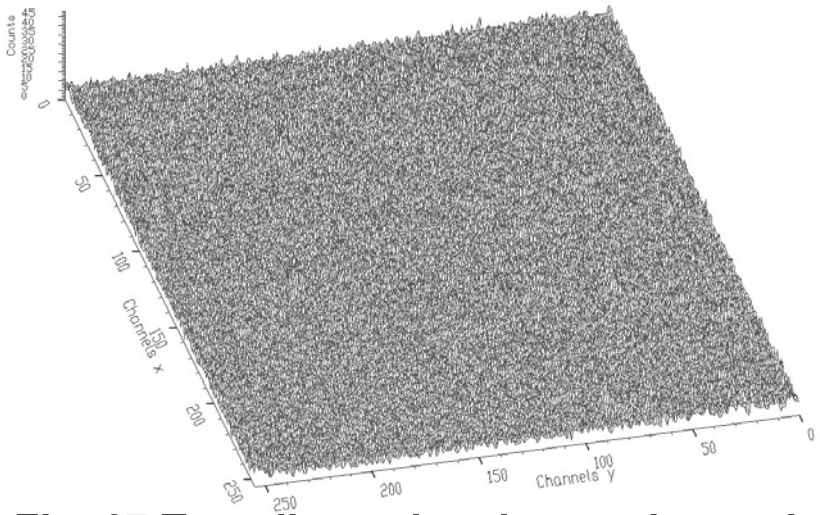

Fig. 67 Two-dimensional experimental γ γ -ray spectrum

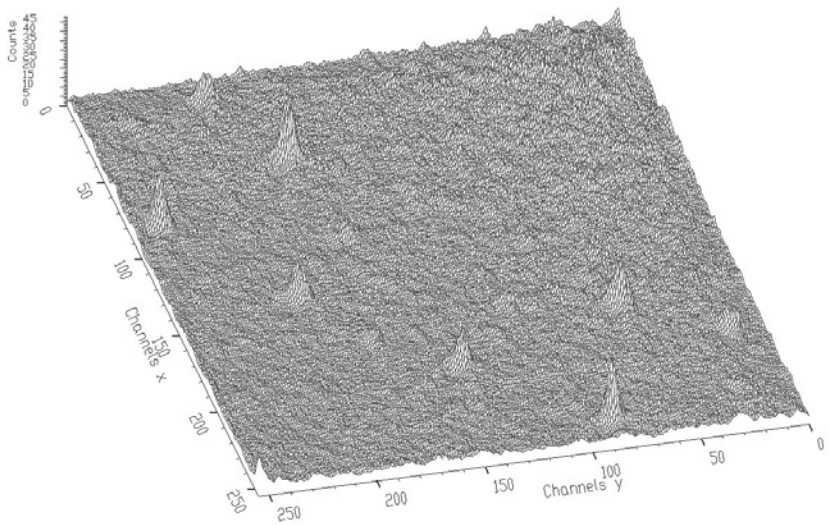

Fig. 68 Markov spectrum (m=5) of the spectrum from Fig. 67

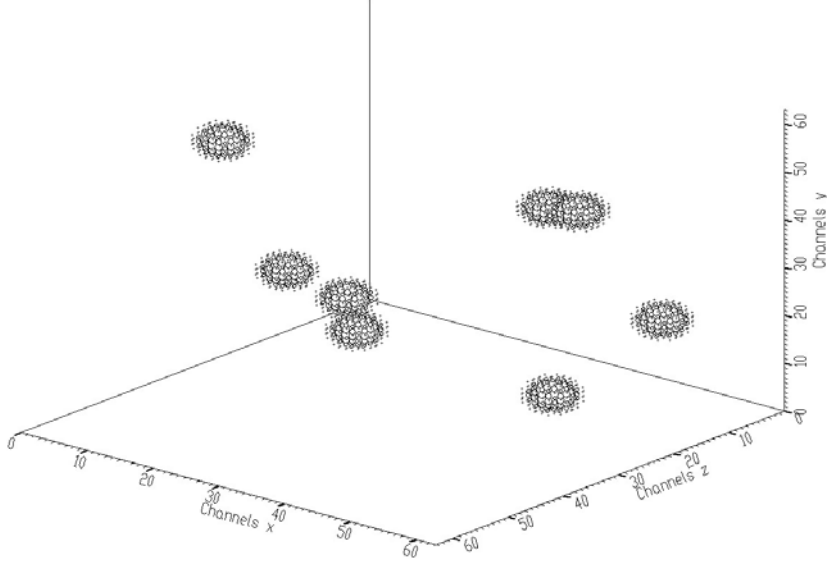

Fig. 69 An example of three-dimensional synthetic spectrum with 8 peaks plus constant background

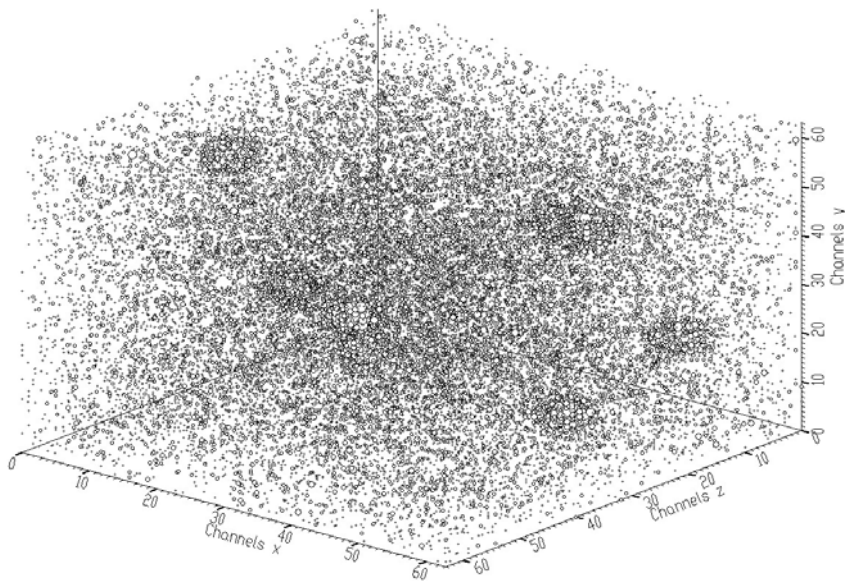

Fig. 70 Spectrum from Fig. 69 plus Gaussian noise (30% of peak amplitudes)

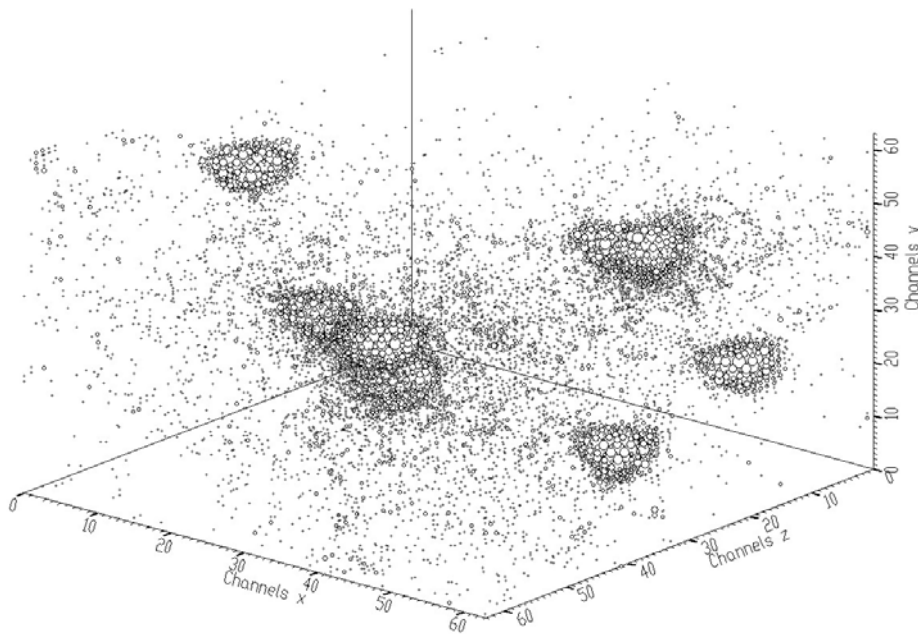

Fig. 71 Markov spectrum (m=3) of the spectrum from Fig. 70

Identification of ridges in two-dimensional spectra of nuclear multifragmentation [20]

Morhac M., Veselsky M. : Identification of isotope lines in two-dimensional spectra of nuclear multifragmentation. NIMA, A 592, pp. 434–450, 2008.

Goal: to determine ridges of corresponding points from very sparsely distributed two-dimensional experimental data

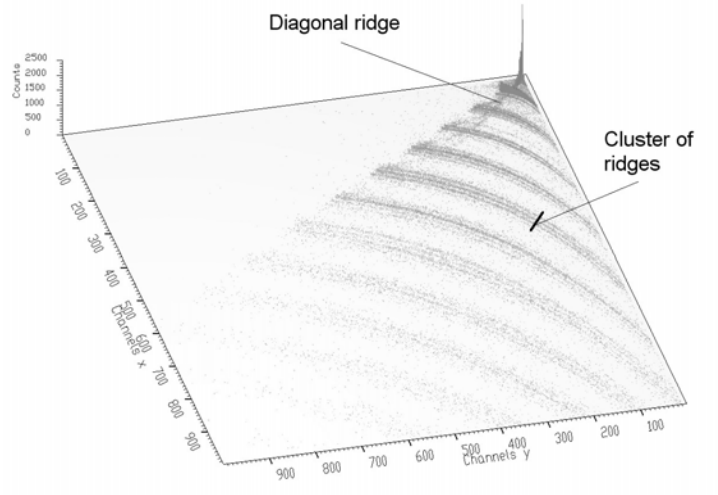

Fig. 72 Original Si-Si spectrum

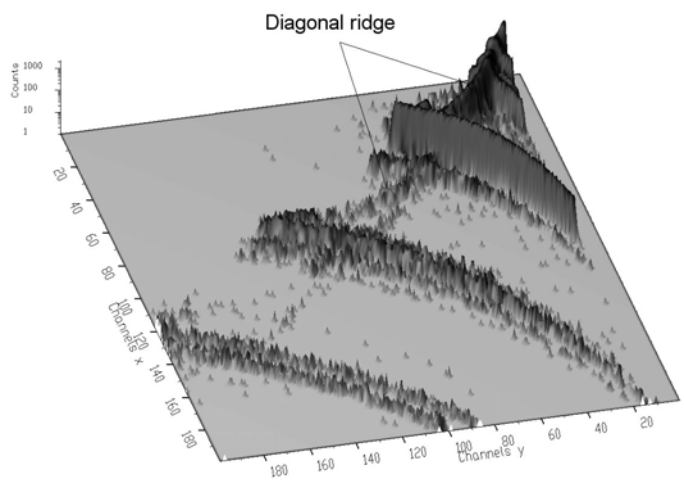

Fig. 73 Detail of the spectrum

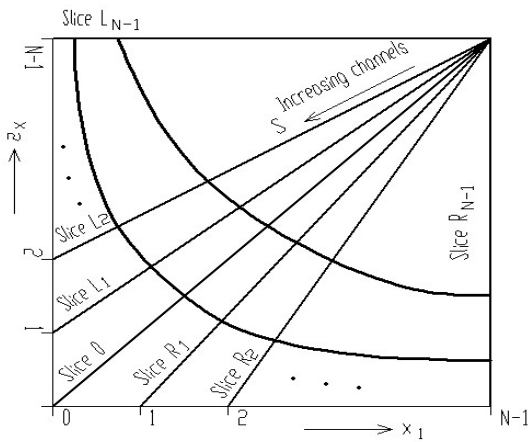

Fig. 74 Principle of slicing

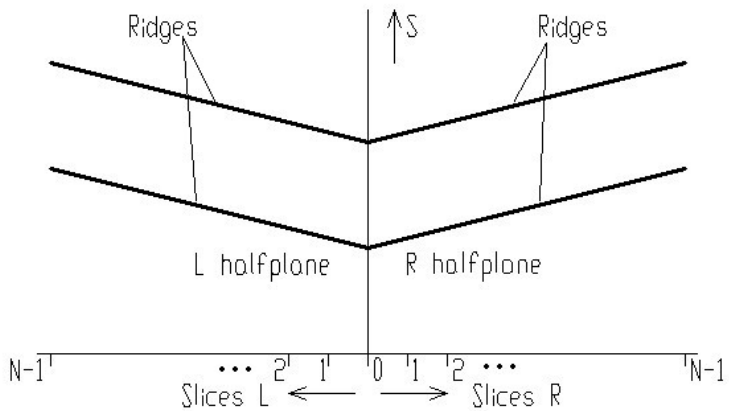

Fig. 75 Linearized ridges in sliced spectrum

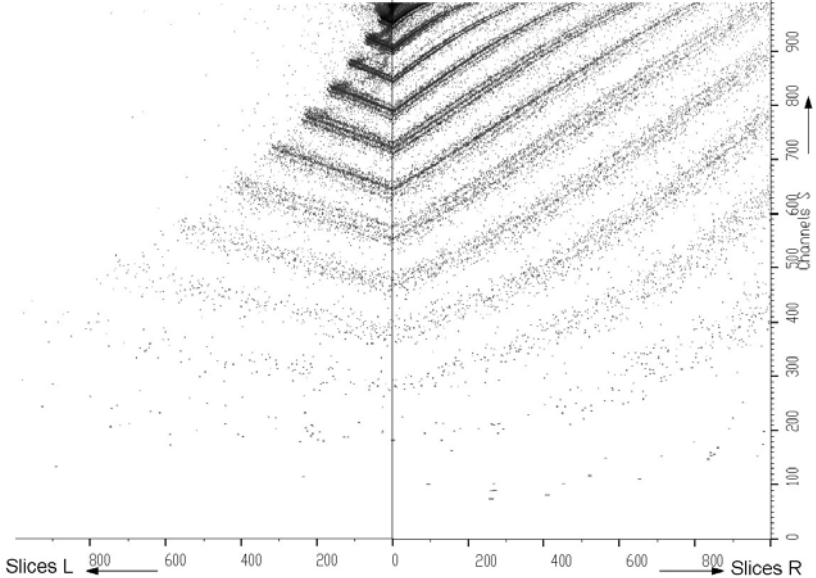

Fig. 76 Spectrum from Fig. 72 sliced and linearized

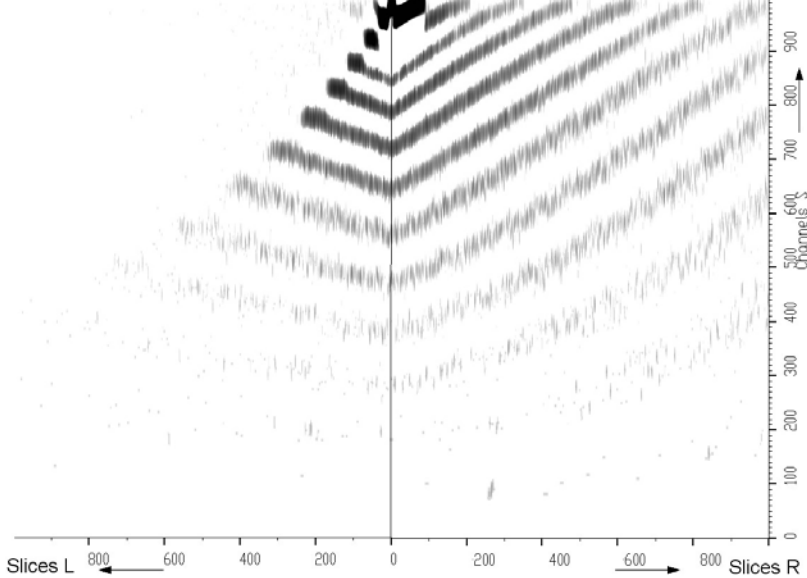

Fig. 77 Data from Fig. 76 smoothed vertically

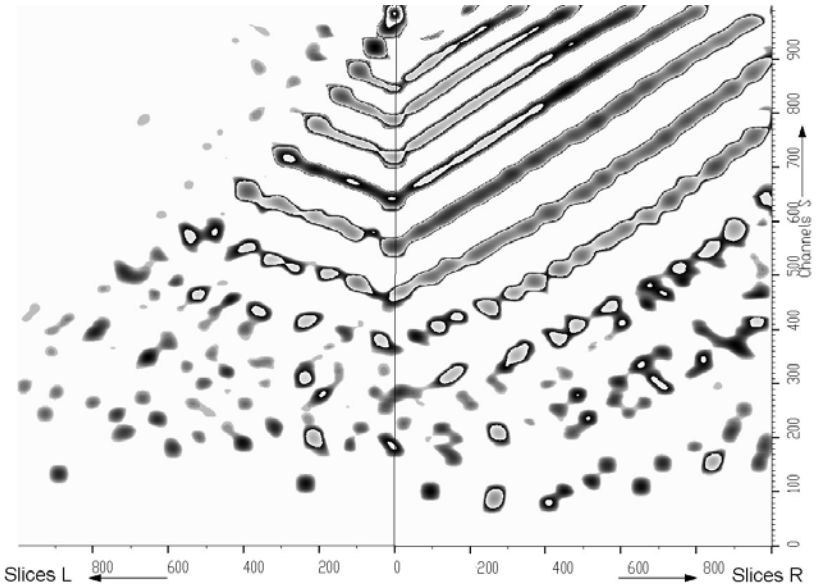

Fig. 78 Data from Fig. 77 smoothed in horizontal direction

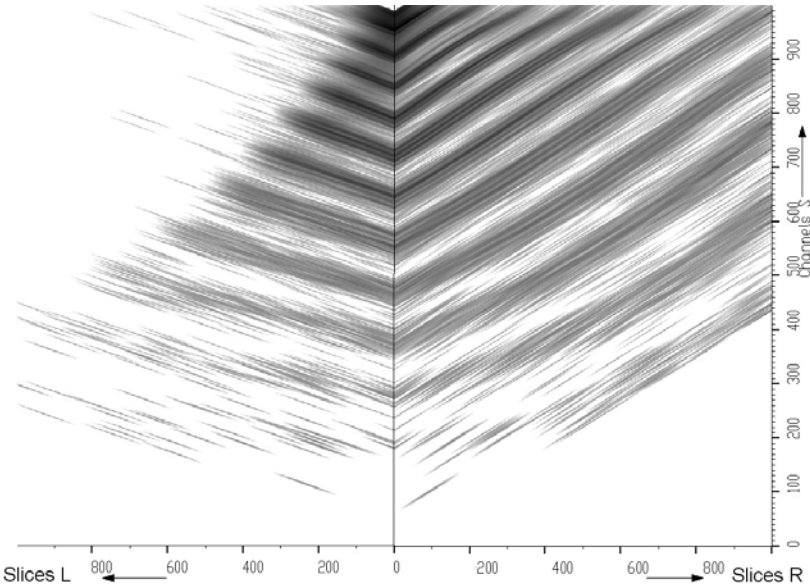

Fig. 79 Data from Fig. 72 after Gaussian smoothing in the direction of ridges

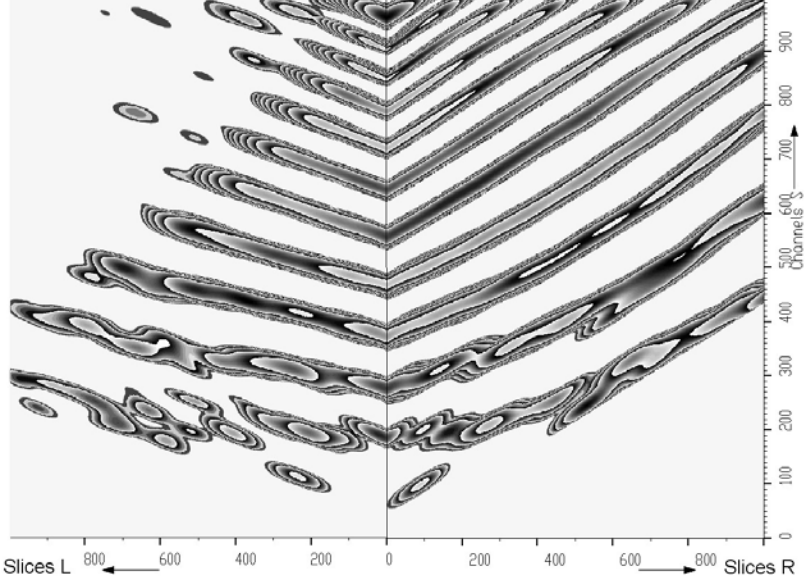

Fig. 80 Data from Fig. 79 after application of inverse positive second derivative filter

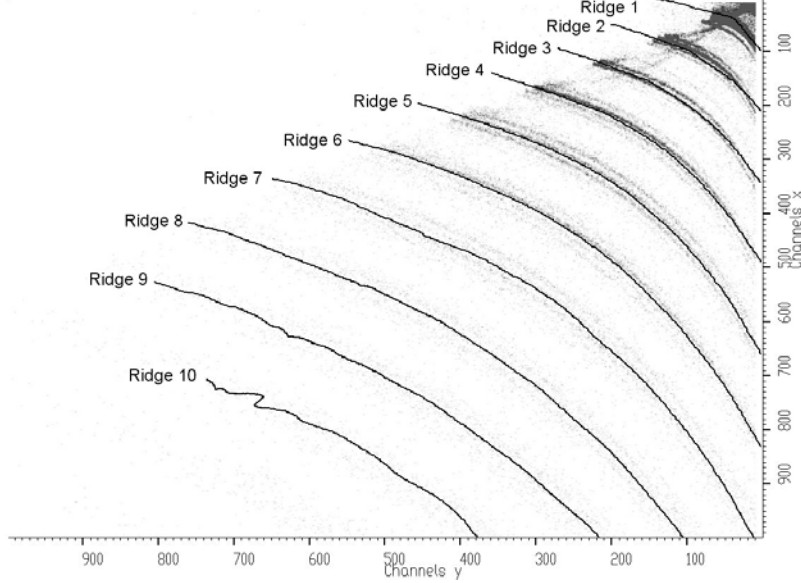

Fig. 81 Spectrum from Fig. 72 with determined ridges

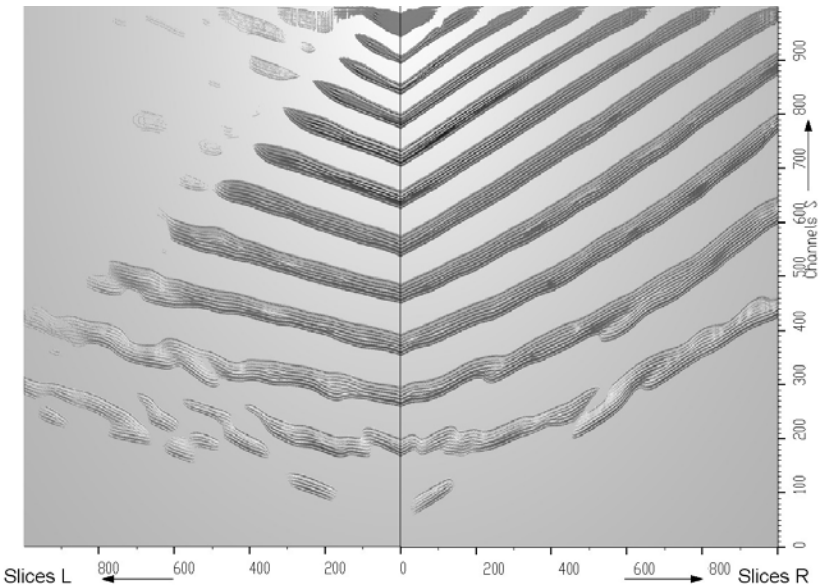

Fig. 82 Data from Fig. 80 after deconvolution of vertical slices

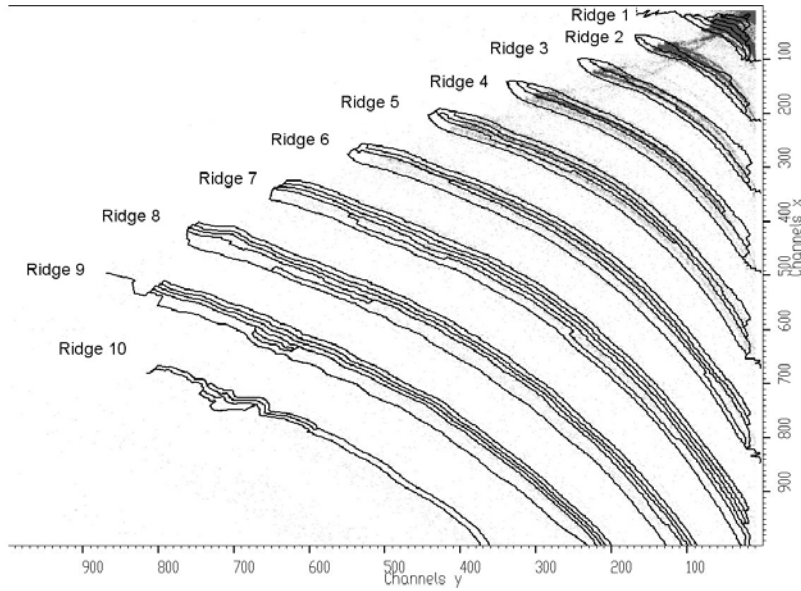

Fig. 83 estimated decomposed slices transformed back to the two-dimensional spectrum

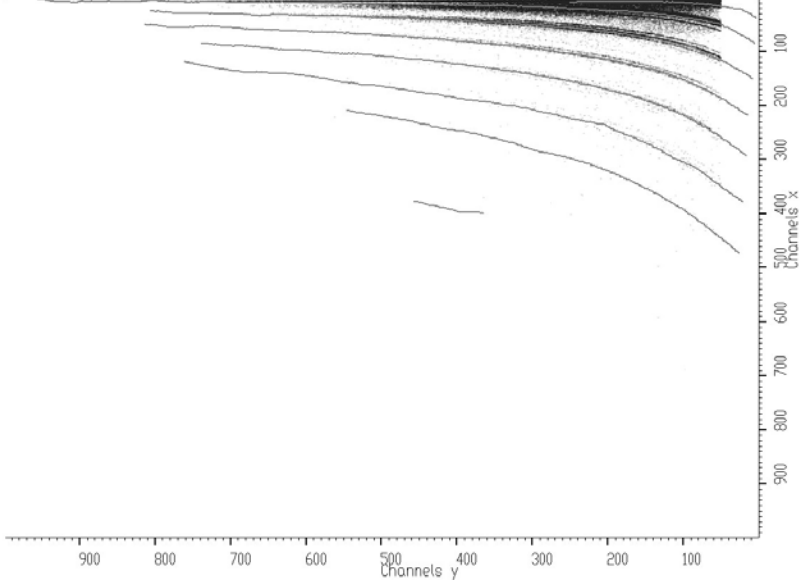

Fig. 84 Spectrum with determined ridges

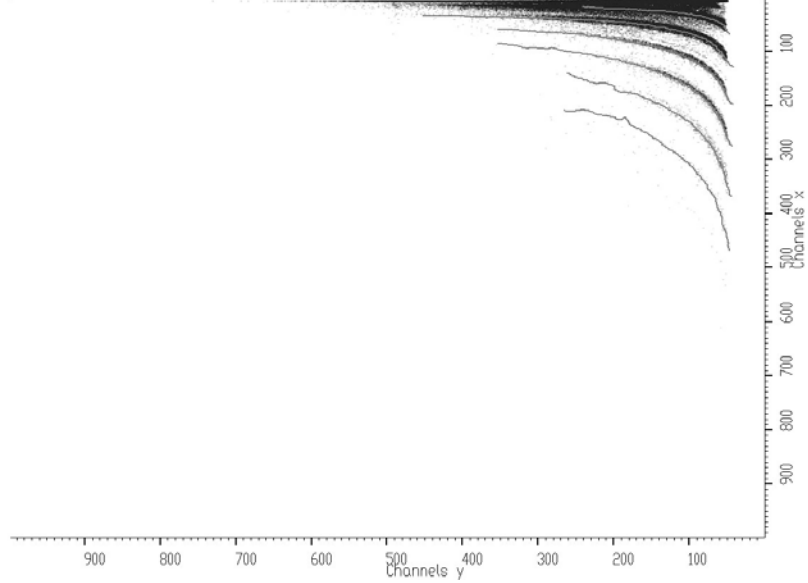

Fig. 85 Spectrum with determined ridges

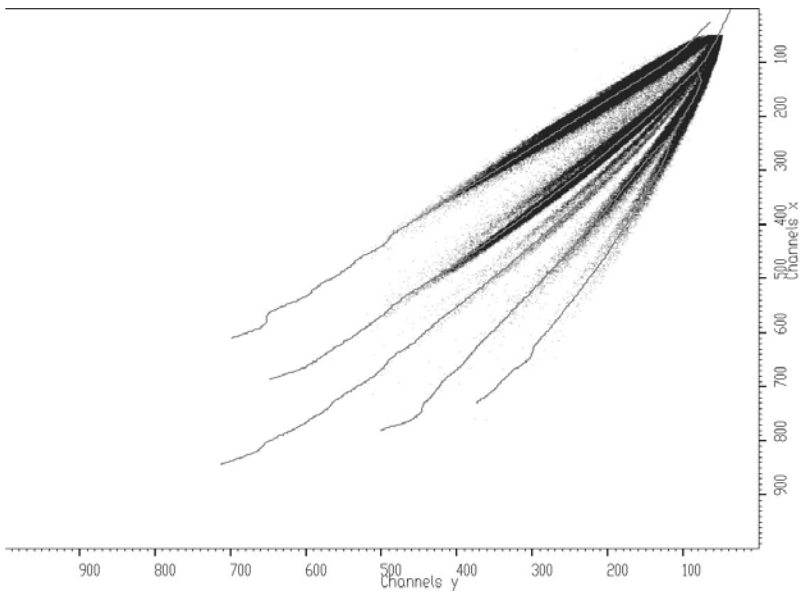

Fig. 86 Spectrum with determined ridges

Identification of rings in two-dimensional spectra from RICH (Ring Imaging Cherenkov) detectors

Goal: to determine rings of corresponding points from very sparsely distributed two-dimensional experimental data

•principle of the algorithm is based on two-dimensional Gold deconvolution

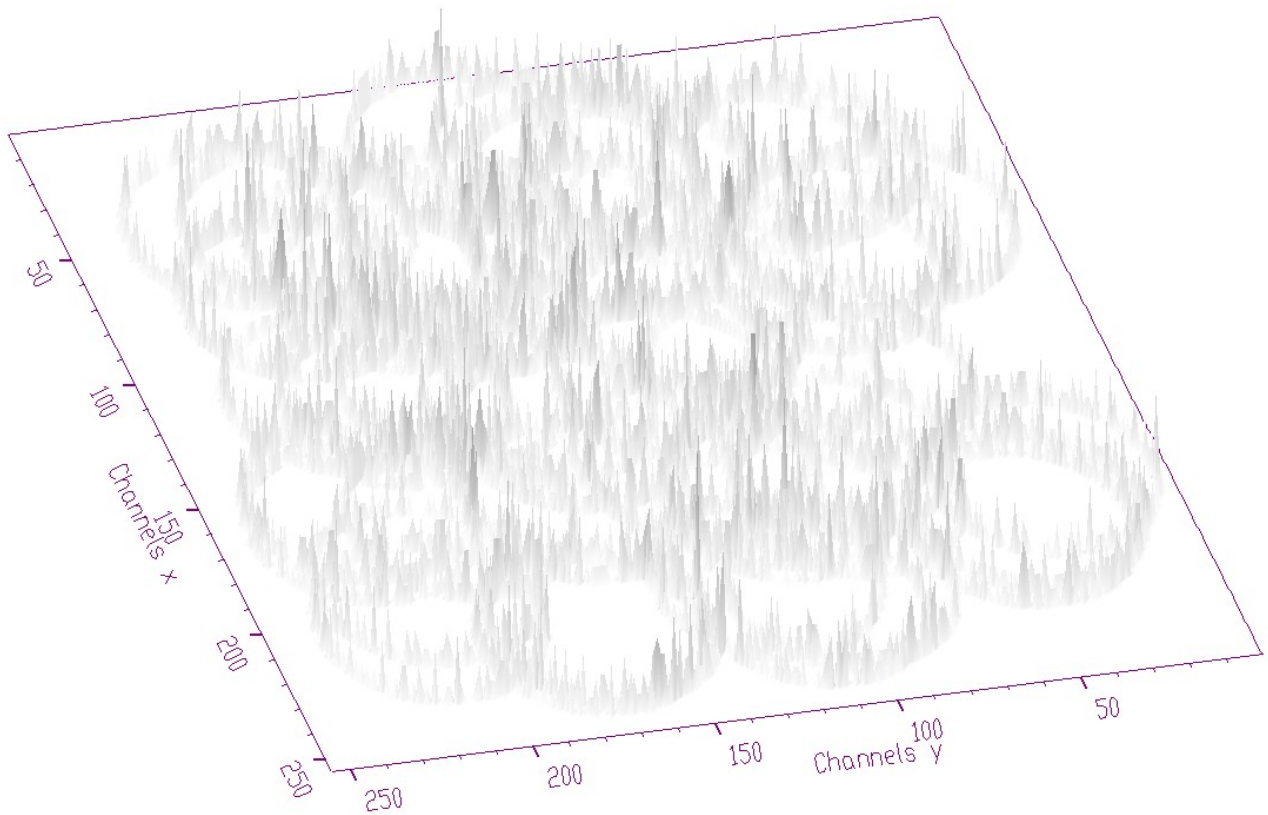

Fig. 87 Synthetic two-dimensional spectrum with 50 rings

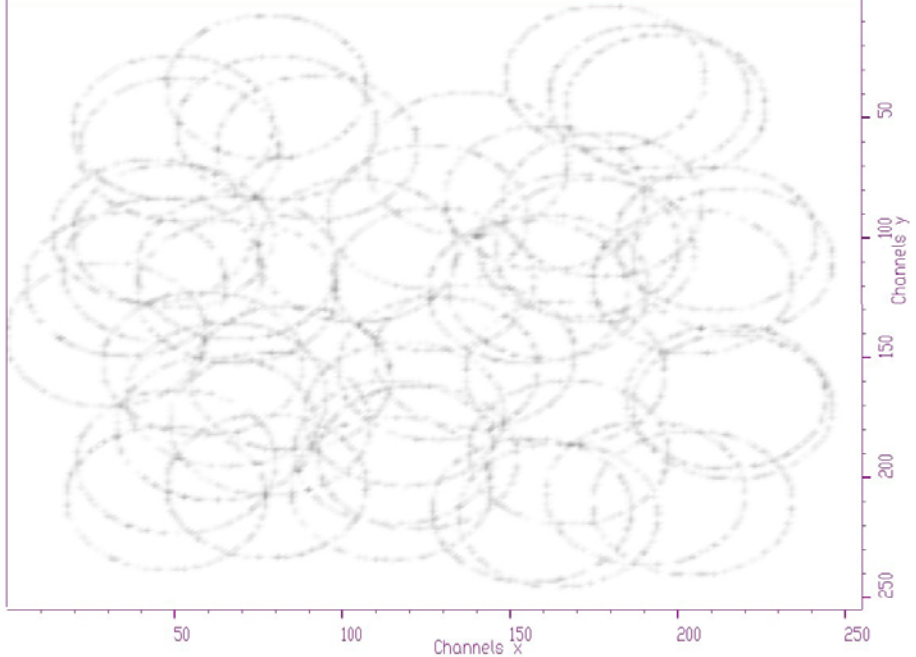

Fig. 88 Orthogonal view of the spectrum from Fig.87

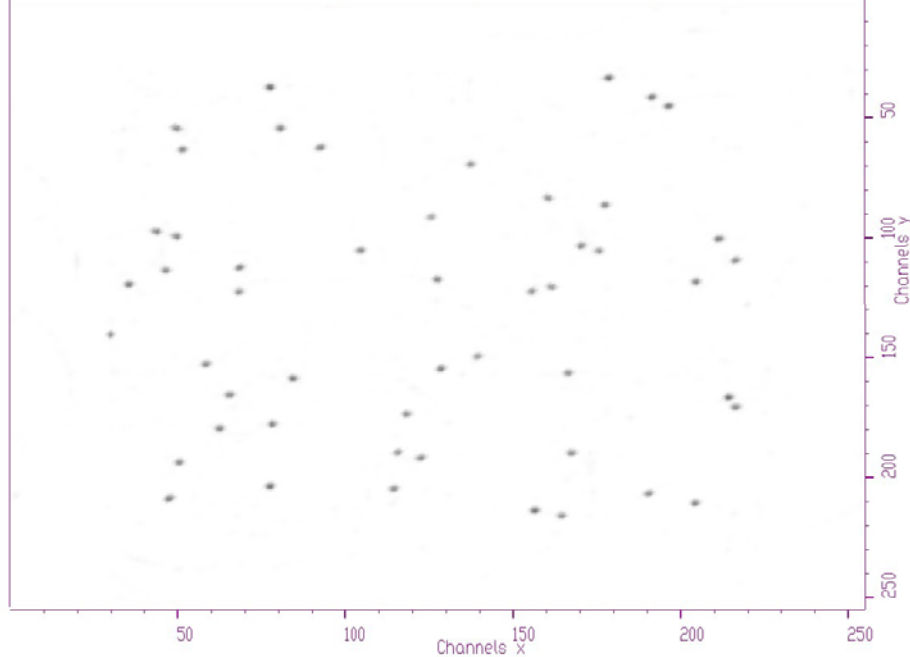

Fig. 89 Deconvolved data of the spectrum from Fig.87

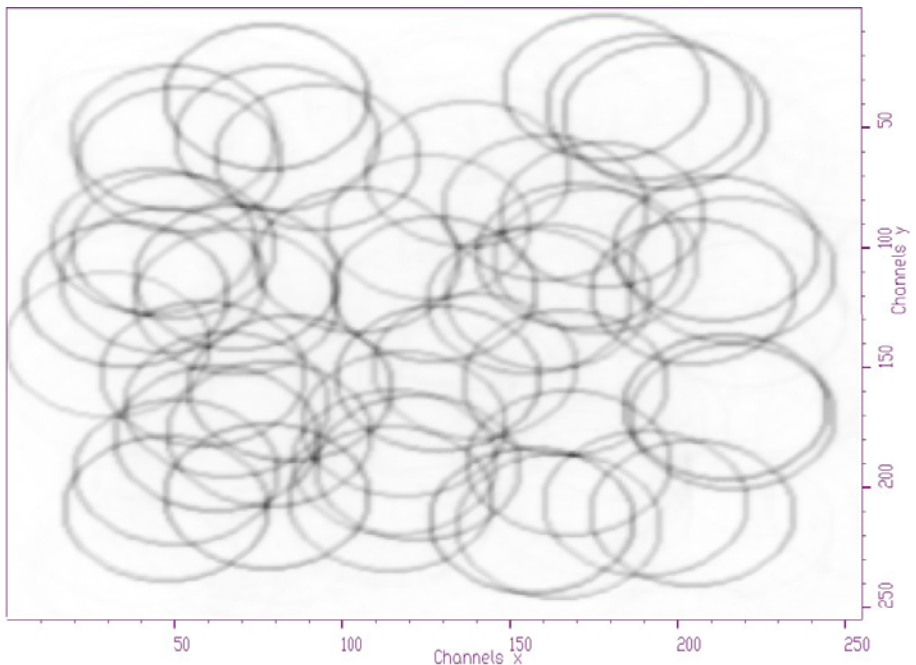

Fig. 90 Identified rings

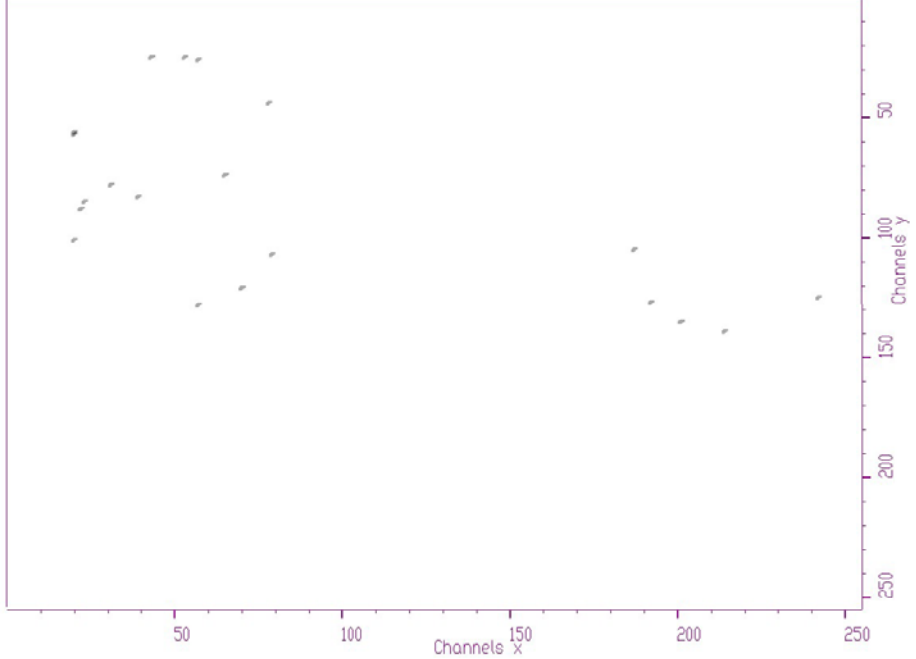

Fig. 91 Several points in 3 rings

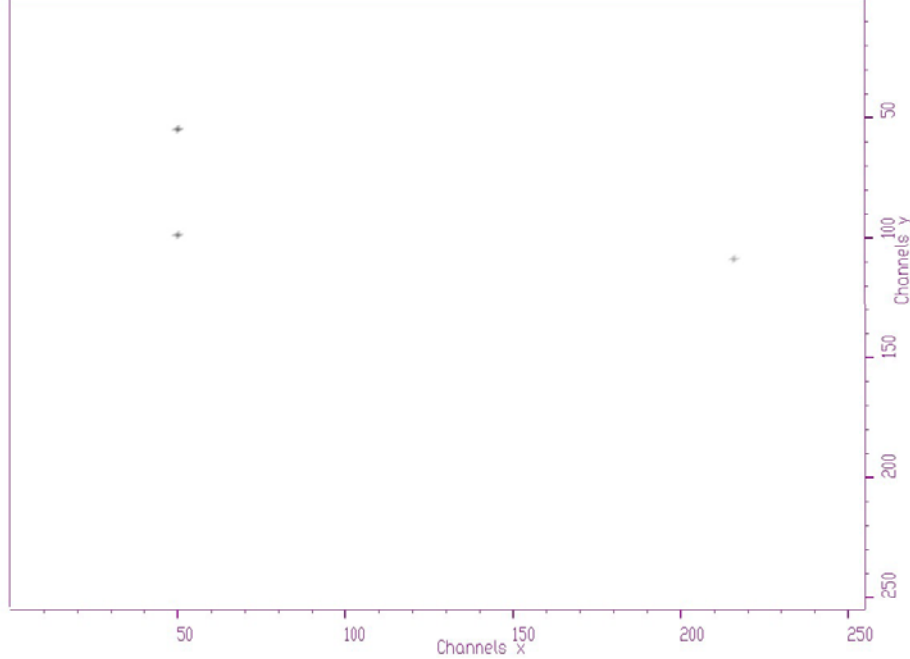

Fig. 92 Deconvolved data (3 points) of the spectrum from Fig.91

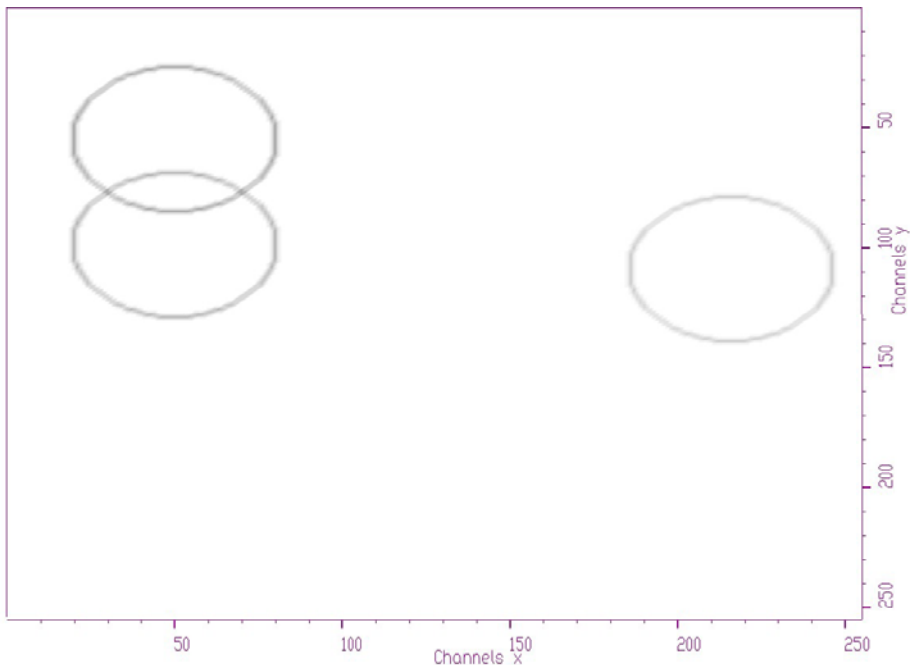

Fig. 93 Identified rings

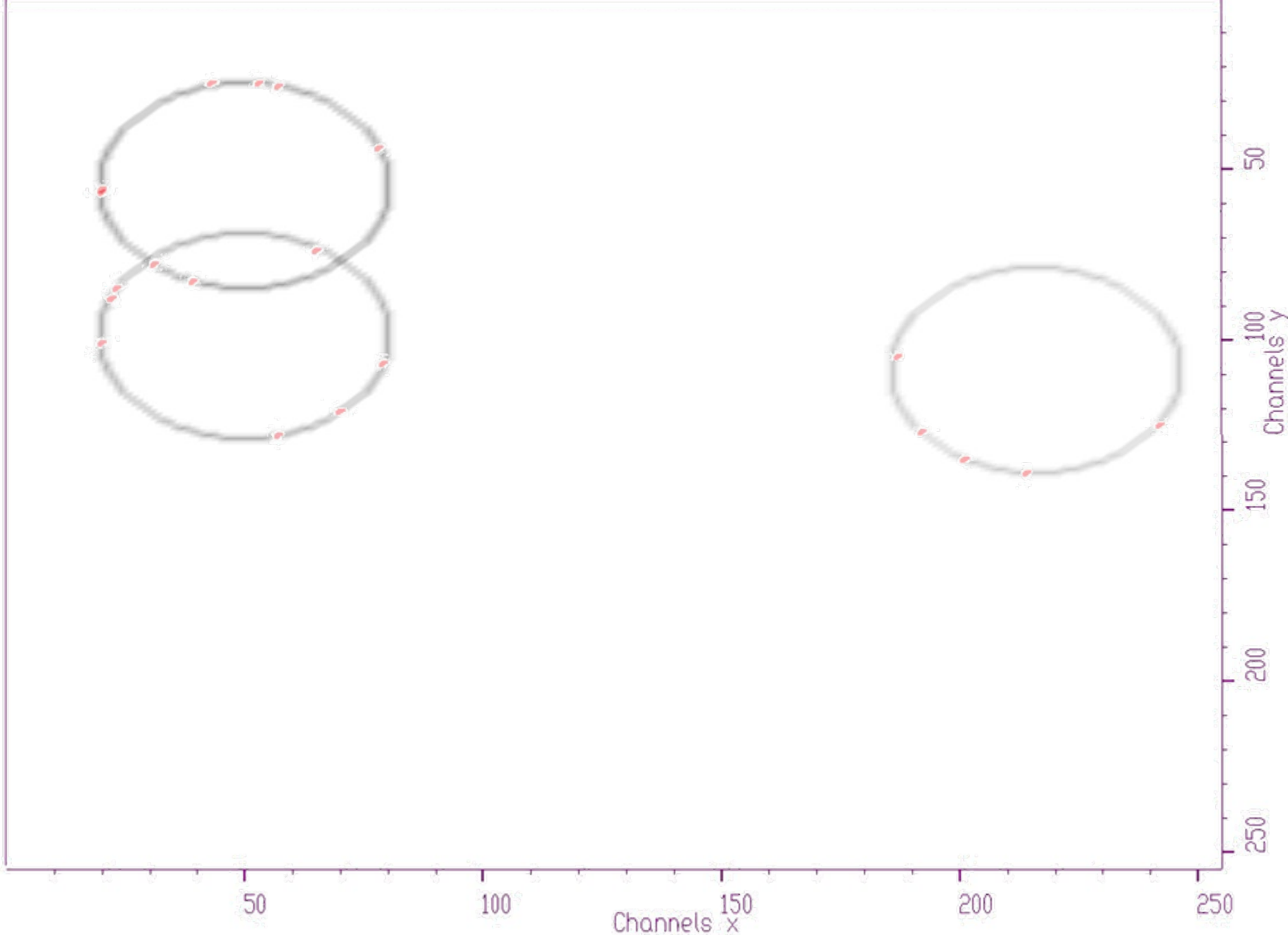

Fig. 94 Original data (red points) and identified rings

Conclusions

•we have generalized and extended the existing basic SNIP algorithm for additional parameters and possibilities that make it possible to improve substantially the quality of the background estimation
•we have included these modifications and derived the algorithms for two-, three-, up to n-dimensional spectra
•we proposed a new approach for the determination of the peak regions
•in the contribution we suggested a new algorithm of background estimation with the clipping window adaptable to the widths of peak regions
•moreover the algorithm for separation of peaks containing regions from peak-free regions can be utilized for fitting purposes to confine the fitting regions
•in the work, we have discussed and analyzed a series of deconvolution methods
•we present a survey of existing deconvolution methods and we emphasize the relations among them
•we have restricted our investigations only to positive definite methods. Though they improve substantially the resolution in the spectra they are not efficient enough to decompose closely positioned peaks.
•we proposed boosted deconvolution algorithms.
•boosted algorithms are able to decompose the overlapped peaks practically to δ functions while concentrating the peak areas to one channel.
•we have derived deconvolution algorithm based on Tikhonov regularization of squares of negative values
•in the contribution we present peak searching algorithm based on smoothed second differences high resolution peak searching algorithm based on Gold deconvolution and peak searching algorithm for low-statistic spectra based on Markov chain method
•all these algorithms have been extended to higher dimensions

•finally in the contribution we present the algorithm of identification of ridges in two-dimensional spectra of nuclear multifragmentation and spectra from RICH detectors

•all the presented algorithms were implemented in DaqProVis system [21]

M. Morháč, V. Matoušek, I. Turzo, J. Kliman J.: DaqProVis, a toolkit for acquisition, interactive analysis, processing and visualization of multidimensional data. NIM A, Vol. 559/1 (2006), pp. 76-80.

•several algorithms were also implemented in ROOT system [22] in the form of TSpectrum, TSpectrum2 and TSpectrum3 classes, developed in collaboration with CERN

•integral part of analysis of spectrometric data is their visualization – poster – hypervolume techniques

References

[1] Ryan C.G., Clayton E., Griffin W.L., Sie S.H. and Cousens D.R., Nucl. Instr. and Meth. B 34 (1988) 396.

[2] M. Morháč, J. Kliman, V. Matoušek, M. Veselský, and I. Turzo, Nucl. Instr. and Meth. **A401**, 113 (1997).
[3] M. Morháč and V. Matoušek, Applied Spectroscopy 62, 91 (2008).
[4] A. Likar and T. Vidmar, J. Phys. D: Appl. Phys. 36 (2003) 1903.
[5] A. Likar, T. Vidmar and M. Lipoglavšek, J. Phys. D: Appl. Phys. 37 (2004) 932.
[6] C. Takiya, O. Helene, E. do Nascimento and V.R. Vanin, Minimum variance regularization in linear inverse problems, Nucl. Instrum. Methods Phys. Res., A 523 (1-2) (2004), pp. 186-192.
[7] G.E. Backus and F. Gilbert, The resolving power of gross earth data, Geophysical Journal of the Royal Astronomical Society 16 (1968) p. 169.
[8] M.K. Ozkan, A.M. Tekalp and M.I. Sezan, POCS-based restoration of space-varying blurred images, IEEE Trans. Image Process. 3 (4) (1994) pp. 450-454.
[9] C. Sanchez-Avila and A.R. Figueiras-Vidal, New POCS algorithms for regularization of inverse problems, Journal of Computational and Applied Mathematics 72 (1) (1996) pp. 21-39.
[10] G.E.Coote, Iterative smoothing and deconvolution of one- and two-dimensional elemental distribution data, Nucl. Instrum. Methods Phys. Res., B 130 (1-4) (1997) pp. 118-122.
[11] R. Gold, ANL-6984, Argonne National Laboratories, Argonne Ill., 1964.
[12] M. Morháč, J. Kliman, V. Matoušek, M. Veselský and I. Turzo, Efficient one and two dimensional Gold deconvolution and its application to gamma-ray spectra decomposition, Nucl. Instrum. Methods Phys. Res., A 401 (2-3) (1997) pp. 385-408.
[13] M. Morháč, Deconvolution methods and their applications in the analysis of gamma-ray spectra, Nucl. Instrum. Methods Phys. Res., Sect. A 559 (1) (2006), pp. 119-123.

[14] W.H. Richardson, Bayesian-based iterative method of image restoration, J. Opt. Soc. Am. 62 (1972) p. 55.
[15] L.B. Lucy, An iterative technique for the rectification of observed images, Astronomical Journal 79 (1974) p. 745-754.
[16] Y. Lin and D.D. Lee, Bayesian regularization and nonnegative deconvolution for room impulse response estimation, IEEE Transactions on Signal Processing 54 (3) (2006) p. 839-847.
[17] B. Hunt, Int. J. Mod. Phys. C 5 (1994) p. 151.
[18] Morháč M., Kliman J., Matoušek V., Veselský M., Turzo I., Identification of peaks in multidimensional coincidence gamma-ray spectra, NIM A 443 (2000) 108.
[19] Morháč M.: Multidimensional peak searching algorithm for low-statistics nuclear spectra. NIM A, Vol. 581, pp. 821-830, 2007.
[20] Morhac M., Veselsky M. : Identification of isotope lines in two-dimensional spectra of nuclear multifragmentation. NIMA, A 592, pp. 434–450, 2008, doi:10.1016/j.nima.2008.04.002.
[21] M. Morháč, V. Matoušek, I. Turzo, and J. Kliman, Nucl. Instr. and Meth. **A559,**1, 76 (2006).
[22] R. Brun, F. Rademakers, S. Panacek, D. Buskulic, J. Adamczewski, M. Hemberger, ROOT, An Object-Oriented Data Analysis Framework, Users Guide 3.02c, CERN, 2002.