Quantification of overlapping chromatographic peaks using a matched filter.

van den Bogaert, B.; Boelens, H.F.M.; Smit, H.C.

DOI
10.1016/0169-7439(94)85049-6
Publication date
1994

Published in
Chemometrics and Intelligent Laboratory Systems

ELSEVIER

Chemometrics and Intelligent Laboratory Systems 25 (1994) 297–311

Chemometrics and intelligent laboratory systems

Quantification of overlapping chromatographic peaks using a matched filter

Bas van den Bogaert, Hans F.M. Boelens, Her C. Smit *

Laboratory for Analytical Chemistry, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands

Received 3 March 1994; accepted 13 June 1994

Abstract

A novel approach to the quantification of overlapping chromatographic peaks is introduced. The sum of the models for the individual overlapping peaks is taken as the signal model in a matched filter. Thus, the signal intensities being the objective of the quantification procedure become parameters in an overall signal model. These and, if necessary, other parameters are adapted by a modified simplex algorithm optimizing the maximum in the output of the matched filter. A prediction of the results can be made on the basis of a noiseless response surface that can be calculated from the models. When shapes and positions of the peaks are known and only their intensities need to be estimated, a quantitative theoretical error estimation is possible. The results thus predicted are considered optimal and are used as a reference in the evaluation of the results of a range of experiments using simulated data containing two overlapping Gaussians and first-order band-limited noise. The proposed procedure works well, the quality of the results usually being on or just little below the theoretical optimum. Under conditions of high overlap or a low signal to noise ratio, the experimental results no longer follow a normal distribution and their quality is lower.

1. Introduction

A novel approach to the quantification of overlapping chromatographic peaks is introduced. It is based on the MFX, a recently introduced extension to the matched filter [1], the operation of which is described below. The procedure has been tested on simulated data. The results are compared to a theoretical reference, not to existing methods such as non-linear regression [2], Kalman filtering [3,4], deconvolution [5] or integration [6,7].

The matched filter (MF) is an optimal linear filter for data consisting of a signal and additive stationary noise. It is based on exact models of both the signal and the noise component in the data [8–10]. The optimality of the MF concerns the signal to noise ratio (S/N) in the output, if defined as the maximum of the signal component divided by the standard deviation of the noise. In other words, true optimality can be achieved only when the position of this maximum is known. Otherwise, its estimation is a source of error that is not accounted for by the MF theory. Nevertheless, it is sometimes assumed that a MF followed by a maximum searcher represents the optimal estimator for both the amplitude and the position of a signal [11].

A MF is rigid, being optimized for the particular signal whose model it contains, e.g. one peak in a chromatogram. The MFX has been introduced to add flexibility to the MF by inverting the principle of the MF: the best estimation of the signal model is the one for which the output is optimal. Rather than being a matched filter, it is a filter being matched. This has been implemented as a modified simplex optimization of the

* Corresponding author.

SSDI 0169-7439(94)00056-O

output S/N of a MF that is equipped with a parameterized signal model. The simplex adapts the values of the parameters in the model. The S/N is approximated as the ratio of the output at the true or estimated position of the maximum in the signal component to the theoretical standard deviation of the output noise. The calculation of this standard deviation, and dividing by it, has been made part of the MF. This can be regarded a normalization of the filter, restricting the amplitude effect of signal model changes to the signal component in the output.

This paper deals with signals that consist of overlapping peaks. When the positions and the shapes of the peaks are known, normal matched filters will suffice, one filter for each shape. In the simplest situation of overlapping peaks of identical shape only one filter is required. The output is measured at the positions of the maxima of the individual peaks. From these values disturbed by overlap, undisturbed values can be calculated by means of the models used. However, when the exact position or shape of one or more peaks is not known, this approach will fail. Using a fixed though incorrect position results in both systematic errors and higher relative standard deviations in the individual peak intensity estimations. Estimating the position from an ensemble of data results in the transfer of part of the systematic error to the random error. Incomplete information on the shape of the peaks has analogous effects. The MFX will be hindered by the overlap, since the contributions from the overlapping peaks are described by neither the peak shape model nor the noise model.

An answer to these problems is found in using one multi-peak MFX instead of several single-peak filters. A normal, rigid MF is no longer realistic with such a model, because not only the positions and the shapes would need to be known, but also the intensity ratios of the peaks. Such a MF would only be able to quantify the entire cluster, not its constituents. In the MFX however, intensity ratios, position differences and, if necessary, other parameters in the multi-peak signal model can be estimated by the normal procedure of optimizing the output S/N. The intensities of the individual peaks can be calculated from the output S/N and the estimations for the model parameters that have arisen from the optimization. In this paper it is assumed that the signal model is correct except for the values of the parameters in the model.

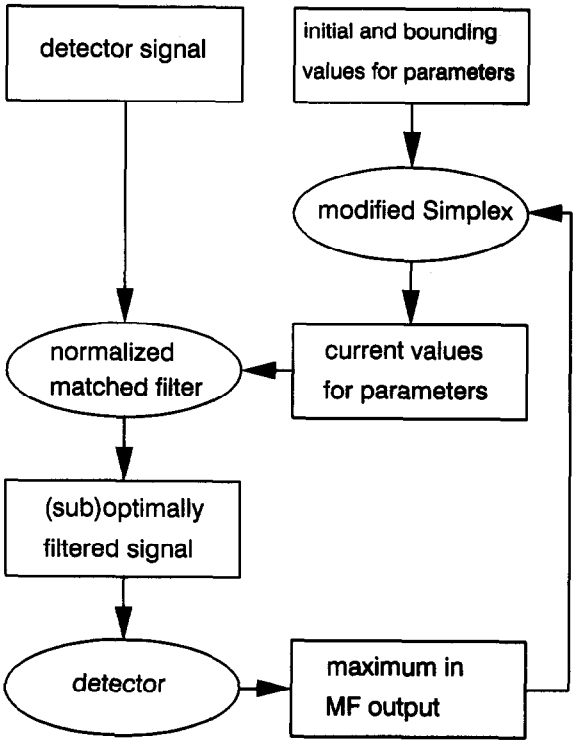

Fig. 1. Flow chart of the MFX. Ellipses and rectangles represent operators and data respectively.

The present research on the multi-peak MFX is limited to signals containing two peaks: two Gaussians of equal width in the presence of first-order band-limited noise with a zero baseline. The working of a double-peak MFX is illustrated in Figs. 1 and 2. A mathematical description will be given in the next section, followed by a discussion of the expected quality of the results.

It is realized more readily than in the single-peak situation, that the MFX is a fitting procedure in which the goodness-of-fit is measured by a S/N. A fitting procedure, moreover, that is able to deal with correlated noise. The comparison with other fitting procedures and quantification methods will not be elaborated in this paper, though a tentative comparison with non-linear regression is made in the discussion. It is clear that overlap poses a serious problem to quantification based on integration, the most popular approach. Most of the integration methods are not documented sufficiently to allow a thorough statistical comparison, the required information usually being proprietary [7,12]. Furthermore, the number of available techniques and the num-

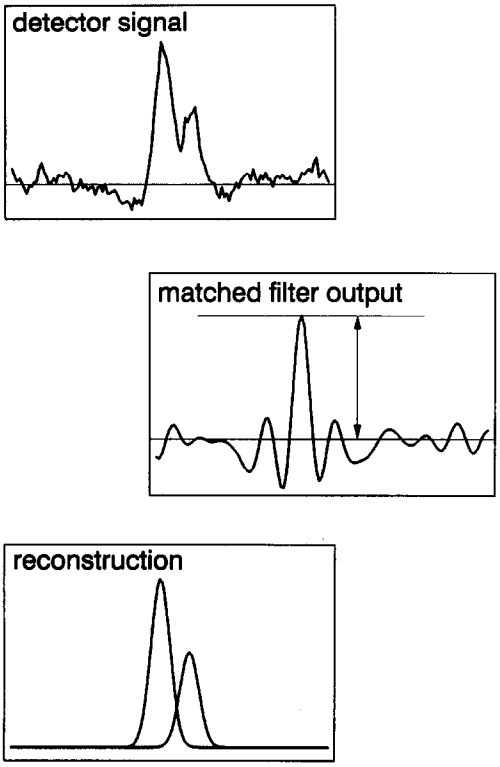

Fig. 2. Input and output of a MF being part of an MFX, for correlated noise.

ber of parameters for each technique are so large that every comparison is in fact arbitrary. A solution to the problem of comparison might be a large public domain data set representing the problems that may be encountered in practice, or the software needed to generate such a data set. Every researcher and supplier could then assess the capabilities of his approach with regard to this data set. Recently, a general chemometrics reference data base was initiated [13]. The first entries are multivariate data sets, but hopefully univariate data will also be added.

2. Mathematical description of a double-peak MFX

A mathematical description will be given of a double-peak MFX without making further assumptions regarding noise or peak models. A signal $x(t)$ consisting of two peaks can be written as:

$$x(t) = A_{1x_1}(t) + A_{2x_2}(t) \tag{1}$$

where the $x_i(t)$ represent the shapes of the peaks and the A_i their amplitudes. The model $m(t)$ is defined analogously:

$$m(t) = m_1(t) + A_m m_2(t) \tag{2}$$

where A_m is the amplitude ratio of the model peaks. The complex frequency response of the normalized filter being matched in the MFX procedure is:

$$H_n(jf) = \frac{M^*(jf)}{\sigma_{\text{out}} S(f)} \tag{3}$$

where $M(jf)$ is the Fourier transform (FT) of the signal model $m(t)$, σ_{out} is the standard deviation of the output noise, $S(f)$ is the power spectral density of the input noise and the asterisk denotes the complex conjugate. The output of the filter is:

$$\begin{aligned} y_n(t) &= \text{FT}^{-1}[H_n(jf)X(jf)] \\ &= \frac{A_1}{\sigma_{\text{out}}}\text{FT}^{-1}\left[\frac{M_1^*(jf)X_1(jf)}{S(f)}\right] \\ &\quad + \frac{A_2}{\sigma_{\text{out}}}\text{FT}^{-1}\left[\frac{M_1^*(jf)X_2(jf)}{S(f)}\right] \\ &\quad + \frac{A_{1A_m}}{\sigma_{\text{out}}}\text{FT}^{-1}\left[\frac{M_2^*(jf)X_1(jf)}{S(f)}\right] \\ &\quad + \frac{A_{2A_m}}{\sigma_{\text{out}}}\text{FT}^{-1}\left[\frac{M_2^*(jf)X_2(jf)}{S(f)}\right] \end{aligned} \tag{4}$$

where $X(jf)$ and $X_i(jf)$ are the FTs of the signal $x(t)$ and the individual peaks $x_i(t)$, and FT^{-1} denotes the inverse FT. The maximum of this function will be found at $t=0$ when the signal model is correct. In general, the signal model is correct when it differs from the true signal only by some multiplication factor. The size of that factor is irrelevant, because it amplifies both signal and noise and does not affect the output S/N. Therefore, the model being correct can be expressed as $x(t) = A_1 m(t)$. In that case, the inverse FTs of Eq. 4 can be written as:

$$R_{ik} = \int_{-\infty}^{+\infty} \frac{M_i^*(jf)M_k(jf)}{S(f)} df \tag{5}$$

where $i \in \{1,2\}$ and $k \in \{1,2\}$. R_{12} and R_{21} are single points in the cross-correlations of the peaks that have been transformed by the division by $S(f)$. The division by the square root of $S(f)$ is often referred to as whit-

ening, because it whitens the input noise prior to the cross-correlation [1]. R_{12} and R_{21} are equal for any combination of peak shapes; the output of the filter with correct signal model, being an auto-correlation, will always be symmetrical. R_{11} and R_{22} are the powers of the peak models, transformed by $S(f)$. Using these definitions, the maximum in the output can be written as:

$$y_{n,\max} = \frac{A_1}{\sigma_{\text{out}}}[R_{11} + 2A_m R_{12} + A_m^2 R_{22}] \tag{6}$$

The term between square brackets represents the power of the entire cluster model transformed by $S(f)$. A property of the MF is that this power is equal to the variance of the noise in the output of the filter without normalization, being the square of the normalization factor σ_{out}. When the optimization comes to an end, it is assumed that the current signal model is correct, and therefore the current normalization factor can be used to estimate the amplitude A_1 from the current $y_{n,\max}$:

$$A_1 = \frac{y_{n,\max}}{\sigma_{\text{out}}} \tag{7}$$

The amplitude of the second peak is calculated from this result simply by using the ratio A_m that was arrived at by the optimization:

$$A_2 = A_m A_1 \tag{8}$$

3. Error estimation for a double-peak MFX

It is expected that the method will work, though not equally well under all conditions. How well it works will, for the moment, be associated with the standard deviations of the results. The factors that determine its success may be expected to be the degree of overlap, the size of the peaks relative to the noise and the amount of a priori information that is supplied to the detector and the simplex optimization. The success will decrease when the peaks are brought closer together and when the size of one or both of the peaks is decreased at a constant noise level. In Fig. 3, an array of noiseless input data consisting of two Gaussians of varying position difference and amplitude ratio is displayed. Resolution is identified with the ratio of peak position difference (t_x) to peak width (σ_x). It may be

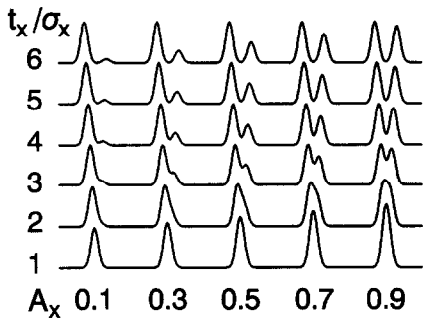

Fig. 3. Signal component in the data in the experiments with the double-peak MFX.

expected that the situations with the lowest resolutions, $t_x/\sigma_x \leq 2$, and those with the smallest second peak, $A_x = 0.1$, will be problematic, as will, to a lesser extent, the combination of $A_x = 0.3$ and $t_x/\sigma_x = 3$. In all of these cases, the difference between the peaks is so small that it will be easily obscured by noise. Especially when the procedure is allowed to adapt the width parameter of the peaks, the clusters with the lowest resolutions may be fitted by a single wider peak.

The key issue is the response surface on which the optimization has to find its way, i.e. the response as a function of the parameters in the signal model. A practical response surface will consist of contributions from both the signal and the noise component in the input data. Here, the signal contribution is used to calculate the distribution of the results. For instance, when optimizing two parameters in an otherwise correct signal model, the response surface is a three-dimensional entity, some mountainous area. The signal contribution will have a global optimum at the location of the correct parameter values. A series of superimposed practical response surfaces can be thought to create a haze over the noiseless surface. The top of the mountain will be hidden in the clouds and a certain area around the top will be indistinguishable.

The cloud around the top is the three-dimensional distribution of the maxima. Each maximum is described by three coordinates, being the two parameters and the amplitude of the cluster output. The estimations of the amplitudes of the individual peaks are calculated from these coordinates. Projection of the cloud on the ground plane gives the two-dimensional distribution of the parameter estimations. Iso-probability contours of this distribution are assumed to coincide with contours of the signal contribution to the response

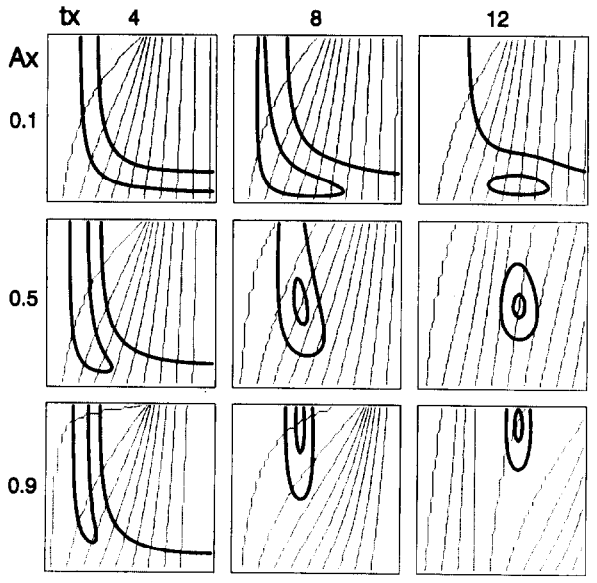

Fig. 4. Amplitude (drawn) and position (dotted) of response of double-peak white noise MFX. X axis: $t_m[0; 20]$; Y axis: $A_m[0; 1]$.

surface. In other words, when one cuts through the noiseless response surface at some level below the top, one obtains a contour of the distribution of the parameter estimates. This is similar to one of the approaches to error estimation in non-linear regression [14,15].

When specific models are substituted in Eq. 4, the signal contribution to the response surface can be investigated mathematically. For two overlapping Gaussians of equal width in the presence of first-order band-limited white noise, the formula is given in the Appendix. No closed analytical solution to the problem of finding the maximum in the output has been found, due to the fact that the output consists essentially of four cross-correlations. The position of the absolute maximum of their envelope depends on the positions and amplitude ratios of the individual peaks. The maximum is found numerically, by scanning the record.

In Fig. 4, contour plots are given for a range of values for the signal parameters t_x and A_x. The noise time constant $\tau = 1$, the width of the peaks $\sigma_x = 4$ and the amplitude of the first peak is always one. The units of time in which σ_x, t_x, t_m and τ are expressed are arbitrary. The contours are located 0.003 and 0.03 below the maximum of the response surface. Fig. 4 also presents the respective positions of the maximum in the output of the filter, as the contours of the position surface at discrete point positions. The time constant $\tau = 1$ corresponds with approximately white noise. In theory, the time constant can be made smaller than unity, but such noise cannot be represented by discrete data with a reasonable sampling frequency [1].

In case of a model that is linear in its parameters, the parameter estimations are normally distributed and the contours of the response surface are elliptical. The quality of the estimations, when measured by the standard deviation, is visible as the relative size of the ellipse. Following the 0.03 contour in Fig. 4, it is observed that the distributions of the parameter estimations are nearly elliptical, except for the situations in which the second peak is only very small ($A_x = 0.1$) or the overlap is high ($t_x = 4$). The shape and size of the contours depend on the noise level, a lower level corresponds with higher contours and vice versa. For instance at $A_x = 0.1$ and $t_x = 12$, the higher contour is elliptical, whereas the lower one is not. It seems that a linear approximation is appropriate under all but the most adverse conditions. The non-normality can be understood as follows: if $t_m = 0$ then A_m may have any value and vice versa. For $t_x = 4$ and $A_x = 0.1$ the probability that this happens is large. When A_x is increased, $A_m = 0$ becomes less probable and when t_x is increased $t_m = 0$ becomes less probable, though at $A_x = 0.1$ the peaks have to be rather well separated in order to cut off the $t_m = 0$ path. When a time constant $\tau = 40$ is used in the calculations, which corresponds with highly correlated noise, slightly less separation is required. This influence on the t_m scale of the contours is visible in all plots: the correlated contours are always narrower than the white ones. For the rest, the response surfaces for white and correlated noise are very similar.

In the situations where the parameters may escape from the simple ellipse the maximum in the filter output may be found on several positions. On the parameter domains that have been plotted the number of different positions ranges from 5 to 10 for the 0.01 contour. In the elliptical situations the maximum is consistently found at one location, being that of the maximum of the noiseless output. A lower noise level will correspond with higher contours and more different positions may be found. In case of correlated noise there are sharp ridges in the position surfaces far from the optimum, caused by jumps from one maximum to another.

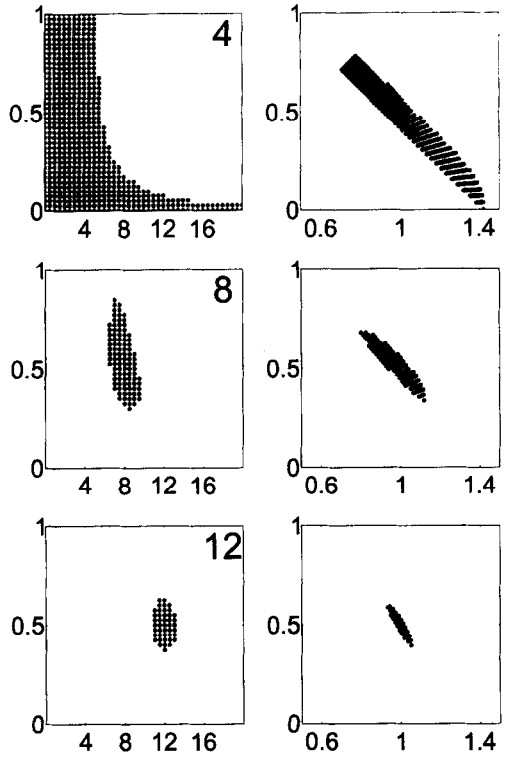

Fig. 5. Responses (left, $A_m - t_m$) mapped onto amplitudes (right, $A_2 - A_1$); white noise, $A_x = 0.5$.

The ultimate goal of the MFX is not to estimate the parameters, but the individual amplitudes of the peaks that are calculated from the parameter estimates and the response itself using Eqs. 7 and 8. The distribution of the amplitudes of the maximum responses is not known, and therefore, in order to obtain an impression of the distributions of the amplitudes, the distribution of the parameters is mapped onto the surface spanned by the individual peak amplitudes by assuming that the response always has the value of the theoretical noiseless maximum. In that way, every parameter combination corresponds with a single amplitude pair. In reality it will correspond with some distribution along the line through the origin and the above amplitude pair, i.e. every point in the mapping of the parameter estimations will be drawn out to form a line. In Fig. 5 the mapping is displayed for white noise, one amplitude ratio and several position differences. The response surface has been scanned to give a grid of points. Only those points that are above the 0.01 contour under the top have been mapped. All features of the contours of the response surface are visible in the mapping. The path along the A_m axis in the response surface translates into smearing of the distribution to form a broad stroke. The path along the t_m axis has little effect, because, with $A_m = 0$, it cannot get off the A_1 axis. It is assumed that the smearing will cause the distribution of the amplitudes to deviate from normality. The cloud of mapped points is smaller when the peaks are further apart, indicating an increasing quality of the results. When the mapping is performed for correlated noise, the same observations can be made.

4. Error estimation for two single-peak MFs

In the presence of complete and accurate a priori information on the data, one or two normal single-peak MFs can be used for the quantification of the individual peaks, and it is possible to make a theoretical estimation of the error, the standard deviation, in the amplitude estimations. These estimations will be used as a reference for the experimental results that are obtained with the double-peak MFX in more demanding situations. It is expected that this reference sets a lower limit to the experimental error. It is assumed that the input noise follows a Gaussian distribution, and so does, therefore, the output noise.

For the signal defined by Eq. 1, two filters are defined instead of one:

$$H_i(jf) = \frac{M_i^*(jf)}{S(f)} \tag{9}$$

$M_i^*(jf)$ is the complex conjugate of the FT of $m_i(t)$ and $i \in \{1,2\}$. The models are correct, i.e. $m_i(t) = x_i(t)$. Since they are fixed filters, they do not need to be normalized. The output of each filter is measured at the position where the respective filtered peak has its maximum. With the filters defined as they are, this is at $t = 0$ in both outputs:

$$y_i = \int_{-\infty}^{+\infty} H_i(jf) X(jf)\,\mathrm{d}f \tag{10}$$

Again $i \in 2\}$. With Eq. 5, this becomes:

$$y_1 = A_1 R_{11} + A_2 R_{12}$$
$$y_2 = A_1 R_{12} + A_2 R_{22} \tag{11}$$

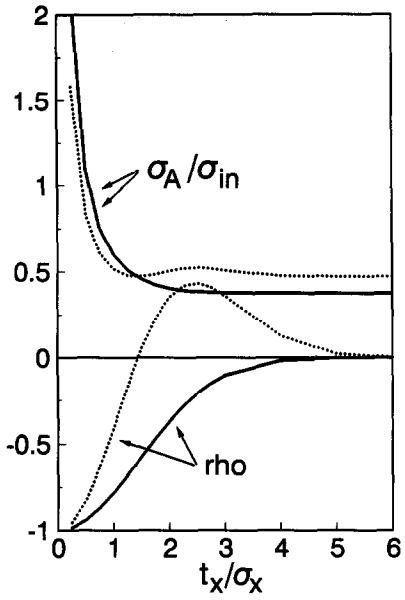

Fig. 6. Error estimation for two single-peak MFs for white (drawn) and correlated (dotted) noise.

Note that R_{12} and R_{21} are identical here as well as in their original context. The equations can be solved for the individual amplitudes:

$$A_1 = (y_1 R_{22} - y_2 R_{12})/D$$

$$A_2 = (-y_1 R_{12} + y_2 R_{11})/D$$

$$D = R_{11} R_{22} - R_{12}^2 \qquad (12)$$

The variances of these estimators, and their correlation, can be calculated using simple error propagation for a linear combination of two stochastic variables. It is easily derived that R_{11} en R_{22} are the variances of the output noises of the two filters [1]. Analogously, R_{12} is the covariance between these noises. In formula:

$$R_{11} = \sigma_{y_1}^2; \quad R_{22} = \sigma_{y_2}^2; \quad R_{12} = \mathrm{cov}_{y_1 y_2} \qquad (13)$$

which leads to

$$\sigma_{A_1}^2 = \frac{R_{22}}{D}; \quad \sigma_{A_2}^2 = \frac{R_{11}}{D}; \quad \rho_{A_1 A_2} = \frac{-R_{12}}{\sqrt{R_{11} R_{22}}} \qquad (14)$$

Even without substituting specific peak models, a trend in the quality of the peak amplitude estimations as a function of overlap can be observed. R_{12} will approach zero when the peaks are far apart, leading to minimal variances:

$$\sigma_{A_1}^2 = R_{11}^{-1}; \quad \sigma_{A_2}^2 = R_{22}^{-1} \qquad (15)$$

When the peaks are closer together, R_{12} is larger, D will become smaller and the variances will be larger. R_{12} and ρ can serve as measures of overlap or resolution. The former depends on the scale of the models, the latter does not.

For two Gaussians of equal width in the presence of first-order band-limited noise, the effect of overlap on the estimated error is plotted in Fig. 6. The calculation has been performed for correlated and approximately white noise. The curves agree with the trend described above. The quality of the estimation decreases significantly only when the peaks are less than about 1σ apart. In case of correlated noise, the standard deviation shows a little bump where the correlation coefficient changes sign. This can be explained by examining the impulse responses of the MF for correlated noise, displayed in Fig. 7. The impulse responses have negative side lobes that can coincide with each others maximum. The presence of the side lobes causes a second overlap and therefore a second increase of the estimation error. The sign of the lobes causes the correlation to become positive instead of negative.

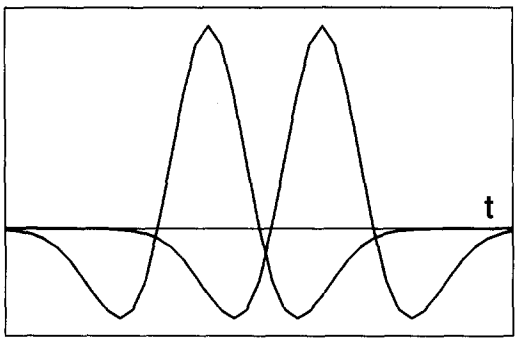

Fig. 7. MF impulse responses for Gaussians with $\sigma_x = 4$ and correlated noise, for $t_x = 9$.

5. Experimental

The overall structure of the experimental setup is: two types of noise (white and correlated), two levels of a priori information, two noise levels, and a range of position differences and amplitude ratios. Details are given below.

The experiments have been carried out with simulated data, 1000 different records each. In each record

the signal component is the same, but it is disturbed by a different noise realization. The noise has been generated using a software pseudo-random generator and shaping filter as described previously [1].

The width of the peaks, σ_x, was 4 points, fulfilling the demands of the sampling theorem [16]. The noise had zero mean, so the amplitude of the output of the filter could be measured simply from zero. As in earlier work [16], two values for the noise time constant have been used: $\tau = 0.01$ points (coded WH), defines noise that is virtually white, $\tau = 40$ points (coded CL), defines very correlated noise.

The experiments are characterized by the S/N for the largest peak, always the first peak in the cluster, and the amplitude ratio of the peaks ($A_x = A_2/A_1$). S/N was set at 10 and 100, A_x was set at 0.1, 0.3, 0.5, 0.7 and 0.9. For each combination, the position difference in the input, t_x, ranged from one to six times the sigma of the Gaussian, i.e. from 4 to 24 points with increment 4. The data defined by the different A_x and t_x values are plotted in Fig. 3. Note that these data are theoretical. When the different resolutions would correspond to separation conditions for some sample, the peaks would also differ in width and height, which they do not. The amplitude of the first peak was always unity, the S/N was controlled via the variance of the noise. The seeds for the pseudo-random generator at $S/N = 10$ where identical to those at $S/N = 100$, so that only the amplitude of the noise contribution was changed between those series.

The input records contain 1024 points, with the peak cluster positioned in the centre. The position difference in the signal model is a continuous variable, but the peaks in the input are always located on discrete point positions. One of the peaks in the model is always on a discrete point position too. A filter length of 128 points allows the filter operation to be performed as a linear convolution in the time domain. The principle of detection was simply to choose the maximum in the entire record.

Two runs of the experiments have been made. In the first, the simplex optimization had to adapt the amplitude ratio and the position difference of the peaks. In the second run, the parameter controlling the width of both peaks was adapted as well. The expectation was that the second run would meet more problems than the first.

The vertices of the initial simplex were calculated according to literature [17]. The initial and bounding values for the parameters were:

	start	step	min	max
A_m	$A_x - \sigma_{in}$	$2\,\sigma_{in}$	0	2
t_m	$t_x - \sigma_x$	$2\,\sigma_x$	0	100
σ_m	$\sigma_x - 1$	3	0	40

where σ_{in} is the standard deviation of the input noise. Part of the function of the selected boundaries is to prevent a complete symmetry in the response surface. The combination of boundaries reflects a preference towards making the first peak the largest. The robustness of the simplex may be checked by restarting it, but any system for automatic generation of alternative initial values will be as arbitrary as the existing single set.

6. Results

The ultimate result of each simulation is a pair of amplitude estimates that are part of a two-dimensional distribution. Normally it is assumed that a description of that distribution by all 1000 points is redundant and that statistics can reduce it to a set of five values: two means, two variances and one covariance. This reduction is possible only when the assumption is valid that the distribution is Gaussian. However, when the amplitude estimations are plotted against each other as scatter plots as in Figs. 8 and 9, it is observed that the assumption of a two-dimensional Gaussian distribution is not always valid, especially at low S/N, when there is a high degree of overlap and the amplitude ratio is low. The non-normality hinders the data reduction necessary for the presentation and the evaluation of the results, but it also represents information on the quality of the results. The non-normality has been tested using the modified Anderson–Darling statistic, with a 2.5% level of significance for each dimension [18].

The results of the Anderson–Darling test are collected in Table 1, together with the test results for the means, standard deviations and correlation coefficients. When a test fails, the magnitude of the test statistic serves as an indication of the degree of failure. The logarithm of the Anderson–Darling statistic is used, because the statistic spans orders of magnitude. The

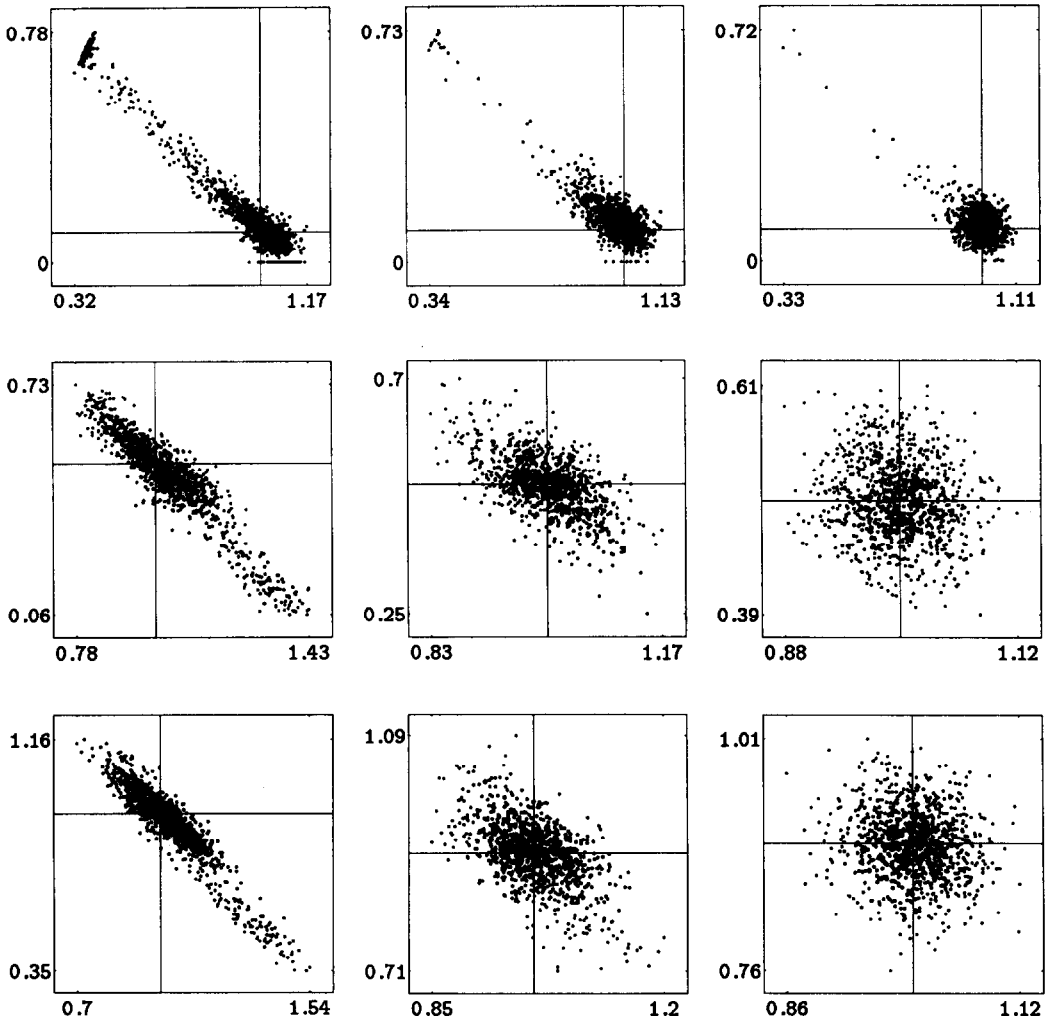

Fig. 8. Amplitude estimations in WH run 1 at $S/N = 10$. X axis: A_1; Y axis: A_2. Graphs ordered as in Fig. 4.

critical value for the statistic at a 2.5% level of significance is 0.873. The means are tested as the difference with the true value, divided by the estimated standard deviation of the mean. The critical value at a 5% significance level in a two-sided test is 1.96. The standard deviations are tested as the ratio of the experimental value to the value calculated for two single-peak MFs. The hypothesis tested is that this ratio equals unity, with critical values 0.955 and 1.043. Table 1 gives the difference between unity and the empirical ratio as a percentage of unity. The correlation coefficient is made to follow a normal distribution by performing the Fisher transformation and the test statistic is the same as for the means. The statistic spans orders of magnitude and can have both positive and negative values. Therefore the logarithm of its absolute value is printed.

7. Discussion

In view of the large number of simulations that has been performed for each experiment, there is a remarkable amount of diversity in the results in Table 1. The expected trends cannot be verified or falsified unambiguously. Nevertheless, the general picture is that the results are in accordance with the expectations: the proposed procedure works, though not equally well under all circumstances, and the distributions of the

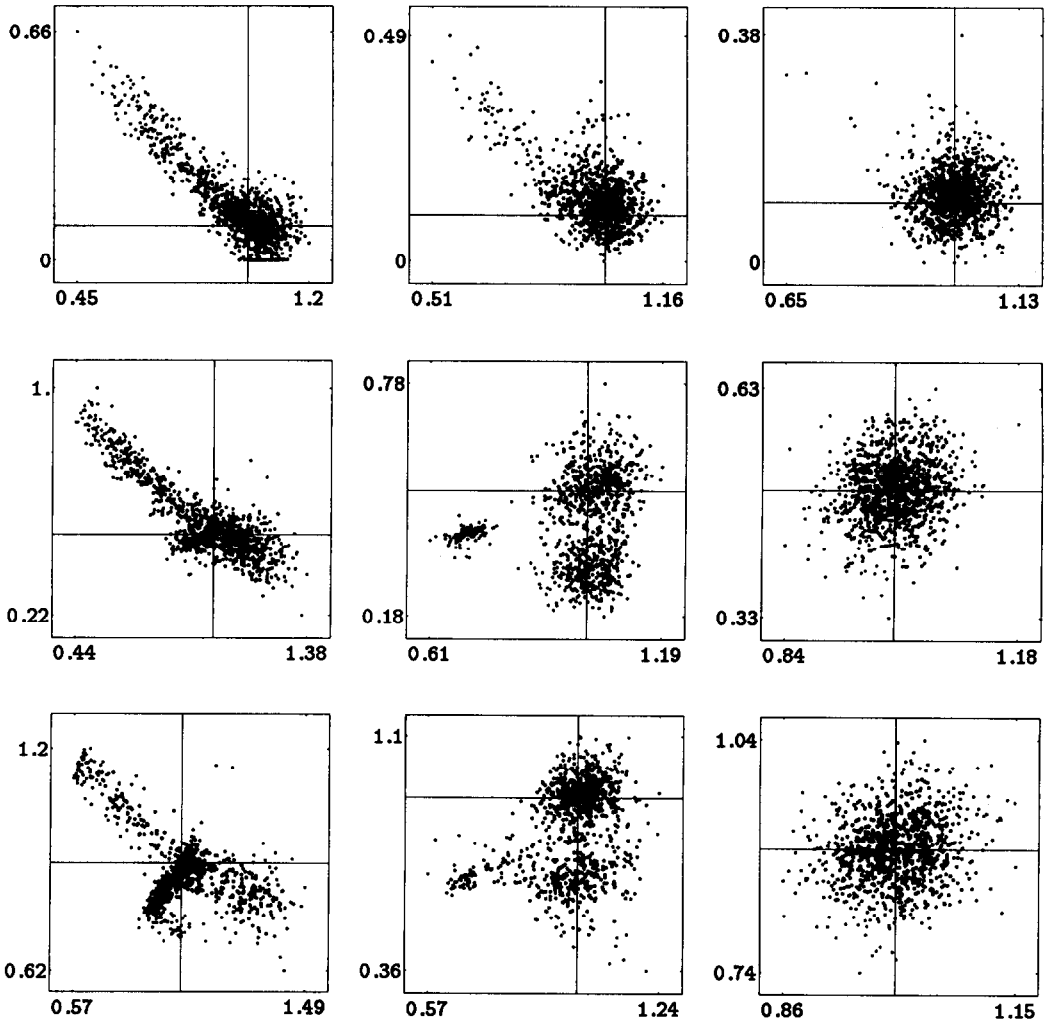

Fig. 9. Amplitude estimations in WH run 2 at $S/N=10$. X axis: A_1; Y axis: A_2. Graphs ordered as in Fig. 4.

amplitude estimations deviate from normality in difficult situations. A situation is called difficult when the resolution is low, $t_x/\sigma_x \leq 2$, or the second peak is small, $A_x=0.1$. Those are the conditions for which the response surfaces indicated that deviations could be expected. Although the abnormalities complicate the quantification of the quality of the results, the conclusion must be that they make the results of lower quality.

The deviations from normality present themselves as distributions that are drawn out in the direction of underestimation of A_1 and overestimation of A_2, as in Fig. 8. The deformation is blocked by the upper limit that has been set to the simplex optimization, being $A_m=2$. In run 1 the escapes tend to be on this limit, whereas in run 2, they mark the path rather than being on the limit. The results for correlated noise are slightly better than for white noise and the splitting up as in the WH run 2 does not occur.

For the simple situations, the results of run 1 are in agreement with the estimations for two separate single-peak MFs. Table 1 shows that neither the means, nor the standard deviations nor the correlation coefficients differ significantly. In other words, the difference between the more complex, essentially non-linear approach and the simple linear approach is in the occurrence of abnormalities in the difficult situations. Associating an elliptical contour of the response surface with normally distributed results appears to be appropriate

Table 1
Statistical testing of the experimental results. (.) success; (*) failure coinciding with failure of normality test; numbers: magnitude of test statistics; (+) infinite normality statistic. All numbers are rounded to the nearest integer, coding described in text

		normality S/N 100				10	mean 100				10	sigma 100				10	rho 100				10
WH run1	a1	4	1	.	0	. 0	2	1	1	1	1	*	-2	.	.	*	*	*	*	*	*	* 1 * * * * * * * * * *
		8	.	.	.	. 0	2	0	0	0	.	.	.	.	.	.	*	*	.	*	.	25 21 25 28 * * * * * 32 * 0 1 1 * * * * * *
		12	.	.	.	. .	2	.	.	.	.	.	.	.	.	.	*	.	.	.	.	. 5 . 6 7 * 6 . . . . . 1 0 * . . . .
		16	.	.	.	. .	+	.	.	.	.	.	.	.	.	.	*	.	.	.	.	. . 4 . 5 * . . . . . 0 * . . 0 .
		20	.	.	.	0 .	.	.	.	.	0	.	.	.	.	.	.	.	.	.	.	. . . . * . . . . . * . . . . .
		24	.	.	0	. .	.	.	.	0	.	.	.	.	.	.	.	.	.	.	.	. . . . . 0 . . . . .
	a2	4	1	.	0	0 0	2	1	1	1	1	*	.	.	.	*	*	*	*	*	*	* 85 * * * * * * * * *
		8	0	.	.	. .	+	0	0	1	1	.	.	.	.	.	*	*	.	.	.	* 10 10 10 5 * * * * *
		12	.	.	.	. .	+	.	.	.	.	-2	.	.	.	.	*	4	.	.	.	5 . . . . * . . . .
		16	0	.	.	. .	+	.	.	.	.	.	.	2	.	.	*	.	3	.	2	* . 7 . . * . . . .
		20	.	.	.	. .	0	.	.	.	.	.	.	.	.	.	*	.	.	.	.	7 6 . 9 5 . . . 5 .
		24	.	.	.	. .	.	.	.	.	.	.	.	.	.	.	7	3	.	.	.	5 . . 4 6 . . -5 . .
WH run2	a1	4	1	1	0	1 1	2	1	1	1	1	*	*	*	*	*	*	*	*	*	*	* * * * * * * * * * * * * * * * * * * *
		8	.	.	+	2 0	1	2	2	2	1	.	.	*	.	*	*	*	*	*	*	39 48 * * * * * * * * * 1 * * * * * * * * * *
		12	.	.	.	. .	0	.	.	.	.	.	.	.	.	.	.	.	.	.	.	25 27 24 26 21 * 25 21 23 22 1 1 1 1 1 . 1 1 * 1
		16	.	.	.	. .	.	.	.	.	.	.	.	.	.	-2	.	.	.	.	.	22 25 18 19 24 25 25 17 16 14 1 1 1 1 . 0 1 1 1 1
		20	.	.	.	1 .	.	.	.	.	.	.	.	.	.	*	.	.	.	.	.	24 20 20 18 * 23 19 18 15 14 . 1 1 1 * . 1 1 1 1
		24	.	.	.	. .	.	.	.	.	.	.	.	.	.	-2	.	.	.	.	.	23 21 20 16 20 21 21 19 14 12 . 0 1 1 . . 0 1 1 1
	a2	4	1	1	0	+ +	2	1	2	2	2	*	*	*	*	*	*	*	*	*	*	* * * * * * * * * *
		8	.	+	2	2 +	1	0	1	1	1	3	.	*	*	.	*	*	*	*	*	32 * * * * * * * * * *
		12	.	.	.	. .	+	.	.	0	.	.	.	.	.	-3	*	4	.	.	.	13 15 17 26 20 * 16 16 * 16
		16	.	.	.	0	.	.	.	.	.	.	.	3	.	*	9	.	3	.	2	6 6 12 9 * . . 10 8 8
		20	.	.	.	1	0	.	.	.	.	.	.	.	.	*	*	.	.	.	.	6 7 10 16 * . 5 5 12 9
		24	.	.	.	+	.	.	.	.	.	.	.	.	.	.	7	3	.	.	.	6 . 5 12 * . . . 8 12
CL run1	a1	4	.	.	.	. .	2	1	0	0	1	-2	.	.	.	.	*	.	*	*	*	94 110 102 114 100 * * * * * 1 1 1 * 1 * * * * *
		8	.	.	.	. .	2	.	0	.	.	.	.	.	.	.	*	.	.	.	.	8 . . 8 5 * 8 . 8 . . . . . * . . . .
		12	.	.	.	. .	2	2	.	.	.	.	.	.	.	.	.	.	.	.	.	. . . . 5 * * . . 5 . * * . . . . .
		16	.	0	.	. .	0	0	.	.	.	.	.	.	.	.	*	.	.	.	.	. . . . . * . . . . . * . . . .
		20	.	.	.	. .	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	. . . . . 0 . . . 0 0 . . . .
		24	.	.	.	. .	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	. . . . 7 . . . . 6 . . . . 0 . . . . .
	a2	4	.	.	0	.	2	+	2	2	2	2	2	.	.	.	*	.	*	*	*	22 28 19 * 28 * * * * *
		8	.	.	.	. .	+	.	.	.	0	.	.	.	.	.	*	4	.	.	.	. 7 5 . . * 5 . 4 .
		12	.	.	.	. .	+	+	.	.	0	.	.	.	.	.	*	*	.	.	.	. . -5 . . * * -5 . .
		16	.	.	.	. .	+	.	.	.	.	3	-2	.	.	.	*	.	3	2	2	. . 5 . . . .
		20	.	.	.	. .	.	.	.	.	.	7	.	.	.	.	.	.	.	.	.	. . . . . -8 . . . .
		24	.	.	.	. .	.	.	.	.	.	2	.	.	.	.	11	4	2	.	3	. 5 . 4 . -7 . . . .
CL run2	a1	4	1	0	0	1 2	1	2	2	1	1	*	.	.	.	*	*	*	*	*	*	* * * * * * * * * * * * * * * * * * * *
		8	.	.	.	. .	1	0	.	.	.	.	.	.	.	.	*	*	.	.	.	19 10 . 9 9 * * 4 8 8 1 0 . 0 . * * . . .
		12	0	.	.	. .	0	.	.	.	.	.	.	.	.	.	*	12	17	11	11	* 13 17 11 11 * 1 1 1 1 * 1 1 1 1
		16	.	.	.	. .	.	.	.	.	.	.	.	.	.	.	4	.	.	.	.	13 18 18 23 23 13 16 18 21 21 1 1 1 1 * * * 1 1 1
		20	.	.	.	. .	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	6 9 9 12 12 7 9 9 10 10 1 . 1 0 . 1 0 1 1 1
		24	.	.	.	1 .	.	.	.	.	.	.	.	.	.	*	.	.	.	.	.	8 9 5 9 * 9 8 5 5 5 . . . . * . . 0 0 .
	a2	4	.	.	0	1 +	2	2	2	2	1	4	4	.	.	*	*	*	*	*	*	32 32 * * * * * * * * *
		8	.	.	.	. .	+	.	.	.	0	.	.	.	.	.	*	8	.	.	.	7 11 . 8 8 * . 7 5 *
		12	.	.	.	. .	+	.	.	.	.	.	.	.	.	.	*	.	.	.	.	7 6 7 12 12 . 5 5 10 10
		16	.	.	.	0	0	0	.	.	.	3	.	.	2	*	*	.	3	3	3	10 13 19 22 * . * 16 15 15
		20	.	.	.	. .	.	.	.	.	.	.	.	.	.	.	8	.	.	2	3	4 . . 6 6 -5 . . . .
		24	.	.	0	1	.	.	.	.	.	3	.	2	.	*	11	4	3	.	3	. 5 6 * * -6 . . 6 6

here. Comparing the standard deviation columns of Table 1 with Fig. 6, it is observed that the increase of the error in the amplitude estimations as predicted for low resolutions is often obscured by the abnormalities. It would be interesting to check if this failure of the linear approximation is also valid in case of non-linear regression. In that case, error estimations for the results of fitting overlapping peaks by means of non-linear regression [19] should be reconsidered.

Estimation of an additional parameter, the peak width, as in run 2, increases both the standard deviations and the occurrence of deviations from normality. The agreement with the theoretical reference as found in the simple situations of run 1 is absent here. Furthermore, in the white noise experiments of run 2 there are some remarkable deviations in the results when the peaks are 2σ apart even at $S/N = 100$. In that case, the distributions are not smeared out in the upper left direction as in the expected problem areas of run 1, but drawn out to form several clusters (Fig. 9). A similar deviation occurs at $S/N = 100$ when the peaks are further apart and $A_x = 0.9$: a small group of points is separated from the main cloud.

The increase in the standard deviations of the amplitude estimates observed in the normal regions of run 2 is independent of the S/N, but larger for A_1 than for A_2.

Table 2
Statistical testing of results of run 2 experiments at $S/N=10$ with different start-up of the simplex. Coding as in Table 1

		normality	mean	sigma	rho
WH	a1 4	1 1 . 1 2	. . . . .	* * 323 * *	* * * * *
	8	0 . . 2 2	. . . . .	* 46 59 * *	* 1 1 * *
	12	. . . . .	. . . . .	24 26 24 24 26	1 1 1 1 1
	16	. . . . .	. . . . .	22 25 19 16 17	0 1 1 1 1
	20	. . . . .	. . . . .	25 20 19 18 17	. 1 1 1 1
	24	. . . . .	. . . . .	21 21 20 17 14	. 0 1 1 1
	a2 4	+ 1 1 + 2	. . . . .	* * * * *	
	8	. . . 2 +	. . . * *	31 46 52 * *	
	12	. . . . .	. . . . .	14 19 18 28 20	
	16	. . . . .	. . . . .	. 6 15 11 14	
	20	0	. . . . .	* 9 8 14 12	
	24	. . . . .	. . . . .	6 . . 9 16	
CL	a1 4	1 0 1 1 1	* . * * *	* * * * *	* * * * *
	8	. . . . +	. . . . *	19 10 . 8 *	1 0 . . *
	12	. . . . .	. . . . .	13 13 17 12 15	1 1 1 1 1
	16	. . . . .	. . . . .	13 17 19 21 17	1 1 1 1 *
	20	. . . . .	. . . . .	6 10 10 12 11	1 . 1 0 0
	24	. . . . .	. . . . .	9 8 5 7 13	. . . . .
	a2 4	. 0 0 1 +	-5 * * * *	31 * * * *	
	8	. . . + +	. . . -1 *	6 10 7 . *	
	12	. . . . .	. . . . .	7 8 7 11 8	
	16	 0	-2 2 . . .	9 12 18 14 16	
	20	. . . . .	2 2 . . .	5 . . 7 5	
	24	. . . . .	-2 . . . -2	. 5 . 10 8	

For CL, it rises to a small maximum at $t_x=16$. For the mean the picture is erratic. The effects visible in the standard deviation and the normality statistic are not consistently reflected by the means. It should be noted that a test on the mean of 1000 observations is rather sensitive. The distributions of the individual observations are 30 times wider, and there are no observations that significantly deviate from the true value, apart from the outliers that occur in case of abnormalities. When the $S/N=10$ and $A_x=0.1$, the means of A_2 generally have a positive bias. The source of this error is found to be the estimation of A_m, which is not correlated with any of the other parameters and has a positive bias that is passed on to A_2. The probable explanation of this bias is noise fitting. Its sign is the result of the maximization inherent to the MFX. Noise fitting can also explain the low standard deviations in the estimations of the second peak: maxima in the noise have a smaller standard deviation than the noise, and A_2, being dominated by a maximized noise contribution, will have relatively little random error.

The splitting-up of some of the distributions in run 2 is one of the most striking features of the results, especially since they are also observed at $S/N=100$. Visualization of the four-dimensional response surface of run 2 by making contour plots of cross-sections did not reveal local maxima, so the effect must be due to the presence of noise. Detection and noise fitting are not expected to be strong enough to have this effect by themselves when the $S/N=100$, but a lack of robustness of the optimization is a likely candidate for causing the trouble. Therefore, run 2 at $S/N=100$ has been repeated with a change in the initial values of the peak-width parameter: the initial step size of σ_m was increased from 3 to 6. The peak-width parameter was selected for variation, because its estimation is the only difference with run 1, where the non-normalities in simple situations are not as strong. The initial step was increased, because this makes the search more global. The changed start-up has many effects, it removes some split-ups, but introduces others, also in the CL experiments. The test results of the experiments are listed in Table 2. The non-normalities that gave rise to the experiments have disappeared all but one. The structure of the standard deviations and the correlation coefficients is unchanged. The means of the WH experiments have been cleaned up. For CL the means are worse at $t_x=4$ and for higher t_x there has been merely a change of the pattern. It is concluded that the method is not robust with respect to the starting values. What ultimately causes this lack of robustness is not elucidated by the experiment, since the optimization cannot be isolated from detection and noise fitting. Yet the simplex, being a local search, may well be a source of error in itself. An interesting experiment therefore, would be to replace it by a very robust global search, e.g. a genetic algorithm [20].

Scatter plots of the parameters in run 1 at $S/N=10$ as in Fig. 10, show deviations from the expected response contours as in Fig. 4, within the difficult areas for both white and correlated noise. In the simple situations all points are situated within the 0.03 contour, whereas in the difficult situations, the shapes of the clouds do not conform with any of the contours, though the points are also roughly within that 0.03 contour. The clouds are not in one piece, but tend to loose points to a region below. It is concluded that the approach of cutting through the noiseless response in order to obtain the distributions of the parameters fails in detail in the difficult situations, though it does have a crude predictive power.

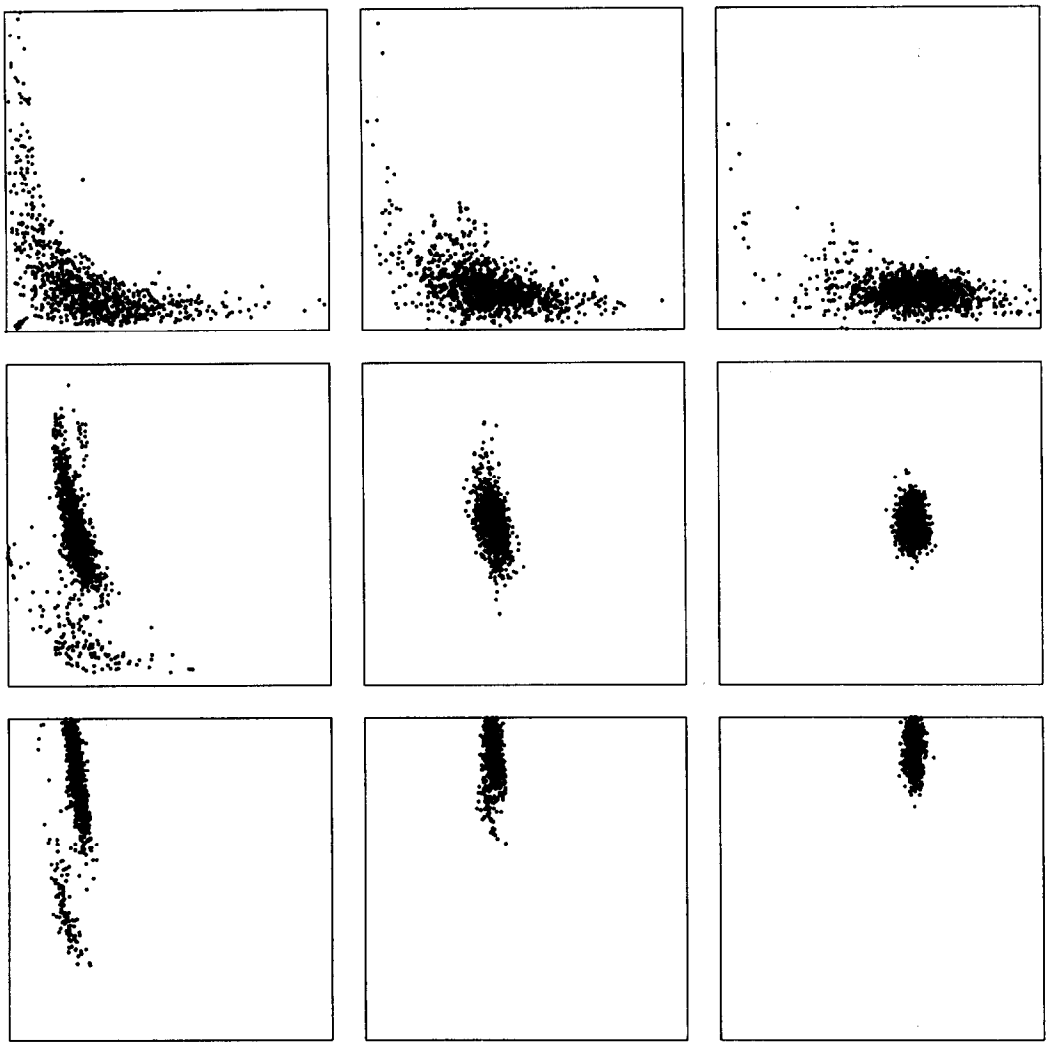

Fig. 10. Parameter estimations in WH run 1 at $S/N=10$. X axis: $t_m[0; 20]$; Y axis: $A_m[0; 1]$. Graphs ordered as in Fig. 4.

Projection of the scatter plots of the parameters (Fig. 10) onto the positions of the maxima in the absence of noise (Fig. 4), leads to the expectation that the MFX would end up with a dozen different positions in the difficult situations. In the simple situations it is expected to find just one position, being the position of the maximum in the absence of noise, further referred to as the true position. The experimental results prove differently. The run 1 WH experiments all end up with three different positions at $S/N=10$, being the true position and the points on either side of it. The true position is the most frequent in all experiments. In the CL equivalent the number of positions varies from one to three. The true position is always one of them. At $S/N=100$ the MFX always finds the true position in run 1. Position estimates that do differ from the true value correspond with A_m-t_m pairs outside the theoretical contours as described in the previous paragraph. The picture that is formed on the basis of these observations is that noise dominates the detection directly, and the optimization indirectly. The detector finds a maximum that consists of a noise maximum close to the top of the noiseless output. Once the procedure has got hold of such a maximum it will be difficult to let it go. The noise contribution rides on top of the noiseless response that is being optimized. Within the bounds that have been set to the parameters in the experiments, and for the noise levels that have been investigated, noise fit-

ting is not expected to be strong enough to remove the noise contribution that pins down the detector. In this view it is remarkable that the detector finds the true optimum so often in the difficult situations.

The question that remains is whether the occurrence of deviations in the difficult situations is due to the double-peak MFX, or inherent to the specific combination of data and a priori information. The expectation is this, the entire set of experiments could be repeated with non-linear regression (NLIN) instead of the MFX. In that case, NLIN should be made capable of dealing with noise that is not white. Although such procedures are described in literature [14,15], they have not found application in chemistry. The gain of using a correct noise model should be evaluated for the double-peak MFX (and NLIN) as it has been done for the single-peak MFX [16].

NLIN and the MFX are related, but differ in details. The response being optimized by NLIN is the sum of squares of the residuals. This is the square of the distance between the data and the orthogonal projection of the data onto the current realization of the model. In MFX, the response being optimized is the length of that orthogonal projection. In other words, the responses of NLIN and MFX are orthogonal. The usual optimization technique in NLIN is the Marquardt–Levenberg algorithm, which is a steepest descent far from the optimum, changing into a linear approximation, i.e. assuming a parabolical shape of the response surface, close to the optimum. In NLIN, the position of the peaks is a parameter just as the other peak parameters, whereas in the MFX, the position is estimated by a separate detector. Probably the most important difference is that the MFX allows the optimization to be bounded. It may be expected that these bounds provide greater robustness and a higher speed of convergence.

8. Conclusions

Under most conditions, the proposed procedure works well for two overlapping Gaussians of equal width in the presence of first-order band-limited noise, the quality of the results being on or little below the theoretical optimum. When the peaks differ less than twice the peak width in position, i.e. the resolution $t_x/\sigma_x \leq 2$, or when the height of one of the peaks is less than 10% of the height of the other, the experimental results may no longer follow a normal distribution and their quality is lower. [21]

Appendix: Output of a double-peak normalized MF for two Gaussians and first-order band-limited noise

The PSD of first-order band-limited noise, normalized to have unit total power is:

$$S(f) = \frac{2\tau}{1+4\pi^2\tau^2 f^2} \tag{16}$$

where τ is the time constant of the noise. The Gaussian peaks of equal width can be defined relative to a central position:

$$x_i(t) = \exp\left[-\frac{1}{2}\left(\frac{2t+kt_x}{2\,\sigma_x}\right)^2\right] \tag{17}$$

where $k=1$ for $i=1$ and $k=-1$ for $i=2$, t_x is the position difference and σ_x the standard deviation, the width of the peaks. The models $m_1(t)$ and $m_2(t)$ are defined analogously, with subscripts m instead of x. Using these definitions and assuming $A_1=1$, Eq. 4 is solved to give:

$$y_n(t) = \frac{\sigma_x\sigma_m\sqrt{\pi}}{\sigma_{\text{out}}\tau\sqrt{s}}\sum_{i=1}^{4}\left[a_i\left(1+\frac{\tau^2}{s}-\frac{\tau^2 b_i^2}{s^2}\right)\exp\left(-\frac{b_i^2}{4s}\right)\right] \tag{18}$$

with the constants

$$\begin{aligned} s &= 2(\sigma_x^2+\sigma_m^2) \\ a_1 &= 1; \quad b_1 = 2t - t_m + t_x \\ a_2 &= A_2; \quad b_2 = 2t - t_m - t_x \\ a_3 &= A_m; \quad b_3 = 2t + t_m + t_x \\ a_4 &= A_2 A_m; \quad b_4 = 2t + t_m - t_x \end{aligned} \tag{19}$$

The maximum of this function is found at $t=0$ when the model is correct, i.e. when $\sigma_x=\sigma_m$, $A_2=A_m$ and $t_x=t_m$. When $\sigma_{\text{out}}=1$ in this expression for the maximum, the maximum equals $\sigma_{\text{out}}{}^2$, the square of the normalization factor.

References

[1] B. van den Bogaert, H.F.M. Boelens and H.C. Smit, Evaluation and correction of signal model errors in a matched filter for the quantification of chromatographic data, *Analytica Chimica Acta*, 274 (1993) 71–85.

[2] P.J.H. Scheeren, P. Barna and H.C. Smit, A software package for the evaluation of peak parameters in an analytical signal based on a non-linear regression method, *Analytica Chimica Acta*, 167 (1985) 65–80.

[3] Y. Hayashi, S.C. Rutan, R.S. Helburn and J.M. Pompano, Information-based prediction of the precision and evaluation of the accuracy of the results from an adaptive filter, *Chemometrics and Intelligent Laboratory Systems*, 20 (1993) 163–171.

[4] Y. Hayashi and S.C. Rutan, Accuracy, precision and information of the adaptive Kalman filter in chromatography, *Analytica Chimica Acta*, 271 (1993) 91–100.

[5] P.B. Crilly, The use of a cross-correlation technique to enhance Jansson's deconvolution procedure, *Journal of Chemometrics*, 4 (1990) 291–298.

[6] A.N. Papas, Chromatographic data systems: a critical review, *CRC Critical Reviews in Analytical Chemistry*, 20 (1989) 359–404.

[7] N. Dyson, *Chromatographic Integration Methods*, RSC Chromatography Monographs, Cambridge, UK, 1990.

[8] R. Deutsch, *System Analysis Techniques*, Prentice-Hall, Englewood Cliffs, NJ, 1969.

[9] B.P. Lathi, *Modern Digital and Analog Communication Systems*, Holt, Rinehart and Winston, 1983.

[10] W.B. Davenport and W.L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York, 1958.

[11] M.U.A. Bromba and H. Ziegler, Variable filter for digital smoothing and resolution enhancement of noisy spectra, *Analytical Chemistry*, 56 (1984) 2052–2058.

[12] P.A. Bristow, Towards a chromatographic quantitator, *Journal of Chromatography*, 506 (1990) 265–277.

[13] Ph.K. Hopke and D.L. Massart, Reference data sets for chemometrical methods testing, *Chemometrics and Intelligent Laboratory Systems*, 19 (1993) 35–41.

[14] Y. Bard, *Nonlinear Parameter Estimation*, Academic Press, New York/London, 1974.

[15] N.R. Draper and H. Smith, *Applied Regression Analysis*, Wiley, New York, 1966.

[16] B. van den Bogaert, H.F.M. Boelens and H.C. Smit, Quantification of chromatographic data using a matched filter: robustness towards noise model errors, *Analytica Chimica Acta*, 274 (1993) 87–97.

[17] L.A. Yarbro and S.N. Deming, Selection and preprocessing of factors for simplex optimization, *Analytica Chimica Acta*, 73 (1974) 391–398.

[18] R.B. D'Agostino and M.A. Stephens, *Goodness-of-Fit Techniques*, Marcel Dekker, New York, 1986.

[19] G. Crisponi, F. Cristiani and V. Nurchi, Reliability of the parameters in the resolution of overlapped Gaussian peaks, *Analytica Chimica Acta*, 281 (1993) 197–206.

[20] C.B. Lucasius, A.P. de Weyer, L.M.C Buydens and G. Kateman, CFIT: a genetic algorithm for survival of the fitting, *Chemometrics and Intelligent Laboratory Systems*, 19 (1993) 337–341.