JOURNAL OF CHEMOMETRICS
J. Chemometrics 2001; **15**: 627–646
DOI: 10.1002/cem.654

Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution

R. Tauler*

Department of Analytical Chemistry, University of Barcelona, Diagonal 647, E-08028 Barcelona, Spain

SUMMARY

A method for the calculation of maximum and minimum band boundaries of feasible solutions corresponding to the species profiles estimated by multivariate curve resolution is presented. The method is based on a non-linear constrained optimization of an objective function defined by the ratio of the signal contribution of a particular species to the whole measured signal. Implementation of constraints such as normalization, closure, non-negativity, unimodality and local rank/selectivity during the non-linear optimization and their effect on the calculation of the feasible solutions are studied in detail. Calculation of the band boundaries is shown for different simulated and real two-way data examples of increasing complexity. Copyright © 2001 John Wiley & Sons, Ltd.

KEY WORDS: multivariate curve resolution; non-linear constrained optimization; band boundaries; non-negativity; unimodality; closure; normalization; selectivity; local rank

1. INTRODUCTION

Multivariate curve resolution techniques have been proposed for the recovery of the profiles (spectra, pH profiles, time profiles, elution profiles, etc.) of more than one component in an unresolved and unknown mixture when no prior information is available about the nature and composition of this mixture [1–4]. Complete resolution of a two-way data set without ambiguities is only possible in some favourable cases where selectivity [5] or local rank conditions [6] are present. When these resolution conditions are not present in the system, resolution without ambiguities is not possible even if constraints such as non-negativity, unimodality or closure are applied [7,8]. In these cases, instead of unique profiles, a range or band of feasible profiles fitting equally well the experimental data and fulfilling the physical and chemical constraints of the system has to be considered. Different procedures have been proposed for the estimation of these bands of feasible solutions in the case of mixtures of two and three components [1,2], and for more complex multicomponent mixtures using different optimization strategies [9–11]. Recently, Gemperline [12] has shown that the calculation of the band boundaries of feasible solutions for every species is possible when an objective optimization

* Correspondence to: R. Tauler, Department of Analytical Chemistry; University of Barcelona, Diagonal 647, E-08028 Barcelona, Spain.
E-mail: roma@quimio.qui.ub.es
Contract/grant sponsor: Ministerio de Educación y Cultura de España; Contract/grant number: PB96-0377

Received 20 February 2000
Accepted 13 September 2000

function is defined in terms of the ratio of the signal contribution of that species to the whole signal contribution for the mixture of all species. Based on this proposal and on recent optimization strategies, a procedure is described which attempts the calculation of the band boundaries of the feasible profiles obtained by curve resolution of a two-way data matrix under non-negativity, closure, unimodality and selectivity/local rank constraints. The method needs firstly the estimation of one of the feasible solutions within the range of all possible solutions, for instance using alternating least squares with constraints [5]. Secondly, once this feasible solution is available, a non-linear constrained optimization is initiated looking for the boundaries of the whole set of feasible solutions. The effect of applied constraints, in particular the effect of local rank and selectivity constraints, is especially examined. The results presented in this paper should be compared with those found previously by Gemperline [12].

Three examples are studied in detail. The first example is a simulated spectroscopic data set of an equilibrium system with two components. The second example is a simulated chromatographic coelution system with multiwavelength UV diode array detection and three highly overlapped components. Finally, in a third example the experimental study of Cu(II) complexation by chloride is carried out at two temperatures, 25 and 80 °C, and 5·0 M $Na^{+}(ClO_4^{-}, Cl^{-})$. In this case the study involved firstly the resolution of the unknown species and secondly the estimation of the band boundaries of the resolved species profiles.

2. METHOD

2.1. *Multivariate curve resolution, rotational ambiguities and feasible bands*

In multivariate curve resolution a bilinear decomposition of the experimental data matrix is performed using the model equation

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathrm{T}} + \mathbf{E} \tag{1}$$

where the dimensions of the matrices are $\mathbf{D}(NR,NC)$, $\mathbf{C}(NR,N)$, $\mathbf{S}^{\mathrm{T}}(N,NC)$ and $\mathbf{E}(NR,NC)$; N is the number of considered components (chemical species contributing to the signal); NR is the number of rows (for spectroscopic data, NR is the number of spectra) in the data matrix $\mathbf{D}$; and NC is the number of columns (for spectroscopic data, NC is the number of wavelengths) in the data matrix $\mathbf{D}$. $\mathbf{C}$ is the matrix describing how the contributions of the N species change in the NR different rows of the data matrix (concentration profiles). $\mathbf{S}^{\mathrm{T}}$ is the matrix describing how the responses of these N species change in the NC columns of the data matrix (pure spectral profiles). $\mathbf{E}$ is the residual matrix with the data variance unexplained by $\mathbf{C}\mathbf{S}^{\mathrm{T}}$.

The problem to solve in multivariate curve resolution may be stated mathematically in the following way. Given the data matrix $\mathbf{D}$: (1) find N, the number of chemical components or species causing the observed data variance $\mathbf{D}$; (2) find the (row) concentration profiles of these species, i.e. matrix $\mathbf{C}$; and (3) find the (column) pure response or spectral profiles of these species, i.e. matrix $\mathbf{S}^{\mathrm{T}}$. However, the solution of Equation (1) for $\mathbf{C}$ and $\mathbf{S}^{\mathrm{T}}$ is ambiguous if no additional information is available; or, in other words, there is rotational and scale freedom in the solutions of Equation (1). This problem is often called the factor analysis ambiguity problem [1–6].

Rotational and intensity ambiguities can be easily shown from the equation

$$\mathbf{D} = \mathbf{C}_{\mathrm{old}}\mathbf{S}_{\mathrm{old}}^{\mathrm{T}} = (\mathbf{C}_{\mathrm{old}}\mathbf{T})(\mathbf{T}^{1}\mathbf{S}_{\mathrm{old}}^{\mathrm{T}}) = \mathbf{C}_{\mathrm{new}}\mathbf{S}_{\mathrm{new}}^{\mathrm{T}} \tag{2}$$

It is clear that if no constraints are considered, there is an infinite number of possible solutions of Equation (2) for any non-singular matrix $\mathbf{T}$, i.e. $\mathbf{T}$ should be invertible; but this is accomplished by an

infinite number of matrices **T**. Interestingly, however, **T** is of reduced dimensions **T**(*N*,*N*), much lower than the dimensions of **C**, $\mathbf{S}^{\mathrm{T}}$ or **D**.

It is usually possible to reduce considerably this infinite number of possible solutions by means of constraints derived from the physical nature and previous knowledge of the problem under study. For instance, only positive values for the concentrations of the chemical components in the mixture have physical meaning; in many spectroscopies, only positive values are allowed in the spectra; concentration profiles are often unimodal; and closure or mass balance equations should be fulfilled for reaction-based systems. However, as already pointed out by various authors [1,5,6], the most important constraints to limit drastically the number of possible solutions are the selectivity and local rank constraints. Under these constraints, unique solutions can be obtained both if the data have such a favourable structure and if a suitable method is used to detect and use such information. Methods to detect and use local rank and selectivity structure for curve resolution have been proposed, based mostly on evolving factor analysis [13,14]. For instance, if selectivity [5] or resolution local rank conditions [6] are present in one of the two orders for every species of the unresolved mixture, the correct recovery of the concentration and/or spectral profiles for all the different species is possible.

For a particular species profile the set of feasible solutions under constraints defines a range or band of feasible solutions, and this band may be considered delimited by a maximum band boundary and a minimum band boundary. These boundaries may be defined in different ways. One of the simplest ways for their definition is that proposed recently by Gemperline [12] in terms of maximum and minimum signal contributions of each species to the whole measured signal (see below). These boundaries will be related to specific rotation matrices **T** for each species k, which will be called $\mathbf{T}_{\max,k}$ and $\mathbf{T}_{\min,k}$. Consider a particular set of solutions of Equation (1) fulfilling the constraints defined by the problem, $\mathbf{C}_{\mathrm{inic}}$ and $\mathbf{S}^{\mathrm{T}}_{\mathrm{inic}}$. The maximum band boundaries $\mathbf{C}_{\max,k}$ (for the concentration profiles) and $\mathbf{S}^{\mathrm{T}}_{\max,k}$ (for the spectral profiles) and the minimum band boundaries $\mathbf{C}_{\min,k}$ (for the concentration profiles) and $\mathbf{S}^{\mathrm{T}}_{\min,k}$ (for the spectral profiles) may be defined by the equation

$$\mathbf{D} = \mathbf{C}_{\mathrm{inic}}\mathbf{S}^{\mathrm{T}}_{\mathrm{inic}} = \mathbf{C}_{\mathrm{inic}}\mathbf{T}_{\min}\mathbf{T}^{-1}_{\min}\mathbf{S}^{\mathrm{T}}_{\mathrm{inic}} = \mathbf{C}_{\min,k}\mathbf{S}^{\mathrm{T}}_{\min,k} = \mathbf{C}_{\mathrm{inic}}\mathbf{T}_{\max,k}\mathbf{T}^{-1}_{\max,k}\mathbf{S}^{\mathrm{T}}_{\mathrm{inic}} = \mathbf{C}_{\max,k}\mathbf{S}^{\mathrm{T}}_{\max,k} \qquad (3)$$

The goal of the method described here is to find a way to calculate these values of $\mathbf{T}_{\max,k}$ and $\mathbf{T}_{\min,k}$ which define the maximum and minimum band boundaries of feasible solutions for profiles of species k under a set of constraints defined for a particular data set.

2.2. *Description of the optimization problem and the optimization method*

In this work the calculation of the maximum and minimum band boundaries of feasible solutions defined by Equation (3) is considered in the frame of a constrained non-linear optimization problem (NCP) [15,16]. This optimization problem is described mathematically using the equation

$$\underset{\mathbf{X}}{\text{minimize}}\ f(\mathbf{X}) \quad \text{subject to} \quad \mathbf{g}_{\mathrm{e}}(\mathbf{X}) = 0 \text{ and } \mathbf{g}_{\mathrm{i}}(\mathbf{X}) \leq 0 \qquad (4)$$

where **X** is a matrix of variables, $f(\mathbf{X})$ is a non-linear scalar function of **X**, $\mathbf{g}_{\mathrm{e}}(\mathbf{X})$ is a vector of equality constraints and $\mathbf{g}_{\mathrm{i}}(\mathbf{X})$ is a vector of inequality constraints. In the general framework, both $\mathbf{g}_{\mathrm{e}}(\mathbf{X})$ and $\mathbf{g}_{\mathrm{i}}(\mathbf{X})$ will be considered non-linear functions of the matrix of variables **X**.

2.2.1. What are the **X** *variables of the problem?* The **X** matrix of variables to optimize will be the elements of the rotation matrix **T** which give the band boundaries of feasible solutions of Equation (1). For each set of species profiles (concentration and spectral), two matrices should be obtained, the $\mathbf{T}_{\max,k}$ matrix which defines the maximum band boundary for this species profile and

the $\mathbf{T}_{\min,k}$ matrix which defines the minimum band boundary for this species profile (see Equation (3)). Therefore two optimizations each giving a rotation matrix $\mathbf{T}$ should be performed for each species k of the unresolved mixture. The number of optimizations to be carried out increases with the complexity of the problem. As pointed out previously, the number of variables in matrices $\mathbf{T}$, hence the complexity of the optimization, increases also with the number of species of the problem. For $k = 1,\ldots,N$ species the number of variables in the $\mathbf{T}$ matrix is equal to $N \times N$ and the number of optimizations will be $2N$.

2.2.2. What is the objective function to optimize? The definition of the objective function is a critical step in the implementation of the optimization algorithm. It should be a scalar function of the variables, the optimization of which allows the calculation of the maximum and minimum band boundaries of the feasible solutions. For a good performance of the optimization algorithm the optimization function should be scaled, for instance between zero and one. Two functions are proposed in this paper to define the function to be optimized and consequently to define also the physical meaning of the band boundaries obtained in such an optimization. The first function is the same as the one proposed previously by Gemperline [12] and gives the ratio between the integrated signal contribution of a particular species and the integrated whole signal of all the species present in the unresolved mixture. The equation for its evaluation for each species is

$$f_k(\mathbf{T}) = \frac{\sum_i c_{i,k} s_{k,i}^{\mathrm{T}}}{\sum_k \sum_i c_{i,k} s_{k,i}^{\mathrm{T}}}, \quad k = 1, \ldots, N \tag{5}$$

The second function is defined similarly using the Frobenius norm of the signal contribution of a particular species with respect to the Frobenius norm of the whole signal for all the considered species:

$$f_k(\mathbf{T}) = \frac{\|\mathbf{c}_k \mathbf{s}_k^{\mathrm{T}}\|}{\|\mathbf{C}\mathbf{S}^{\mathrm{T}}\|}, \quad k = 1, \ldots, N \tag{6}$$

Both functions will give identical results for positive concentration and spectral profiles. For some spectroscopic signals giving negative values, Equation (6) should be preferred since it is valid also for negative profiles.

The optimization of this objective function, either maximized or minimized, will give respectively the maximum and the minimum boundary. For a particular species these boundaries will define the feasible profiles (concentration and spectral) fulfilling the constraints of the problem and giving a maximum and a minimum signal contribution.

The definition of both objective functions requires initial values of the variables in $\mathbf{T}$ and also requires initial values for the species profiles, $\mathbf{C}_{\mathrm{inic}}$ and $\mathbf{S}_{\mathrm{inic}}^{\mathrm{T}}$ (see Section 2·2·5). When the optimization is implemented as a minimization of the objective function as in Equation (4), Equations (5) and (6) are appropriate to find the minimum band boundaries, whereas if the optimization is implemented as a maximization of the objective function, the functions given in Equations (5) and (6) should be changed in sign.

2.2.3. What are the constraints of the problem? Are they equality or inequality constraints? As previously stated, the solutions of Equation (2) are unbounded if no additional constraints are considered; there is an infinite number of possible solutions and it is not possible to find the band

boundaries of the feasible solutions. The optimization of the functions defined by Equations (5) and (6) without any additional constraints will not be possible under these circumstances. Fortunately, in most curve resolution problems it is possible to constrain appropriately the set of feasible solutions and allow the calculation of at least one of them using an appropriate algorithm [5,7,8]. In this work it is shown that, using these constraints, it is also possible to find the band boundaries of feasible solutions. The following constraints are considered.

2.2.3.1. Normalization and closure constraints. Intensity ambiguities arise because for any scalar m and species k profiles $\mathbf{c}_k$ and $\mathbf{s}_k^{\mathrm{T}}$,

$$\mathbf{c}_k \; m \; m^{-1} \; \mathbf{s}_k^{\mathrm{T}} = \mathbf{c}_k \mathbf{s}_k^{\mathrm{T}} \tag{7}$$

For a particular species k the concentration profile $\mathbf{c}_k$ can be arbitrarily increased in an unlimited way by multiplying it by an arbitrary scalar number m if at the same time its spectrum $\mathbf{s}_k^{\mathrm{T}}$ is decreased by the same amount by dividing it by the same number m. It is possible to limit the size of $\mathbf{c}_k$ or $\mathbf{s}_k^{\mathrm{T}}$ profiles using appropriate normalization and closure constraints. A spectral normalization constraint may be implemented using the norm of the species spectra, for instance forcing it to be equal to one, i.e. $\|(\mathbf{s}_k(\mathbf{T}))\| = 1$. This is an equality constraint and is implemented in Equation (4) by the equality constraint function

$$g_{\mathrm{norm},k}(\mathbf{T}) = 1 - \|(\mathbf{s}_k(\mathbf{T}))\| = 0, \quad k = 1, \ldots, N \tag{8}$$

If N species are present, this will give N equality constraints. Alternatively, other possible normalization constraints can be applied by constraining the signal height or maximum intensity of a profile to be equal to a constant value. Obviously this will also give N equality constraints implemented in a very similar way to the one given previously in Equation (8).

A completely different constraint is the closure constraint, which is usually implemented on the rows of matrix $\mathbf{C}$. By this constraint the sum of the elements of each row of matrix $\mathbf{C}$ is equal to a known constant. This is the case for instance in reaction-based systems, where a mass balance equation is obeyed by the concentration profiles of the species present in the system, i.e. $\sum_{k=1}^{N} c_{i,k}(\mathbf{T}) = \mathrm{TOT}_i$. This closure constraint is implemented by the equation

$$g_{\mathrm{clos},i}(\mathbf{T}) = \mathrm{TOT}_i - \sum_{k=1}^{N} c_{i,k}(\mathbf{T}) = 0, \quad i = 1, \ldots, NR \tag{9}$$

where TOT_i is the known total concentration of the species in the considered spectra (row of matrix $\mathbf{D}$). This will give NR equality constraints, one for each row of matrix $\mathbf{C}$. The number of equality constraints increases considerably with this constraint, decreasing the degrees of freedom of the optimization.

Either a normalization or a closure constraint should be applied during the optimization of the functions defined by Equations (5) and (6) to avoid scaling problems in the variables during the optimization. If neither or both constraints are applied, the optimization will not work because of the absence of scale boundaries. On the other hand, only one of the two constraints, spectral normalization or concentration closure, is usually applied, otherwise the system becomes overly constrained. Closure constraint can only be applied if this is the case for the data, otherwise the solutions of Equation (5) will not be correct. Conversely, when there is no closure in the concentrations, normalization of the species spectra $\mathbf{S}^{\mathrm{T}}$ is highly recommended to achieve a

successful optimization and recover the shape of the species profiles, even if the true solutions are obviously not fulfilling this normalization constraint.

2.2.3.2. Non-negativity constraints. Non-negativity constraints are probably the most commonly used constraints in curve resolution since the initial work of Lawton and Sylvestre [1]. Physical concentrations can be only positive or zero ($\mathbf{C} \geq 0$), and in many spectroscopies, spectral values can also be only positive or zero ($\mathbf{S}^{\mathrm{T}} \geq 0$). Both constraints can be implemented as inequality constraints:

$$g_{\mathrm{cneg},i,k}(\mathbf{T}) = -c_{i,k}(\mathbf{T}) \leq 0 \tag{10}$$

$$g_{\mathrm{sneg},j,k}(\mathbf{T}) = -s_{j,k}(\mathbf{T}) \leq 0 \tag{11}$$

$$i = 1, \ldots, NR, \quad j = 1, \ldots, NC, \quad k = 1, \ldots, N$$

According to Equation (10) for non-negative concentrations, for each species there will be NR inequality constraints, and for N species this will give a total of $N \times NR$ inequality constraints. According to Equation (11) for non-negative spectra, for each species there will be NC inequality constraints, and for N species this will give a total of $N \times NC$ inequality constraints. These two constraints together with the normalization/closure constraints are the most frequently found and easily applied in multivariate curve resolution problems. The number of inequality constraints given by both equations is very high and limits considerably the size of the bands of feasible solutions. However, they do not constrain the problem sufficiently to give unique solutions [1,5].

2.2.3.3. Selectivity, local rank and known value constraints. Two types of constraints are distinguished here. The first type of constraints frequently used in curve resolution, allowing in many cases the partial or even total elimination of rotational ambiguities, comprises the so-called 'selectivity' and 'local rank' constraints [5,6]. In all circumstances, selectivity and local rank constraints have a tremendous effect of narrowing considerably the bands of feasible solutions, eventually collapsing them in a unique solution. Selectivity and local rank constraints refer to the fact that in certain windows or regions of the data matrix $\mathbf{D}$ a particular species is known to exist while other species are known not to exist. For instance, in reaction-based systems it is a common situation that some of the species are not present at the beginning or at the end of the reaction process. Also, for spectroscopic signals it can happen that a particular species does not absorb in a particular spectral range. In these cases it is extremely helpful to use this information as a constraint defining for each species the concentration window where it exists and the concentration window where it does not (or where it does exist only at very negligible concentrations), and the spectral window where it contributes to the measured signal and the spectral window where it does not. As has been demonstrated previously [5,6], when the resolution conditions are appropriate, the use of the local rank information allows the recovery of the true species profiles without ambiguities. In practice, however, this is not completely achieved in the general case, as in the case of embedded peaks in chromatography [6], or when the species profiles are extremely overlapped as in kinetic reaction-based systems. The two terms selectivity and local rank refer to the same type of constraints. Local rank constraints identify some of the regions where different species do not exist, or better, regions where different species are practically non-existent. Selectivity constraints are only the extreme case of local rank conditions, when only one species exists or contributes to the signal in a particular concentration or spectral window of the data matrix. Methods to detect and use local rank and selectivity in the data have been proposed, such as those derived from evolving factor analysis

[13,14]. Difficulties may arise because of noise, because of the difficulty of detection of low concentrations and low contributions to the signal, because of the difficulty in defining precisely the beginning and the end of the concentration and spectral windows, or because of rank deficiency problems [17,18]. Apart from these sometimes important difficulties, the use of the local rank information provided by these methods is extremely helpful for resolution of the system. Gemperline [12] examined also the effect of measurement noise in the calculation of band boundary uncertainties.

Our experience indicates that selectivity/local rank constraints are better implemented in iterative algorithms as inequality constraints than as equality constraints. Equality constraints are sometimes too strong during the optimization and, in fact, in many circumstances we cannot be completely sure that the values are exactly equal to zero or to another value in a certain data window. Usually what we really know is that a particular species does not exist at appreciable concentrations or that it does not contribute to the signal in an appreciable way. However, this does not mean that the concentration or the signal should be exactly zero in that window, but better, below a threshold value. For this reason it is usually better just to select a threshold value and to force the values of the profile in the window to be equal to or lower than this threshold value. If that threshold value is equal to ϵ, then the selectivity/local rank constraints can be described as

$$g_{\mathrm{csel},i,k}(\mathbf{T}) = c_{\mathrm{sel},i,k}(\mathbf{T}) - \varepsilon \leq 0 \tag{12}$$

$$g_{\mathrm{ssel},j,k}(\mathbf{T}) = s_{\mathrm{sel},j,k}(\mathbf{T}) - \varepsilon \leq 0 \tag{13}$$

$$i = 1, \ldots, ncsel_k, \quad j = 1, \ldots, nssel_k, \quad k = 1, \ldots, N$$

The number of inequality constraints will depend on the number of values included in the selected windows, $ncsel_k$ and $nssel_k$, for each species k considered.

A second type of possible constraints would be those values in $\mathbf{C}$ or $\mathbf{S}^{\mathrm{T}}$ which were exactly known. It may happen for instance that the spectrum of one of the species is known in advance. In that case such information can be used to simplify the optimization problem through equality constraints such as

$$g_{\mathrm{cknown},i,k}(\mathbf{T}) = c_{i,k}(\mathbf{T}) - c_{\mathrm{known},i,k} = 0 \tag{14}$$

$$g_{\mathrm{sknown},j,k}(\mathbf{T}) = s_{i,k}(\mathbf{T}) - s_{\mathrm{known},j,k} = 0 \tag{15}$$

$$i = 1, \ldots, NR, \quad j = 1, \ldots, NC; \quad k = 1, \ldots, N$$

If the whole concentration profile of one species is known, this gives *NR* equality constraints. If the whole spectrum of one species is known, this gives *NC* equality constraints

2.2.3.4. Unimodality constraints. In a wide variety of situations, concentration profiles have a unimodal shape, i.e. they only have a maximum. This is the case for instance for chromatographic elution profiles and also for concentration profiles in reaction-based systems. In contrast, spectral profiles are usually not unimodal in the general case. Some other types of instrumental signals may also be unimodal, as in electrochemistry. Whenever the shape of the profiles is unimodal, an additional constraint which may be useful is unimodality. In that case, what is intended is to avoid the formation of secondary maxima. The way used to implement unimodality constraints has been algorithmic. First the highest maximum is detected and then all the departures from the unimodal condition are constrained, discarding left and right maxima: at the left of the peak maximum,

$$g_{\text{unimod},i,k} = c_{i-1,k}(\mathbf{T}) - c_{i,k}(\mathbf{T}) \leq 0 \tag{16}$$

and at the right of the peak maximum,

$$g_{\text{unimod},i,k} = c_{i+1,k}(\mathbf{T}) - c_{i,k}(\mathbf{T}) \leq 0 \tag{17}$$

This constraint can include also a small local departure from unimodality tolerance because of noise, which gives the amount of local increasing concentration tolerated at the right or at the left of the maximum concentration. See References [8,19] for more detailed descriptions of the unimodality constraint. The number of inequality constraints from Equations (16) and (17) will change during the optimization depending on the shape adopted by the concentration profile $\mathbf{c}(\mathbf{T})$.

2.2.3.5. Set-up of the vector of constraints. Two vectors of constraints are finally built up, one for the equality constraints,

$$\mathbf{g}_{\text{eq}} = \mathbf{g}_{\text{norm}} + \mathbf{g}_{\text{clos}} + \mathbf{g}_{\text{known}} = \mathbf{0} \tag{18}$$

and the other for the inequality constraints,

$$\mathbf{g}_{\text{ineq}} = \mathbf{g}_{\text{neg}} + \mathbf{g}_{\text{sel}} + \mathbf{g}_{\text{unimod}} \leq \mathbf{0} \tag{19}$$

2.2.4. Initial values of the variables. The constrained non-linear optimization described in Equation (5) needs initial estimates of the variables $\mathbf{T}$ and of the profiles $\mathbf{C}_{\text{inic}}$ and $\mathbf{S}^{\text{T}}_{\text{inic}}$. An obvious way to facilitate the optimization is that both $\mathbf{C}_{\text{inic}}$ and $\mathbf{S}^{\text{T}}_{\text{inic}}$ be feasible, i.e. that they already fulfil the constraints of the problem. In this way, if matrix $\mathbf{T}(N,N)$ is selected to be equal to the identity matrix, the starting values of the variables will give feasible solutions of the problem, and the initial optimization steps are facilitated. For simulated data, $\mathbf{C}_{\text{inic}}$ and $\mathbf{S}^{\text{T}}_{\text{inic}}$ may be chosen as those which are used to built up the data matrix $\mathbf{D}$ (see Section 3). Obviously, these profiles already obey the constraints and fit the data appropriately. The goal of the optimization is then to find the maximum and minimum boundaries of all the alternative solutions which fit the data matrix $\mathbf{D}$ equally well and also fulfil the postulated constraints. In order to check that the feasible solutions do not depend on the initial estimates, and also, more importantly, to have starting profiles for real experimental data, the profiles obtained using an alternating least squares procedure (MCR-ALS [5,7,8]), $\mathbf{C}_{\text{ALS}}$ and $\mathbf{S}^{\text{T}}_{\text{ALS}}$ are also used as initial estimates. These profiles fulfil the constraints of the problem and fit the data optimally. In fact, the MCR-ALS procedure [5] was developed with this idea in mind, to find an optimal solution of the multivariate curve resolution problem which optimally fits the data and also fulfils the constraints of the problem. The MCR-ALS method has been described in previous works and applied successfully to different data types [5,7,8,18], so it is not described further in the present work.

2.2.5. Optimization method and initial optimization parameters. The constrained non-linear optimization problem described by Equation (5) is solved using a sequential quadratic programming (SQP) algorithm with a mixed quadratic and cubic line search method. The background of these methods is described in detail in the book by Gill *et al.* [15] and implemented in the MATLAB [16] Optimization Toolbox [20], function constr.m (version 1.5.2, 1990) and, more recently, function fmincon.m (version 2.0, 1998). SQP methods are recognized as the state of the art in non-linear programming optimization methods [21]. At each major iteration

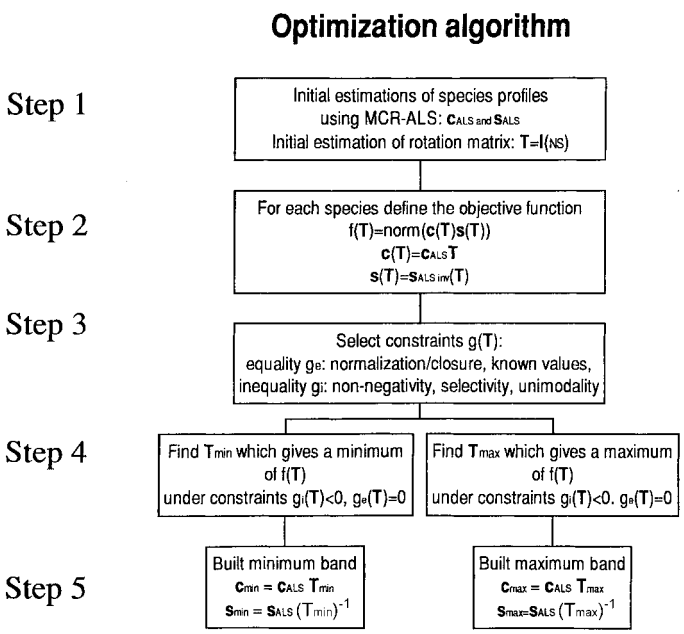

Figure 1. Steps of the optimization algorithm.

of the quadratic programming optimization procedure an approximation is made of the Hessian of the Lagrangian function using a quasi-Newton updating method. The Lagrangian function is obtained from the definition of the necessary and sufficient optimality conditions in constrained optimizations derived from Kuhn–Tucker equations [15,20,21]. A quadratic programming (QP) subproblem is generated iteratively whose solution is used to form a search direction for a line search procedure. More details of the algorithms used in the SQP optimizations are given in References [15,20–22].

The particular implementation flow chart of the algorithm developed in the present work to find the band boundaries of feasible solutions is shown in Figure 1.

In step 1, feasible solutions $\mathbf{C}_{\text{inic}}$ and $\mathbf{S}^{\text{T}}_{\text{inic}}$ are initially postulated or found using the MCR-ALS optimization algorithm [5] under a set of preselected constraints. The rotation matrix $\mathbf{T}(N,N)$ is initially set equal to the identity matrix of the same dimensions. In step 2 the objective function is defined. Two objective functions should be set up for each species profile (concentration and spectral), one to evaluate the maximum band boundary and the other to evaluate the minimum band boundary. In step 3 the set of equality and inequality constraints is formulated. These constraints should be in agreement with the initial feasible solutions postulated in step 1. Step 4 is the core of the algorithm, where the maximum and minimum band boundaries of the feasible solutions are obtained by minimization of the objective functions defined in step 2 under the constraints defined in step 3. This non-linear optimization is carried out using the appropriate external subroutines. From the output of these optimization subroutines the optimal values of the rotation matrix, $\mathbf{T}_{\max}$ and $\mathbf{T}_{\min}$, are obtained for the maximum and minimum band boundaries of the species profiles (concentration and spectral). In step 5 these band boundaries are finally evaluated.

The correct implementation of the previously described optimization method requires that several parameters be considered, such as the termination tolerance for the variables in the $\mathbf{T}$ matrix, the termination tolerance for the objective function $f(\mathbf{T})$, the termination criterion of constraint violation,

the maximum number of iterations, the minimum and maximum changes in variables for difference gradients and the initial step length. From our experience, usually for a well-defined problem with an appropriate scaled objective function, default values provided by the MATLAB routine are adequate (see MATLAB Optimization Toolbox). In some cases, however, if the optimization does not run properly, changes in some of the default values may be attempted. This is largely problem-dependent and no general rules can be given at the moment.

3. DATA

Three application examples are given showing typical situations and constraints found in curve resolution problems.

Example 1. Mixture of a two-component equilibrium system

Figure 2 gives the first data example where the proposed method for the calculation of the band boundaries of feasible solutions was applied. In this example the data come from a hypothetical chemical equilibrium monitored spectrophotometrically, with two species contributing to the measured signal with the concentration ($\mathbf{C}_1$) and spectral ($\mathbf{S}_1^{\mathrm{T}}$) profiles shown in Figure 2. The data matrix was obtained simply from $\mathbf{D}_1 = \mathbf{C}_1\mathbf{S}_1^{\mathrm{T}} + \mathbf{N}$, where the noise $\mathbf{N}$ is Gaussian-distributed with a

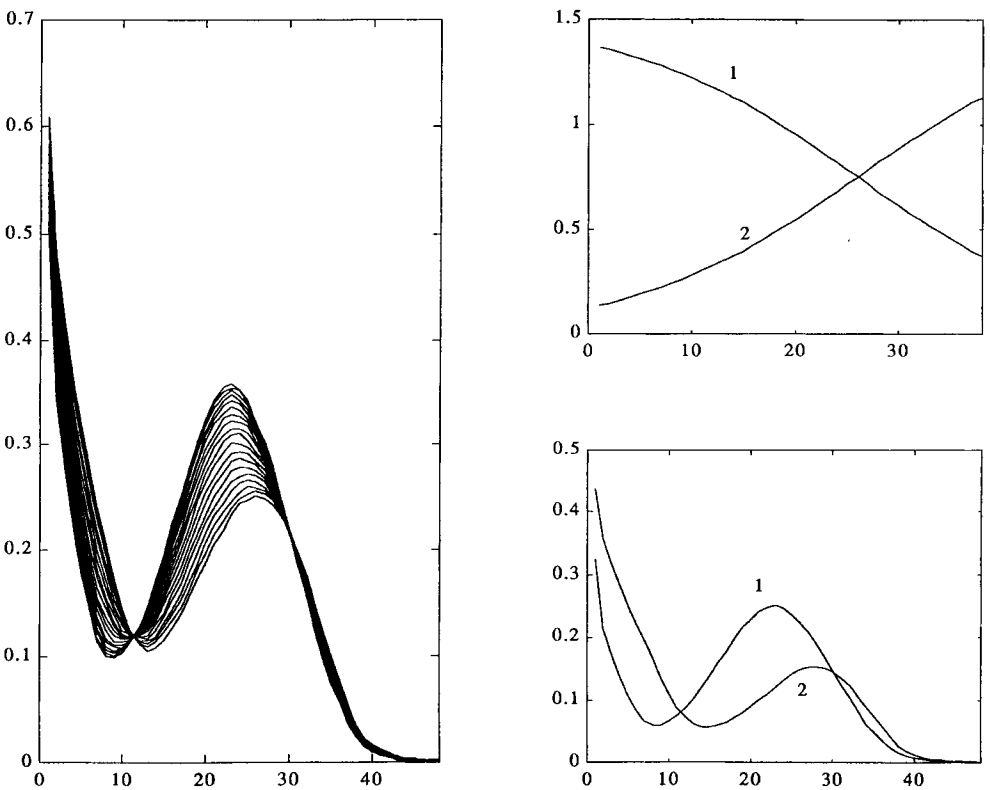

Figure 2. Data example 1: simulation of a two-species (1 and 2) equilibrium system. Left: plot of 38 simulated mixture spectra with 48 wavelengths (data matrix $\mathbf{D}_1$). Top right: plot of species distribution (concentration profiles, $\mathbf{C}_1$). Bottom right: plot of species spectra ($\mathbf{S}_1^{\mathrm{T}}$).

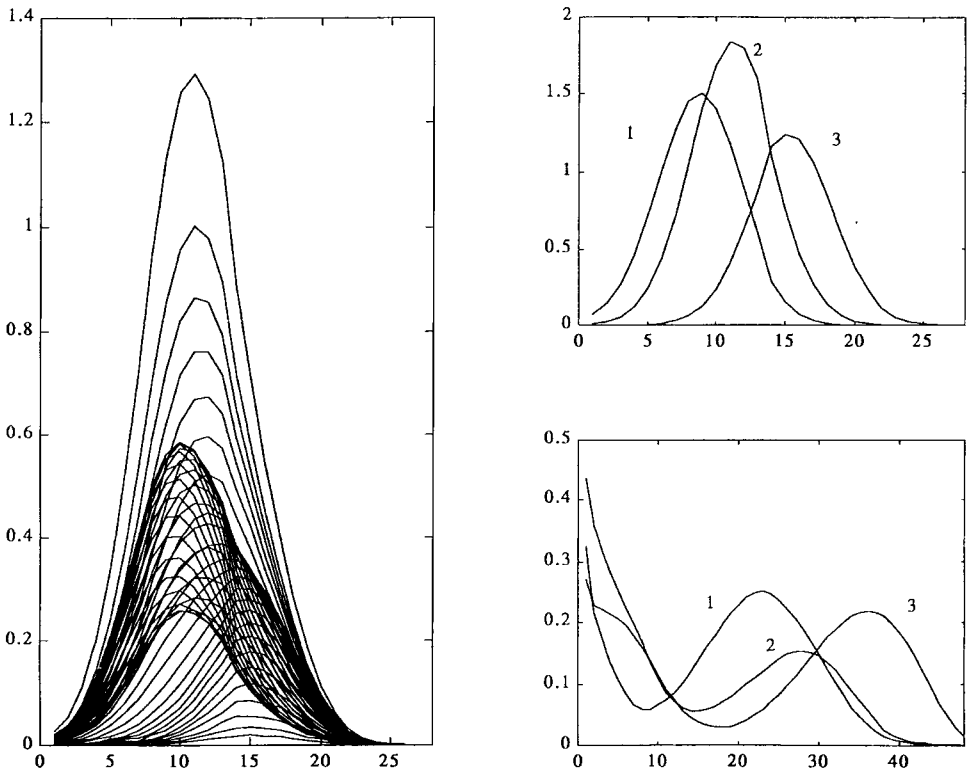

Figure 3. Data example 2: simulation of a three-component chromatographic coelution system (data matrix $\mathbf{D}_2$). Left: plot of 26 simulated chromatographic profiles at 48 different wavelengths. Top right: plot of elution profiles ($\mathbf{C}_2$). Bottom right: plot of spectra of each component ($\mathbf{S}_2^{\mathrm{T}}$).

standard deviation of 0·001 units. The dimensions of this matrix were 38 rows (spectra) × 48 columns (wavelengths). The profiles used for the data generation had no selectivity either in the spectra or in the concentrations. The system is closed, since the sum of the concentrations of the two profiles is always equal to a constant value (mass balance equation). Pure species spectra were also normalized to unit length. In this example the band boundaries of the feasible solutions were calculated considering the following situations: (a) non-negativity and spectral normalization constraints; (b) non-negativity and concentration closure constraints; and (c) selectivity constraints. Initial values of the feasible profiles used for the calculation of the bands of feasible solutions were obtained from the profiles used in the data simulation and/or from the profiles obtained using MCR-ALS.

Example 2. Mixture of three coelution chromatographic components (chromatographic coelution)

Figure 3 shows the simulated data matrix $\mathbf{D}_2$ obtained from the concentration/elution profiles $\mathbf{C}_2$ and spectral profiles $\mathbf{S}_2^{\mathrm{T}}$ corresponding to a mixture of three coelution compounds analysed using chromatography with diode array UV detection. The noise $\mathbf{N}$ was also assumed to be Gaussian-distributed with a standard deviation of 0·001 units. The dimensions of this matrix were 26 rows (spectra) × 48 columns (wavelengths). The chromatographic elution profiles were poorly resolved at all the wavelengths; only for the third species is there a small region at the end of the peak where this

species is the only one present at low concentrations. This third species is also the species which has a less overlapped spectrum, with a selectivity region at one of the ends of the absorption band (Figure 3). From this preliminary analysis it was clear that species 3 was the species which might be resolved with fewer ambiguities and that the resolution without ambiguities of spectral and concentration profiles of species 1 and 2 would be more difficult. For this example the following situations were studied: (a) non-negativity and spectral normalization constraints; (b) unimodality constraint; (c) local rank/selectivity constraints; and (d) initial estimates (from simulation and from EFA-ALS).

Example 3 (real data). Formation of Cu(II)–chloride complexes studied by UV Spectrometric titration ([Cl^- = 0·0–5·0 M, T = 25 and 80 °C)

The formation of relatively weak Cu(II)–chloride complexes in solution has been studied previously by various authors [23–26]. Complex formation between Cu(II) and chloride ion takes place gradually as the chloride concentration increases. There is still some controversy about the number and nature of the species formed at high chloride concentrations. Various experimental methods have been proposed to study this problem, including UV-vis spectroscopy. The UV region between 230 and 500 nm gives information about the evolution of complex formation and about its dependence on the temperature, with complex formation being augmented by increasing temperature.

Figure 4 gives the plot of the two data matrices $\mathbf{D}_3$ and $\mathbf{D}_4$ obtained experimentally in the UV spectrometric titration of samples of Cu(II) ($Cu(ClO_4)_2$) at a concentration of $2{\cdot}5 \times 10^{-5}$ M. The chloride concentration was increased gradually in each sample from 0·0 to 5·0 M using NaCl. A total of 20 spectra and 321 wavelengths were measured in each case. The ionic strength was kept equal to 5·0 M in all the samples using appropriate mixtures of $NaClO_4$ and NaCl stock solutions. The data matrix $\mathbf{D}_3$ was obtained at 25 °C and the data matrix $\mathbf{D}_4$ at 80 °C using a Peltier-based apparatus (Perkin-Elmer). UV-vis spectra were measured for each sample between 230 and 550 nm using a Perkin-Elmer Lambda-19 spectrometer.

Spectra at low chloride concentrations gave a low absorption, increasing considerably when the chloride concentration was increased, especially in the 230–300 nm range. Spectral shifts to higher wavelengths were also observed. When the two experiments are compared, it is clear that in the case of the experiment at 80 °C the shifts are higher and that the complexation process at a particular chloride concentration is more developed at 80 than 25 °C. Observe for instance the higher intensity of the second band at 400 nm, scarcely visible at 25 °C but clearly apparent at 80 °C. All this indicates that the complexation process is not the same at the two temperatures.

The constraints considered during the calculation of the species profiles and their band boundaries were: (a) non-negativity for the concentration and spectral profiles; (b) closure for the concentration profiles, since the sum of the concentrations of the species containing Cu(II) ion should be equal to the known total concentration of Cu(II) (mass balance equation assuming that all species are mononuclear in Cu(II) ion); and (c) selectivity and/or local rank constraints for those regions where some of the species are or are not present.

4. RESULTS AND DISCUSSION

Example 1. Mixture of a two-component equilibrium system (Figure 2)

Figures 5A and 5B give the band boundaries of feasible solutions obtained using the optimization algorithm previously described when non-negativity constraints were applied to both the concentration and spectral profiles and the pure species spectra were also constrained to norm one. Full lines give the maximum and minimum band boundaries. These band boundaries give the profiles giving maximum and minimum signal contributions and, in this case, also the concentration profiles

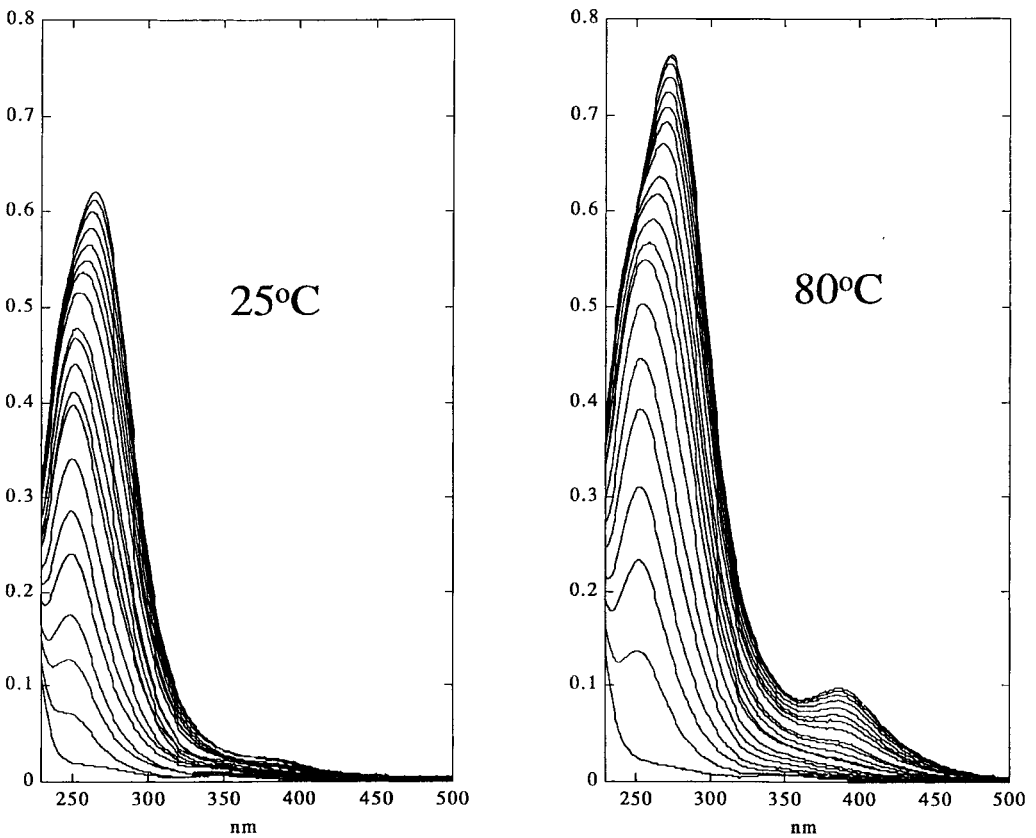

Figure 4. Data example 3: experimental study of Cu(II)–chloride complexation system at different concentrations of chloride (from 0·0 to 5·0 M NaCl). The concentration of Cu(II) in all the spectra was $2{\cdot}5 \times 10^{-4}$ M and the ionic strength was kept to 5·0 M with an appropriate mixture of NaCl and $NaClO_4$ 5·0 M stock solutions. Left: plot of 20 measured spectra (230–500 nm) at 25 °C (data matrix $\mathbf{D}_3$). Right: plot of 20 measured spectra (230–500 nm) at 80 °C (data matrix $\mathbf{D}_4$).

with maximum and minimum areas, since the spectra were always constrained to be of unit area. Dotted lines give the initial profiles used in the optimization, which are in fact the same as those used for the data simulation of Figure 2. These initial profiles were always within the calculated boundaries. Also, observe that the maximum band boundary for the concentration profile of species 1 is a concentration profile which has full selectivity (only one species present) at the beginning of the reaction experiment. The same happens for the minimum band boundary for the concentration profile of species 2, which is a concentration profile which has full selectivity at the end of the reaction experiment. When, instead of starting the optimization with the profiles used for the data simulation, the optimization was started with an estimation of the species profiles obtained from the analysis of the data matrix $\mathbf{D}_1$ using MCR-ALS (broken lines in Figures 5A and 5B), identical boundary solutions were found. The corresponding spectral profiles estimated from MCR-ALS cannot be seen in Figure 5B since they are completely overlapped with the minimum boundary for species spectrum 1 and with the maximum boundary for species spectrum 2. When, instead of spectral normalization, the closure constraint was considered, the calculated boundary bands were only slightly different from those previously found when the spectral normalization constraint was applied instead. This is logical,

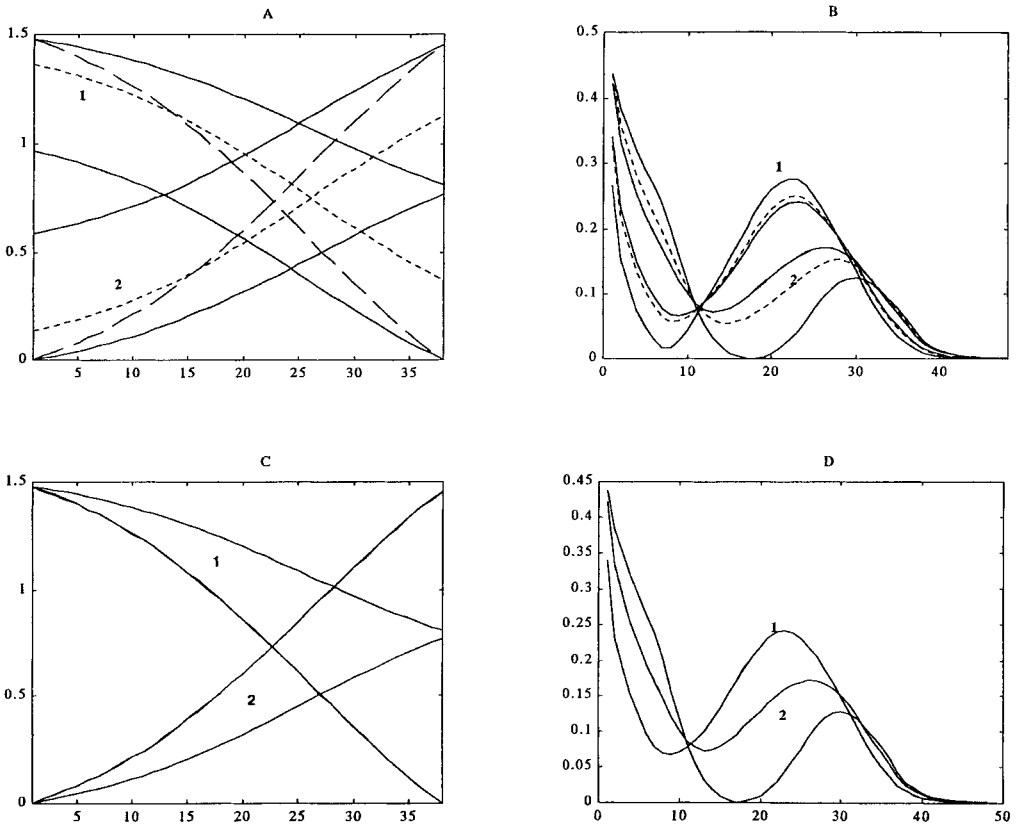

Figure 5. Band boundaries calculated for the species profiles of example 1 (data matrix $\mathbf{D}_1$, Figure 2). Full lines are the band boundaries obtained using the proposed procedure. Dotted lines are the 'true' profiles. Broken lines are the profiles obtained using MCR-ALS. A: band boundaries of concentration profiles obtained under non-negativity and spectral normalization and/or concentration closure constraints. B: band boundaries of spectral profiles obtained under the same constraints as in A. C: band boundaries of concentration profiles obtained under non-negativity, spectral normalization and selectivity constraints. D: band boundaries of spectral profiles obtained under the same constraints as in C.

since for this particular case the data matrix $\mathbf{D}_1$ was in fact built up from concentration profiles fulfilling the closure constraint and with spectral profiles normalized to unit length (Figure 2).

If data selectivity were present at the beginning of the equilibrium process (only species 1 present), the band boundaries of feasible solutions would change to those shown in Figures 5C and 5D, especially for the spectrum of the first species where the solution was obviously unique and the band collapsed to a single curve (only one full line for species spectrum 1 is shown in Figure 5D). However, there was still rotational freedom for the other concentration and spectral profiles as shown by the other full lines in Figures 5C and 5D. The initial profiles used for the band calculations were totally overlapped with the full lines and are not distinguished in these plots. If selectivity were also considered at the end of the equilibrium process (absence of species 1 at the end of the equilibrium process), all the feasible bands would collapse to a single line, since selectivity constraints for the two species concentration profiles would force unique solutions. All this is in agreement with previous studies on selectivity and local rank effects in factor analysis ambiguities [5,6].

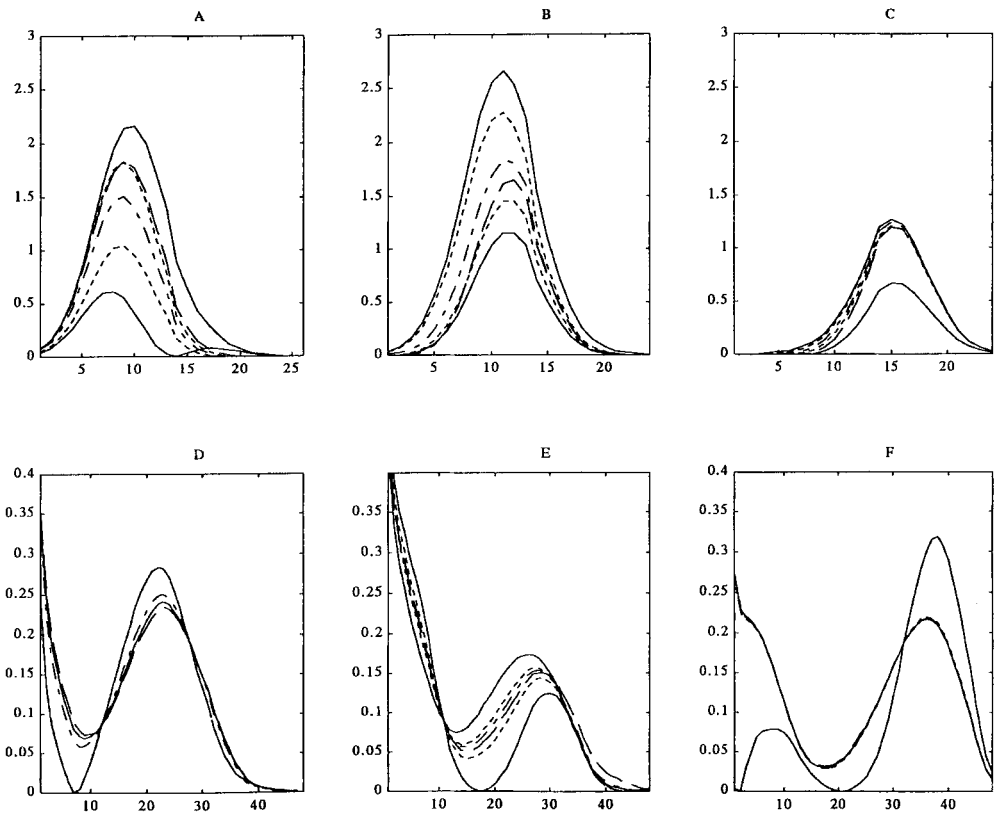

Figure 6. Band boundaries calculated for the component profiles of example 2 (data matrix $\mathbf{D}_2$, Figure 3). Full lines are the band boundaries obtained using non-negativity and spectral normalization constraints. Dotted lines are the band boundaries obtained using non-negativity, spectral normalization and selectivity/local rank constraints. Chain lines are the 'true' profiles. Broken lines are the profiles obtained using MCR-ALS. A–C: elution profiles of species 1–3 respectively. D–F: corresponding spectral profiles.

Example 2. Mixture of three coelution chromatographic components (chromatographic coelution, Figure 3)

Figure 6 gives the results obtained in the calculation of the band boundaries for the elution and spectral profiles of example 2 (Figure 3) under different constraints. Full lines describe the band boundaries when non-negativity constraints were applied to both elution and spectral profiles and the spectral profiles were normalized to unit length. The broad separation between these boundaries indicates that, using only these two constraints, the results obtained by curve resolution would be rather ambiguous. The minimum band boundary of the feasible solution for the elution profile of species 1 showed a double peak, i.e. the solution was not unimodal. When the unimodality constraint was applied, during the optimization this band boundary was corrected accordingly and changed to fulfil this unimodality constraint. This problem, however, may be better solved using appropriate local rank/selectivity constraints.

Band boundaries obtained using local rank constraints (apart from non-negativity and normalization constraints) are plotted using dotted lines in Figure 6. In this particular case the tails of the

elution profiles for the first and second species were constrained to be below 0·01 units for the last four points of the elution profile of the first species and for the last three points of the elution profile of the second species (see Figure 3). Also, the first three points of the front of the elution profile of the third species were constrained to be below 0·01 units. In this way the band of feasible solutions corresponding to the spectral (Figure 6F) and elution (Figure 6C) profiles of the third species collapsed to a unique solution, and the band boundaries of the spectral profiles of the first (Figure 6D) and second (Figure 6E) species were now also narrower. The feasible regions for the elution profiles of the first and second species, although narrower than when no selectivity/local rank constraints were applied (full lines compared to dotted lines in Figures 6A and 6B), were still wide, showing that rotational ambiguities were not totally solved for them yet. In Figure 6 the initial profiles used for the optimization (chain lines) and the MCR-ALS solutions (broken lines) are also plotted. It is seen that these lines appear in most cases to be within the calculated boundaries. It may happen, however, as in Figure 6E, that feasible solutions appear outside the boundary plots. In this case the feasible solutions are still numerically within the boundaries, but when they are plotted as in Figure 6E, they may appear to be outside the boundary plots. The reason for this is that the boundaries give the maximum and minimum of the optimization function defined in terms of signal contribution, but these numerical limits may not coincide exactly with the physical limits in the graphical representation of feasible solutions in two-dimensional plots such as those shown in Figure 6E. For instance, although the boundaries define the maximum and minimum species signal contributions and all the feasible solutions should be within these boundaries, the graphical representation of the boundaries for a particular species in a two-dimensional plot could apparently appear to be outside the boundaries. This is a problem related only to the graphical representation of the boundaries and not to their intrinsic meaning and interpretation. Alternative ways of representation are outside the scope of this work and are left for further discussion.

Finally, the band boundaries obtained for the second chromatographic elution profile using different starting values are compared in Figure 7. This species profile was recovered with large rotational ambiguities, giving wide boundary bands (see Figures 6B and 6E). Figure 7 gives the band boundaries obtained under non-negativity and normalization constraints using the true solution as initial estimates (full lines), as well as (lines with crosses) the band boundaries obtained using the MCR-ALS solution as initial estimates. As can be seen, the band boundaries calculated using these two initial solutions were very similar, thus reinforcing the reliability of the proposed method. Finally, Figure 7 also gives the band boundaries obtained using the local rank constraints previously described for this system (dotted lines).

Example 3 (real data). Formation of Cu(II)–chloride complexes studied by UV spectrometric titration ($[Cl^-]$ = 0·0–5·0 M, T 25 and 80 °C)

Initial profiles for the optimization were obtained in this case from the application of MCR-ALS [5,7,8] to the experimental data matrices $\mathbf{D}_3$ (25 °C) and $\mathbf{D}_4$ (80 °C). Three and four species respectively were initially proposed for the resolution of these two data matrices using MCR-ALS. Constraints applied in both cases were non-negativity, closure and selectivity. For data matrix $\mathbf{D}_3$ (25 °C) the first species (free Cu(II)) was considered to be the only species present in the first spectrum (selectivity), since no chloride had yet been added. In contrast, this species was considered not to be present in the last spectrum at 5·0 M $[Cl^-]$. For data matrix $\mathbf{D}_4$ (80 °C) the concentration of Cu(II) and also the concentration of the second species (first complex Cu–Cl species) were considered negligible in the last spectrum at high chloride concentration. On the other hand, as for data matrix $\mathbf{D}_3$, only the first species (free Cu(II)) was considered to be present in the first spectrum of the series, since no chloride had yet been added. The results confirmed that the selection of the number of

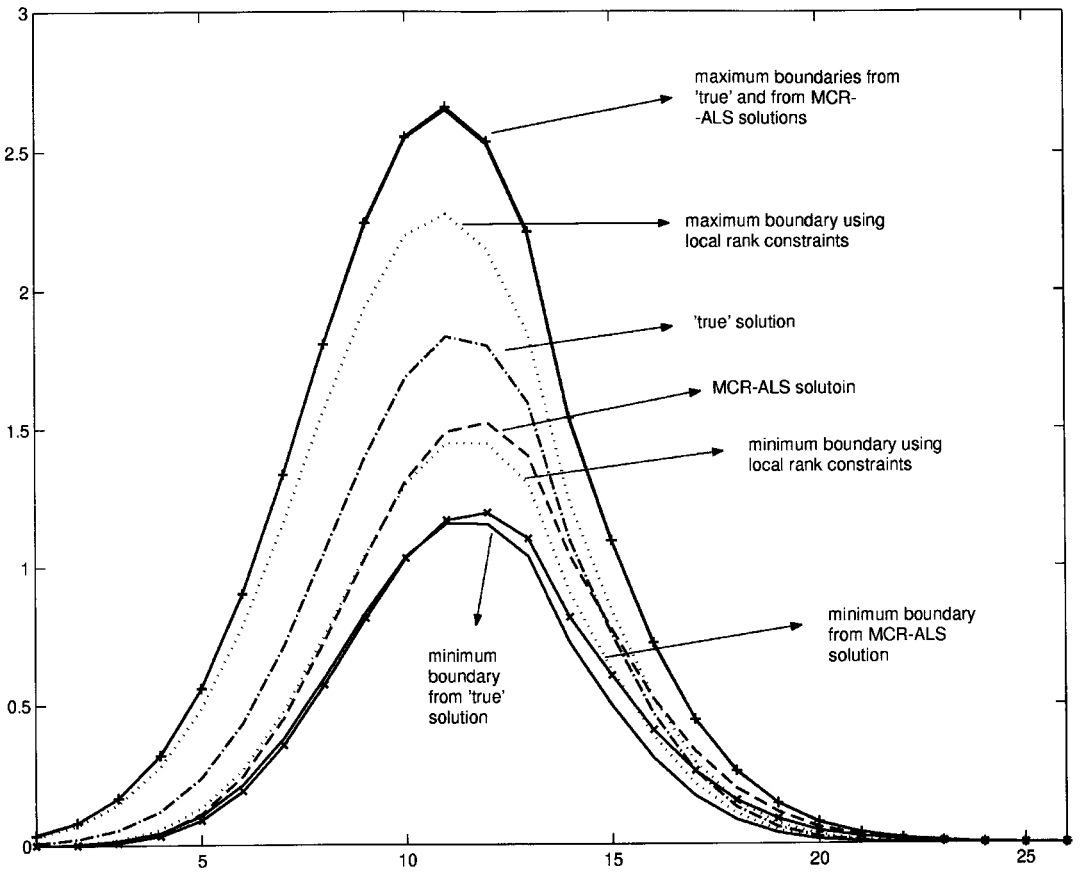

Figure 7. Detail of band boundaries obtained for the elution profile of component 2 using different initial estimates and different constraints.

species in the two experiments was reasonable. A higher or lower number of species gave worse results, both from a data fitting point of view and in terms of the shape of the concentration and spectral profiles recovered in the MCR-ALS analysis. However, the presence at low concentrations of a fourth species at the end of the first experiment at 25 °C cannot be totally excluded and is left for further study. In Figure 8 the finally resolved profiles after the MCR-ALS analysis of the two data matrices are shown as broken lines. The explained data variance was $R^2 = 99{\cdot}79\%$ and 99·98% for $\mathbf{D}_3$ and $\mathbf{D}_4$ respectively.

Using these initial MCR-ALS estimates and using the same constraints as in the MCR-ALS analysis, i.e. non-negativity, closure and selectivity, the band boundaries obtained for the different species profiles using the proposed procedure are given in Figure 8. For data matrix $\mathbf{D}_3$, unresolved rotational ambiguities gave bands that were broader at higher chloride concentrations (Figure 8A), especially for the concentration profiles of the second and third species. The spectrum for the first species (free Cu(II)) is recovered without ambiguities (Figure 8B), since there is selectivity for this species in the first measured spectrum. Moreover, the concentration profile of this first species was also recovered without ambiguities because of the applied local rank constraints [6]. The feasible band for the spectral profile of the second species (Figure 8B) was very narrow, giving a practically

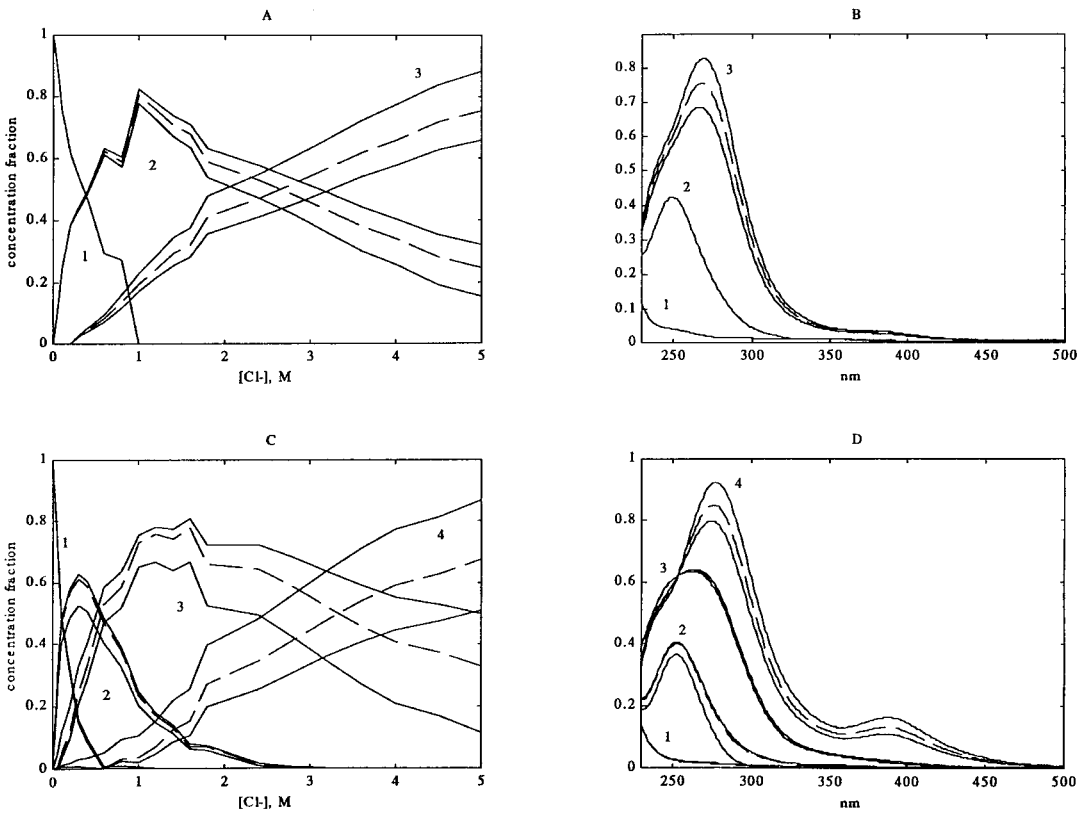

Figure 8. Band boundaries calculated for the species profiles of example 3 (data matrices $\mathbf{D}_3$ and $\mathbf{D}_4$, Figure 4). Full lines are the band boundaries obtained using non-negativity, closure and local rank/selectivity constraints. Broken lineas are the profiles obtained using MCR-ALS. A, B: results at 25 °C (matrix $\mathbf{D}_3$). C, D: results at 80 °C (matrix $\mathbf{D}_4$).

unique solution. The reason for this is also explained in terms of the local rank resolution constraints [6], since the concentration profile of this species always had a region where the other two species did not exist (the first spectrum did not have species 3 and the last spectrum did not have species 1).

Calculation of the band boundaries for species in data matrix $\mathbf{D}_4$ was more difficult owing to the higher complexity of the system. In Figure 8C the band boundaries for the concentration profiles are given. At high chloride concentrations the width of the bands for the third and four species concentration profiles was large, since rotational ambiguities could not be solved for them. As for matrix $\mathbf{D}_3$, the concentration profile for species 1 (free Cu(II)) was recovered without ambiguities. The concentration profile for species 2 (first metal complex) still had some ambiguities, which were not solved with the proposed constraints. The spectrum for species 1 was recovered without ambiguities (Figure 8D), whereas some ambiguities still remained for the spectra of the second, third (only a little) and fourth species. These results are in agreement with the applied local rank constraints.

Comparison of Figures 8A and 8B with Figures 8C and 8D shows clearly the temperature effect in the equilibrium study of the same system. Whereas at 25 °C only two complex species apart from the free metal ion (three species in total) appear to be present at significant concentrations, at higher

temperature (80 °C) the formation of a third complex (four species in total) is very significant, and the formation constants of the different complexes increase considerably with the temperature. On the other hand, the species spectra of Cu(II) and of the first complex (species 2) are very similar in Figures 8B and 8D, whereas the spectrum of the second complex (species 3) appears to be slightly different at the two different temperatures. This, however, needs further study considering the possibility of the simultaneous analysis of the two experiments using MCR-ALS and an extension of the proposed method to the simultaneous analysis of different data matrices (three-way data analysis). Some work is planned in this direction.

5. CONCLUSIONS

Calculation of the maximum and minimum band boundaries of feasible solutions obtained by multivariate curve resolution has been shown to be possible for different simulated and experimental data. This calculation allows the assessment of the effect of constraints on decreasing the rotational ambiguities associated with multivariate curve resolution solutions and allows their quantitative calculation. This opens up the possibility for multivariate curve resolution solutions to be reported accompanied by indeterminations associated with non-solved rotational ambiguities. From the obtained results it is deduced that usual constraints such as normalization, closure, non-negativity, unimodality and, especially, local rank/selectivity decrease considerably the extent of rotational freedom (rotational ambiguities) in multivariate curve resolution solutions.

ACKNOWLEDGEMENTS

The author gratefully acknowledges financial support from the Ministerio de Educación y Cultura de España (DGICYT Project PB96-0377).

REFERENCES

1. Lawton WH, Sylvestre EA. Self modeling curve resolution. *Technometrics* 1971; **13**: 617–633.
2. Sylvestre EA, Lawton WH, Maggio MS. Curve resolution using a postulated reaction. *Technometrics* 1974; **16**: 353–368.
3. Borgen OS, Kowalski BR. An extension of the multivariate component-resolution method to three components. *Anal. Chim. Acta* 1985; **174**: 1–26.
4. Borgen OS, Davidsen N, Myngyang Z, Oyen O. The multivariate *N*-component resolution problem with minimum assumptions. *Mikrochim. Acta* 1986; **II**: 1–6.
5. Tauler R, Smilde A, Kowalski BR. Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution. *J. Chemometrics* 1995; **9**: 31–58.
6. Manne R. On the resolution problem in hyphenated chromatography. *Chemometrics Intell. Lab. Syst.* 1995; **27**: 89–94.
7. Gargallo R, Cuesta-Sànchez F, Massart DL, Tauler R. Validation of alternating least squares multivariate curve resolution for the resolution and quantitation of overlapped peaks obtained in liquid chromatography with diode array detection. *Trends Anal. Chem.* 1996; **15**: 279–286.
8. De Juan A, Vander H, Tauler R, Massart DL. Assessment of new constraints applied to the alternating least squares (ALS) method. *Anal. Chim. Acta* 1997; **346**: 307–318.
9. Henry RC, Kim BM. Extension of self-modeling curve resolution to mixtures of more than three components. Part 1. Finding the basic feasible region. *Chemometrics Intell. Lab. Syst.* 1990; **8**: 205–216.
10. Wentzell PD, Wang J, Loucks LF, Miller KM. Direct optimization of self modelling curve resolution: application to the kinetics of the permanganate–oxalic acid reaction. *Can. J. Chem.* 1998; **76**: 1144–1155.
11. Kim BM, Henry RC. Extension of self-modeling curve resolution to mixtures of more than three components. Part 2. Finding the complete solution. *Chemometrics Intell. Lab. Syst.* 1999; **49**: 67–77.
12. Gemperline P. Computation of the range of feasible solutions in self-modeling curve resolution algorithms. *Anal. Chem.* 1999; **71**: 5398–5404.

13. Maeder M. Evolving factor analysis for the resolution of overlapping chromatographic peaks. *Anal. Chem.* 1987; **59**: 527–530.
14. Keller HR, Massart DL. Peak purity control in liquid chromatography with photodiode-array detection with a fixed size moving window evolving factor analysis. *Anal. Chim. Acta* 1991; **246**: 379–390.
15. Gill PE, Murray W, Wright MH. *Practical Optimization*. Academic Press: New York, 1981.
16. *MATLAB, Version 5.3*. The Mathworks: Natick, MA, 1999.
17. Amhrein M, Srinivasan B, Bonvin D, Schumacher MM. On the rank deficiency and rank augmentation of the spectral measurement matrix. *Chemometrics Intell. Lab. Syst.* 1996; **33**: 17–33.
18. Saurina J, Hernández-Cassou S, Tauler R, Izquierdo-Ridorsa A. Multivariate curve resolution of rank-deficient kinetic spectrophotometric data from first-order kinetic decomposition reactions. *J. Chemometrics* 1998; **12**: 183–203.
19. Bro R, Sidiropoulos N. Least squares algorithms under unimodality and non-negativity constraints. *J. Chemometrics* 1998; **12**: 223–247.
20. *Optimization Toolbox, Version 2.0*. The Mathworks: Natick, MA, 1998.
21. Schittowski K. NLQPL: a Fortran-subroutine solving constrained non-programming problems. *Ann. Oper. Res.* 1985; **5**: 485–500.
22. Dixon LCW. *Non-linear Optimization*. The English Universities Press: London, 1972.
23. Khan MA, Schwing-Weil M. Stability and electronic spectra of copper(II) chloro complexes in aqueous solutions. *Inorg. Chem.* 1976; **15**: 2202–2205.
24. Ashurst KG, Hancock RD. Characterization of inner- and outer-sphere complexes by thermodynamics and absorption spectra. Part 2. Chloro-complexes of copper(II). *J. Chem. Soc., Dalton Trans.* 1981; 245–250.
25. Bjerrum J. Determination of small stability constants. A spectrophotometric study of copper(II) chloride complexes in hydrochloric acid. *Acta Chem. Scand. A* 1987; **41**: 328–334.
26. Tauler R, Rode BM. Reactions of Cu(II) with glycine and glycylglycine in aqueous solution at high concentrations of sodium chloride. *Inorg. Chim. Acta* 1990; **173**: 93–98.

