White, grey and black multicomponent systems

A classification of mixture problems and methods for their quantitative analysis

Yi-Zeng Liang [1], Olav M. Kvalheim and Rolf Manne

Department of Chemistry, University of Bergen, N-5007 Bergen (Norway)

"Every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational and contrary to the spirit of chemistry... If mathematical analysis should ever hold a prominent place in chemistry — an aberration which is happily almost impossible — it would cause a rapid and widespread degeneration of that science."

Auguste Comte, *Philosophie Positive* (1830)

Abstract

Liang, Y.-Z., Kvalheim, O.M. and Manne, R., 1993. White, grey and black multicomponent systems. A classification of mixture problems and methods for their quantitative analysis. *Chemometrics and Intelligent Laboratory Systems*, 18: 235–250.

Multivariate calibration and resolution methods for handling samples of chemical mixtures are examined from the point of view of the analytical chemist. The methods are classified concordant to three different kinds of analytical mixture systems; i.e. 'white', 'grey', and 'black' analytical systems. Advantages and limitations of available multivariate calibration and resolution methods are discussed with respect to the proposed classification of the analytical mixture problem.

CONTENTS

1. INTRODUCTION

Classical methods for mixture analysis are based upon univariate (scalar) data. If there exists a quantitative numerical relationship between the variation of the concentration of a particular analyte and a chemical or physical signal, then this can be used for accurate assessment of the concentration of this chemical species. Unfortunately, such relationships are rarely encountered in the analysis of complex real-world samples. This has lead to the use of various spectroscopic techniques where the chemical composition of a sample is mapped as a spectral profile. Some techniques even generate data in matrix form. Well-known examples of the latter are hyphenated chromatography (liquid chromatography with diode array detection (LC–DAD)), emission–excitation fluorescence spectroscopy and multipulse nuclear magnetic resonance spectroscopy. The development of numerical methods for extracting relevant information from multivariate chemical measurement techniques has thus become a new challenge for the analytical chemist.

The calibration [1] and resolution [2] of multicomponent systems are two of the most important research areas in chemometrics. The methods developed differ both with respect to aim and performance. A method that is excellent for a certain purpose, under certain conditions and assumptions, may be useless when the aims of the analysis, the assumptions and/or the experimental conditions change. Thus, in order to evaluate multivariate resolution methods from the point of view of analytical chemistry, it is necessary first to classify the different mixture problems. We must ask questions like: (1) Is the aim to quantify all or just some selected analytes in the mixture? (2) Do we have to account for the possibility of the presence of unknown (or even unmodelled) interferents in the mixtures? (3) Do we have access to concentrations and reference spectra of a training set of standards (pure chemical species or mixtures) or not?

2. CLASSIFICATION OF MIXTURE PROBLEMS AND METHODS

Mixture samples commonly encountered in analytical chemistry fall into three categories which we, following Chinese usage, shall call *white, grey* and *black multicomponent systems.*

The multicomponent systems which we label white have the following characteristics: the spectra of the chemical species present in the samples, or, spectra from samples that contain possible coexisting interferents in addition to the desired species, are all available. In addition, the concentrations of all the desired analytes must be known for the training set samples. The aim of mixture analysis of white systems is to determine quantitatively the concentrations of some or all chemical species. Methods for this kind of analysis are relatively mature and do almost always provide excellent results [1]. Exceptions are ill-conditioned [1] and seriously nonlinear systems [3].

We give the name black systems to those analytical systems for which there is no a priori information regarding the chemical composition. The aim may nevertheless be to determine simultaneously spectra (resolution) and concentrations (quantitation) of all coexisting chemical species. After the introduction of the self-modeling curve resolution (SMCR) technique [4], several methods for resolving this type of analytical systems have been developed. Most of them are based on instrumental techniques giving data of matrix form, e.g. hyphenated chromatography. In the quantitation step, one needs samples of known concentrations for the resolved chemical components. Effectively, this strategy transforms the black system into a white or grey system, depending on whether the system is fully or only partly resolved.

Grey analytical systems embrace the area between the two kinds defined above. The basic characteristic of these analytical systems is that incomplete knowledge is available regarding the qualitative chemical composition. The aim of the analysis is to determine quantitatively the desired analytes in the presence of unknown coexisting interferents. Thus, spectra or models for the desired analytes are available, but no information regarding possible interferents in the samples. The difference between white and grey lies in their ability to handle interferents in prediction samples. For white systems only modelled interferents in the training set can be accounted for, while grey systems are not constrained by this limitation. Grey systems are frequently encountered in the analysis of real-world samples. Several powerful multivariate calibration methods have been developed both to cope with vectorial data (each sample being represented by a spectrum, for instance) [5] and with bilinear data of matrix form (each sample being represented, for example, by a matrix of chromatographic elution profile time spectra) [6–11].

The classification of mixture problems given here is based solely on the amount of a priori spectral information available for the multicomponent system under investigation. Indeed, the partition into three categories of multicomponent systems defined above can be motivated simply as: (i) white systems are those for which spectral information exists for all the chemical species present in the systems; (ii) grey systems are those for which partial spectral information is available; and (iii) black systems are those for which no spectral information about the pure species is available.

The division between white and grey systems can also be made on the basis of a priori knowledge about the interferents. Thus, for a white system, the spectra of the interferents are known (or modelled), while for a grey system they may be completely unknown.

The classification above will form the framework for our discussion of available methods of multivariate analysis. The situation is sketched in Fig. 1. The vertical axis in Fig. 1 presents the degree of complexity with respect to the data (from scalar via vectorial to matrix form), while the horizontal axis accounts for the degree of complexity of the multicomponent systems (from white via grey to black). Multivariate methods included in the same frame provide almost the same functionality with respect to a particular kind of multicomponent system. Unshaded frames are used for methods which can provide unique
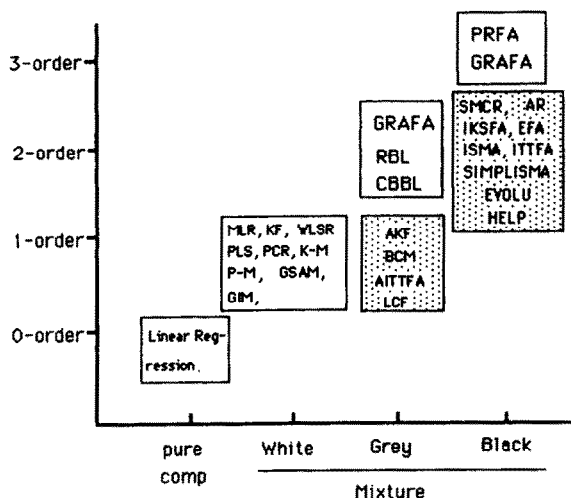
Fig. 1. The relationship between the degree of complexity of analytical system (from 'white' to 'black') and data type required. The multivariate calibration and resolution methods included in each frame have almost same functionality with respect to one kind of analytical systems. Methods encircled by unshaded frames can provide unique solution in the mathematical and chemical sense. Shaded frames indicate methods that provide unique solution for certain special cases and/or only possible solutions or solution regions.

solutions in a strict mathematical and physical sense. Shaded frames indicate methods that provide resolution for certain special cases and/or possible, but not necessarily unique, solutions or solution regions. Fig. 1 implies that for unique solution of a given multicomponent system the order of data needed depends inversely on the spectral information available a priori. Common abbreviations used for the methods and corresponding references cited in Fig. 1 are given in the Appendix.

3. CALIBRATION METHODS FOR WHITE SYSTEMS

With a priori spectral information about all chemical species existing in the samples, regression techniques may be used for determination of concentrations. Multivariate data obtained from spectral analysis can usually be approximated as linear systems [1]:

$$y^T = \sum_{i=1}^{N} c_i s_i^T + e^T \tag{1a}$$

where y^T is the response vector of M sensors (wavelengths) due to N known analytes with known sensitivity coefficient (spectral) vectors $\{s_i^T, i = 1,2,\ldots,N\}$ present at concentrations $\{c_i, i = 1,2,\ldots,N\}$. The residual vector e^T contains measurement noise. Note that T implies transposition, i.e. vectors into row vectors.

The concentrations can be collected in a vector c and the spectra in a matrix S:

$$y^T = c^T S^T + e^T \tag{1b}$$

For a univariate calibration model, Eqns. 1a and b reduce to a scalar equation as follows,

$$y = \sum_{i=1}^{N} c_i s_i + e \tag{2}$$

It is obvious that N unknown concentrations c_i cannot be estimated from one equation. Only when one of the analytes, say j, has a very high sensitivity coefficient compared to the others, can Eqn. 2 provide an estimate of that analyte:

$$y = c_j s_j + e \tag{3}$$

This defines j as a selective sensor (wavelength) for that analyte. The search for chemical or physical measurement methods of high selectivity for a given analyte has haunted analytical chemists for decades.

The difference between a univariate and a vector calibration model (Eqn. 1) can be illustrated in geometric terms. Fig. 2 implies that the vector calibration model can increase the component selectivity significantly and this is, in our opinion, one of the two main advantages of the vector calibration technique compared to the classic (scalar) calibration techniques [12]. (The other main advantage is the possibility of detecting a sample as atypical, i.e. an outlier [1].) Depending on the choice of experimental design the vector calibration methods for treating white analytical systems can be divided into four categories [13]: (i) direct calibration methods; (ii) indirect calibration methods; (iii) generalized standard addition methods; and (iv) generalized internal reference methods.
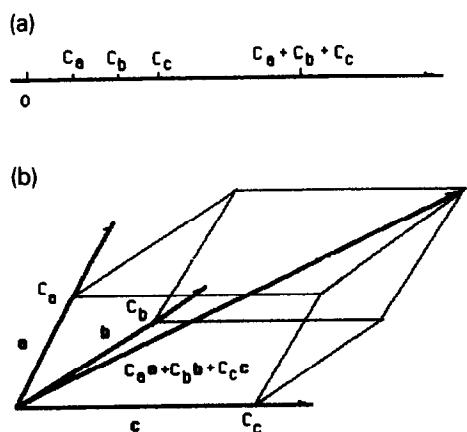
(a)

(b)

Fig. 2. Geometric illustration for univariate (scalar) and multivariate (vectorial) calibration. (a) The univariate calibration model cannot identify or distinguish the three chemical species, a, b and c in a mixture sample. We observe simply the sum of the species and no resolution or quantification in terms of single species is possible. (b) The unit vectors a, b and c represent the spectra of three chemical species in multivariate space. Each species has a unique (and different) vectorial presentation in multivariate space, and, thus, the mixture sample ($c_a a + c_b b + c_c c$) can be resolved and quantified by using multivariate calibration methods.

3.1. Direct calibration methods

Direct calibration methods assume that spectra are available for all chemical species existing

in the mixture. The spectra of the pure components are then used to estimate directly the concentrations of the species from the measured response vector. Multiple linear regression (MLR) represents the standard method of solving white systems under this condition [14]:

$$c^T = y^T S(S^T S)^{-1} \qquad (4)$$

MLR can also be used as a weighted least squares (WLS) procedure [15]. A third alternative is the use of Kalman filtering (KF) [16]. The major advantage of these methods is their simplicity. If the given analytical system is linear and obeys the Beer–Lambert law, one obtains quantitative results for most white systems. If some species interact, indirect calibration methods present a better alternative than direct methods, since we can make a training set of mixture samples and thus accomodate moderate degrees of nonadditivity, which is normally not accounted for by the direct methods.

3.2. Indirect (inverse) calibration methods

The indirect (or inverse) calibration methods assume that we can design and measure the responses of a training set of mixtures with known concentrations (standards). The transformation matrix connecting the spectral response matrix X

TABLE 1

Matrix representations of indirect vector calibration methods *

	K-matrix method	P-matrix method
Beer–Lambert law	$X = CS^T$	$C = XP^T$
Calibration	$C^T X = C^T C S^T$ $S^T = (CC^T)^{-1} C^T X$	$X^T C = X^T X P^T$ $P^T = (XX^T)^{-1} X^T C$
Unknown	$X = CS^T$ $XS = CS^T S$ $C = XS(S^T S)^{-1}$	$C = XP^T$

Improvement of PCR and PLS by means of orthogonal decomposition of measurement matrix X		
PCR	PLS2	PLS1
$X = UGP^T$ $X^+ = PG^{-1}U^T$ $P = CX^+$ G: diagonal matrix	$X = URW^T$ $X^+ = WR^{-1}U^T$ $P = CX^+$ R: triangular matrix	$X = URW^T$ $X^+ = WR^{-1}U^T$ $p_i = c_i X^+$ R: bidiagonal matrix

* X, S and C denote the measurement matrix, the sensitivity coefficient matrix and concentration matrix, respectively.

and the concentration matrix **C** can then be estimated by means of a least-squares procedure. The concentrations of the analytes in prediction samples can then be obtained by use of the transformation matrix. For ease of comparison important relations for some common indirect calibration methods are collected in Table 1.

The first indirect calibration methods to be developed were the so-called **K**-matrix and **P**-matrix methods [17–19]. The **K**-matrix method requires two matrix inversions (see Table 1). This may lead to propagation and accumulation of errors, and for this reason the **P**-matrix method seems to be a better choice. A well-known problem with the **P**-matrix method, however, is the presence of collinearities in the matrix of spectral responses **X**. Collinearity leads to problems with the inversion of $\mathbf{XX}^T$ and thus, possibly, to erroneous regression coefficients.

As a remedy against this pitfall, biased regression methods such as principal component regression (PCR) [20], partial least-squares (PLS) regression [1,21,22] and, recently, principal covariates regression [23] have been developed. Several of the biased regression methods decompose the spectral responses into orthogonal latent variables, which are subsequently used in the regression step. The estimation procedure for the indirect calibration methods PCR and PLS is illustrated in Table 1. Thomas and Haaland [24] recently compared the PCR and PLS methods. It appears that PLS usually provides better concentration estimates than PCR, and that the optimum number of latent components for prediction is smaller for PLS than for PCR.

3.3. Generalized standard addition and internal reference method

The standard addition method and internal reference methods are well-known in analytical chemistry. Their extensions to multivariate calibration are called the generalized standard addition method (GSAM) [25] and generalized internal reference method (GIRM) [26], respectively. They aim at quantification of selected analytes in the presence of known interferents. The basis of the methods is the comparison of the responses of a mixture sample of unknown concentrations with the responses obtained after adding standards of known concentrations. The methods can be used to account for matrix effects or drift of the sensors.

3.4. Nonlinearity and calibration

All methods discussed so far are based on the linear model; nonlinearity presents a problem in multivariate calibration. Mark [3] investigated the effects of nonlinearity in multivariate calibrations by means of artificially induced nonlinearity. Recently Long et al. attempted to handle nonlinearity by means of neural networks [27].

A third way to accommodate a moderate degree of nonlinearity is to expand the spectral description of the systems with squared terms. This can be done on the original spectral variables after a variable reduction [28,29] or by decomposing the spectral information in terms of latent variables and including squared terms in the score vectors [30]. To our knowledge, no one has investigated this approach so far.

It is well known that nonlinear models can be approximated by linear models in small regions. Næs [31], and later Danielson and Malmquist [32], have utilized this fact to handle moderate degrees of nonlinearity within the constraints of a linear model. Their solution is to estimate local calibration models. The only drawback of the methods is the need for rather large training sets and the fact that a new calibration model must be estimated for every prediction sample.

A special kind of nonlinearity is introduced by closure. When there is no absolute scale of spectral intensity, spectra are often normalized to constant intensity or constant height of a certain peak, for example. Such normalizations may destroy the linearity normally assumed between spectral intensity and component concentration. Karstang and Manne [33] have recently developed two methods of optimized scaling by which this linearity is recovered.

For really serious nonlinearity, there seems to be no other solution than to turn to nonlinear regression methods. Collinearity represents a major problem in these methods. In order to circum-

vent this problem, Wold [34] and Taavitsainen and Korhonen [35] have developed nonlinear methods that solve the collinearity problems by introducing latent variables.

3.5. Evaluation of performance

The calibration methods discussed above are all quantitative methods. However, their applicability depends upon the problem under study. Thus, their reliability, adaptability and limitations are major issues for practical applications. Lorber [36] discussed the problems in terms of selected figures of merit, e.g. sensitivity, selectivity, detection limit and error propagation. Such figures of merit are crucial for evaluating the reliability of analytical results and research in this direction is continuing [12,37]. The topic of how to obtain reliable quantitative results for white analytical systems has been well reviewed by Gemperline [38].

4. CALIBRATION METHODS FOR GREY SYSTEMS

Grey multicomponent systems are characterized by incomplete or partial a priori knowledge of the spectra of the chemical species. The aim of the analysis is to estimate concentrations of some chemical species in the presence of unknown or, more precisely, unmodelled interferents. There are two distinct categories of methods for treating grey systems. One is based on vector calibration techniques [5,39–41], while the other uses matrix calibration techniques and thus presupposes characterization techniques which provide data in the form of a matrix [6–11,42–44]. Both categories require a training set with known concentrations and spectra. Because of lack of information, however, vector calibration methods do not give unique solutions for grey systems, while the matrix calibration methods do so in a strict physical and mathematical sense. The two types of methods will now be examined in some detail.

4.1. Calibration methods for vectorial data

The vector calibration model in the presence of unknown background interferents can be ex-

pressed as

$$y^T = \sum_{i=1}^{N} c_i s_i^T + f^T + e^T \tag{5}$$

The only difference between Eqn. 5 and Eqn. 1 is the inclusion of the row vector f^T of length M containing all the contributions from the background interferents. The contributions from the background can be broken down into contributions from each interferent:

$$f^T = \sum_{i=N+1}^{N+P} c_i s_i^T \tag{6}$$

where P is the number of the unknown background interferents.

With knowledge only about y^T and s_i^T ($i = 1, 2, \ldots, N$), the concentrations c_i of the desired analytes cannot be determined uniquely even for the simplest case ($N = 1$) because of the unknown interference vector f^T. The situation with only one analyte in the presence of one or more interferents is illustrated in Fig. 3. Apart from the experimental spectrum y^T, we know only the spectrum of the analyte, s_1^T, i.e. the direction of the vector s_1^T but not its magnitude c_1, and, thus, not its contribution to y^T. Fig. 3 shows that c_1 can be varied over an interval constrained by the condition of positivity of the elements of the background vector. This is a valid constraint in most spectroscopic techniques.

Fig. 3 implies that the concentration estimate obtained by any standard fitting technique will be meaningless unless one has definite spectral knowledge of the background spectrum. Rutan and Brown [39] showed that if the response of analyte is selective for some wavelength region one can find solution for the problem by adaptive Kalman filtering technique. Unfortunately this precondition is often too strong for the spectral analysis of real samples. Osten and Kowalski [40] proposed two possible methods — a perpendicular projection technique and an extreme-vertex projection technique — in order to obtain quantitative results from samples with a background constituent. Combining the GSAM [25] and iterative target testing factor analysis technique (IT-TFA) [45], Liang et al. developed an algorithm to
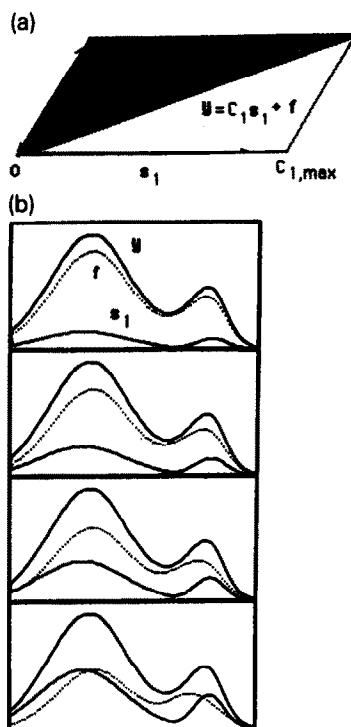
(a)

(b)

Fig. 3. (a) Geometric illustration for the case of quantification of one chemical analyte (standard) in the presence of an unknown interferent. The figure shows the vectorial presentation y of a mixture spectrum; the vectorial presentation s_1 of the spectrum of the standard and a region of vectorial presentations of different background spectra f. As the spectrum of the interferent is unknown, the concentration estimates of the standard can take all values between 0 and $c_{1,max}$. The upper estimate is obtained by the positivity constraint on the spectrum of the interferent. Thus, only a region of possible solutions is obtained. (b) Illustration showing that the mixture spectrum y can be reconstructed from different concentrations of the standard by varying the background spectrum f (dotted).

attack the problem [41]. Unfortunately, all the methods above can only give a possible solution because of lack of information needed for unique resolution.

Karstang and Kvalheim [5] developed a method for local curve fitting (LCF) which makes only two assumptions: (i) the spectra of the interferents are positive at all wavelengths; and (ii) there are local spectral maxima for the interferents. The method will be illustrated here for a one-component calibration model with one interferent.

We write the model for our example as:

$$y^T = c_1 s_1^T + f^T \qquad (7)$$

The first derivative of Eqn. 7 is:

$$d y^T / dw = c_1 \, d s_1^T / dw + d f^T / dw \qquad (8)$$

For every wavelength k where the interferent has a maximum $d f_k / dw = 0$ and $c_1 = (d y_k / dw)/(d s_k / dw)$, i.e. we have created a model where the influence of the interferent is eliminated. In order to be able to find these wavelengths, we need to model the interferent. In LCF this is done by projecting the response vector y^T for the mixture sample on s_1^T to obtain a first estimate of f^T. This estimate is improved by using the positivity constraint for the spectrum of the interferent [5]. Although the LCF method may be the best one available at present for background correction for vectorial data, the method is sensitive to large overlap between the spectra of the interferents and the predicted analytes.

4.2. Calibration methods for two-way bilinear data

If bilinear data are available, the situation changes dramatically. Such data are produced, for instance, by 'hyphenated' chromatographic methods.

Bilinear data are collected in two dimensions, e.g. time and wavelength. Assuming both variables to be continuous we may write the intensity for a pure component at time t and wavelength w as

$$X(t, w) = c(t) s(w) \qquad (9)$$

where $s(w)$ represents the spectral intensity of a standardized sample and $c(t)$ the concentration of the substance eluted at time t. With discrete variables we write Eqn. (9) as

$$X_{jk} = c_j s_k \text{ or } \mathbf{X} = \mathbf{c s}^T \qquad (10)$$

i.e. the data matrix of a pure substance is of rank one. For a mixture one may sum over the components i

$$X_{jk} = \sum_{i=1}^{N} c_{ji} s_{ik} \text{ or } \mathbf{X} = \sum_{i=1}^{N} \mathbf{c}_i \mathbf{s}_i^T \qquad (11)$$

Provided that spectral and concentration profiles

are all different, the rank is equal to the number of components. The one-to-one correspondence between rank and number of chemical species expressed by Eqn. 11, makes a crucial distinction between matrix calibration techniques and vector calibration techniques.

4.2.1. Rank annihilation methods

The rank-one property (Eqn. (10) of a matrix acquired for a pure chemical species, was first utilized for quantitative analysis of grey multicomponent systems by Ho et al. in their rank annihilation factor analysis (RAFA) method [6,7]. They estimated the concentrations of standards using the relation

$$\text{rank}(\mathbf{Y} - c_i\mathbf{X}_i) = \text{rank}(\mathbf{Y}) - 1 \qquad (12)$$

where $\mathbf{Y}$ is the matrix obtained for the mixture and $\{\mathbf{X}_i, i = 1,2,\ldots,N\}$ are matrices of rank one obtained for the pure chemical species. Eqn. 12 provides the conditions for accurate estimates of the concentrations.

Ho et al. used an iterative approach to estimate the concentrations at which Eqn. 12 is fulfilled. Lorber presented a noniterative RAFA based on a singular-value decomposition of the measurement matrix $\mathbf{Y}$ and reformulating the estimation of the concentrations into a generalized eigenvalue problem [8]. Sanchez and Kowalski have extended RAFA into multi-analyte estimation calling their extension generalized rank annihilation factor analysis (GRAFA) [43–45]. One might add that Kubista's method for the analysis of correlated spectra [46] is closely related to GRAFA.

Methods for rank annihilation are simple and powerful methods. A great advantage of the iterative rank annihilation method is that quantification is possible without prior estimation of rank of the two-way data. For techniques where drift is a problem, however, rank annihilation methods break down without proper data pretreatment. Under such circumstances it may be better to use methods for black systems first, in order to resolve the system completely. After identification of the desired analytes, absolute quantification is

achieved by comparing the integrated concentration profiles from the mixtures with those calculated from standards with known concentrations.

4.2.2. Bilinearization methods

In analogy with Eqn. 5, a calibration model for two-way data can be written as

$$\mathbf{Y} = \sum_{i=1}^{N} c_i\mathbf{X}_i + \mathbf{F} + \mathbf{E} \qquad (13)$$

In Eqn. 11, $\mathbf{Y}$ is the data matrix of the mixture sample with unknown background interferents, $\{\mathbf{X}_i, i = 1,2,\ldots,N\}$ are the data matrices of the standards of known analytes, $\mathbf{F}$ represents the background and $\mathbf{E}$ is the residual matrix representing measurement noise. For bilinear data, the rank P of the background matrix is obtained simply as

$$P = \text{rank}(\mathbf{F}) = \text{rank}(\mathbf{Y}) - N \qquad (14)$$

The rank information can be used to model the background matrix in terms of principal components. Thus, Eqn. 13 can be written as:

$$\mathbf{Y} = \sum_{i=1}^{N} c_i\mathbf{X}_i + \sum_{i=N+1}^{N+P} t_i p_i^{\mathrm{T}} + \mathbf{E} \qquad (15)$$

Only when the concentrations $\{c_i, i = 1,2,\ldots,N\}$ are estimated accurately, will the background matrix have rank P and the residual matrix $\mathbf{E}$ be at the level of measurement noise so that the rank relation expressed by Eqn. 14 is fulfilled. Note that Eqn. 14 is basically of the same type as Eqns. 12 used in RAFA. The bilinearity constraint for $\mathbf{F}$ makes the estimation of $\{c_i, i = 1,2,\ldots,N\}$ possible by means of optimization techniques. The so-called residual bilinearization (RBL) [9,10] and constrained background bilinearization (CBBL) methods [11] are based on this approach. The assumptions used in these methods are: (i) bilinearity for all measured two-way data; (ii) that the correct rank can be determined for the mixture data; and (iii) that all the N known analytes are present in amounts sufficient for detection from the analyzed mixture data. For hyphenated chromatographic data, the so-called zero-component

regions can be used to estimate a realistic lower limit for detection [47,48].

Because of the use of matrices acquired in separate runs for the mixtures and standards, the bilinearization methods, just as RAFA, will be vulnerable to instrumental drift. However, the study of Öhman et al. [10] suggests that the bilinearization methods are more robust towards drift than RAFA. As discussed above, the performance of both noniterative rank annihilation factor analysis and bilinearization methods depends strongly on correct estimation of rank.

5. METHODS FOR RESOLVING BLACK SYSTEMS

As defined above, there is no a priori information concerning the chemical composition of the samples available for 'black' systems. The aim of the analysis is nevertheless to estimate spectra and concentrations of all the coexisting chemical species.

There are three distinct categories of black mixture systems, corresponding to partitions into one-way, two-way and three-way instrumental data. Resolution of one-way data demands that a set of mixture samples is available with the same chemical species differing only in concentrations. With two-way or three-way data resolution may be possible even from a single mixture sample. Characterization techniques providing one-way or two-way data both demand a matrix of chemical measurement data, with row vectors corresponding to spectral profiles of mixtures with varying concentrations of the pure species. The difference is that one-way characterization demands many mixture samples, while two-way characterization techniques extract the evolving concentration profile from one mixture sample. Most methods that can be used for one-way data measured on several mixtures can also be used for two-way data obtained for a single mixture. For this reason we treat the two categories of systems together under the name of matrix resolution methods.

The third category of methods uses three-way data obtained as matrices of spectral measurements for one sample at different concentrations,

i.e. chromatography in combination with a hyphenated spectroscopic technique. The corresponding methods will be called three-way resolution methods.

5.1. Matrix resolution methods

For matrix data, the aim of the analysis can be summarized in mathematical form as:

$$\mathbf{X} = \sum_{i=1}^{N} c_i s_i^T + \mathbf{E} \tag{16a}$$

The measurement matrix is decomposed into concentration vectors $\{c_i, i = 1,2,\ldots,N\}$ and sensitivity vectors $\{s_i^T, i = 1,2,\ldots,N\}$ (T implies row vectors) for the pure chemical species. The matrix $\mathbf{E}$ accounts for instrumental and experimental noise.

By collecting the concentration vectors and the spectral profiles of the pure chemical species into a concentration matrix $\mathbf{C}$ (dimension $K \times N$) and a spectral matrix $\mathbf{S}$ (dimension $N \times M$), respectively, Eqn. 16a can be rewritten as:

$$\mathbf{X} = \mathbf{C}\mathbf{S}^T + \mathbf{E} \tag{16b}$$

The objective of matrix resolution methods for black analytical systems can be formulated as: (i) partitioning the data matrix into structure and noise; and (ii) finding the concentration and spectral vectors corresponding to the pure chemical species. Most methods use a principal component decomposition in order to accomplish (i) and then use some criterion for rotating the orthonormal principal components into concentration vectors and spectra corresponding to pure chemical species.

5.1.1. Self-modeling curve resolution methods

Almost all available methods for resolving data of matrix form are based upon the self-modeling curve resolution (SMCR) approach developed by Lawton and Sylvestre [4]. Their method for resolving two-component mixtures makes two basic assumptions: (i) non-negative concentrations and spectra of the resolved components; and (ii) additivity of the component spectra. Thus, SMCR can be used for most spectral data without the need

for any assumptions regarding peak shape, location, or identity. Principal component analysis is used to determine the number of components present in the mixture(s) [2]. The loading and score vectors are subsequently rotated to give vectors fulfilling the non-negativity constraints. In favorable cases, possible solutions are restricted to narrow bands by the non-negative constraints mentioned above. Unique resolution requires at least one selective spectral variable for each component. Thus, the 'art' of SMCR is to find the best rotation of the principal components. Physically correct solutions, however, are never guaranteed.

SMCR was first extended to three-component systems by Ohta [49]. Meister then developed an SMCR algorithm that produced unique solutions for N-component mixtures using a criterion of maximum spectral dissimilarity between mixture samples [50]. This assumption tended to exaggerate differences between the pure spectra of the chemical components. Although the spectra of the pure chemical components are not necessarily determined by Meister's criterion, the method provided useful insight. Kawata et al. devised an entropy optimization technique to produce unique estimates of pure-component spectra for three-component systems [51]. The spectra of the pure components obtained in this way were considered to be of the simplest form in agreement with the mixture spectra. Independently, Vandeginste et al. developed a similar constraint; the spectra of the pure component should have the 'simplest' profile, defined mathematically as the curve having the smallest area-to-norm ratio [52].

Borgen and Kowalski investigated the potential of the non-negativity constraints for spectra and concentrations for locating solution bands for a three-component system in a more detailed way [53]. They were able to reduce a three-dimensional vector space to a two-dimensional by using special normalization conditions. Later, Borgen et al. extended the technique to systems with more than three components [54].

Karjalainen [55–57] developed a method of alternating regression (AR), which does not require an initial decomposition into principal components. Eqn. 14b solved with respect to concentration vectors and spectra of pure components gives:

$$C = XS(S^TS)^{-1} \qquad (17a)$$

$$S^T = (C^TC)^{-1}C^TX \qquad (17b)$$

By randomly choosing starting vectors for the pure species, iteration between Eqns. 17a and b continues until convergence. So far, alternating regression has not been compared with other methods. Problems are anticipated if the initial rank estimate is erroneous and for the case with strong spectral or chromatographic overlap.

The iterative key-set factor analysis (IKSFA) method of Schostack and Malinowski [58] shares some features with the AR method, but instead of starting from randomly chosen pure spectra, IKSFA starts with a principal component decomposition and subsequently uses a criterion of maximum spectral dissimilarity to obtain pure spectra. A direct alternative to IKSFA is proposed below.

During the last decade, Windig and coworkers have developed several SMCR methods making extensive use of graphic procedures to provide a more detailed evaluation than possible by numeric methods. In early work, Windig et al. [59] developed a graphic rotation method whereby principal components axes were rotated in increments of 10° to visually determine the orientation of the pure component axis by looking for dense clusters of large loadings. The graphic maximum variance criterion used was related to numeric criteria used in VARIMAX and similar methods [60]. The graphic rotation procedure was then developed into a spectral isolation method employing so-called variance diagrams to detect the direction in multivariate space with maximum information [61]. Using variance diagrams and key-set spectra to reveal the loadings with maximum determinant, an interactive self-modeling multivariate analysis (ISMA) method was presented by Windig et al. [62] which has been further developed into the SIMPLISMA method [63]. The basic assumption in this last method is that pure spectral variables can be detected using a maximum-variance criterion. The pure variables are detected one by one, and a variable projec-

tion is performed for each dimension until the residual matrix corresponds to noise level. The method is related to the successive orthogonal projection procedure developed by one of the present authors [63] with pure variables as marker variables. Thus, one might extend Windig's approach to accommodate the possibility of one-component regions and using the corresponding marker object (pure mixture) projection [4], in which whole spectra or selective parts of spectra are selected by maximizing spectral dissimilarity [50,63].

5.1.2. Resolution methods for hyphenated techniques

Although, in principle, the methods mentioned in the previous section can be used for the resolution of data from hyphenated chromatography, several new methods have been developed taking advantage of the special shape of the concentration profiles of the pure components in chromatography. We thus know a priori that the chromatographic concentration profile for each pure component has one and only one maximum. Furthermore, the continuous nature of chromatographic peaks implies sequential appearance and disappearance of the concentration profile of a pure component as an additional constraint, that can be utilized in the resolution process. A third important point is the increased possibility of finding selective chromatographic and/or spectral regions, which are crucial for unique resolution of black mixture samples [47]. Some or all of these constraints have been used in several novel methods, such as iterative target testing factor analysis (ITTFA) [45,65–67], evolving factor analysis (EFA) [68,69] and, recently, heuristic evolving latent projections (HELP) [47,48].

5.1.2.1. Iterative target transformation factor analysis.
Iterative target testing factor analysis (IT-TFA) was developed independently by Gemperline [64] and by Vandeginste et al. [45,66,67]. Gemperline constructed test vectors for the concentration profiles using a form of the uniqueness test described by Malinowski and Howery [2]. All values in a test vector were set to zero except one element which was given a value of 1.0. The

resulting vector was used in a target testing procedure [2] to locate the concentration profile in a hyphenated chromatographic data matrix giving the minimum sum of squares for the difference between the test vector and the predicted vector. The test vectors matching the concentration profiles were then iteratively refined to produce non-negative concentration profiles of the pure chemical components. The ITTFA method has been shown to resolve mixtures of two, three and four overlapping peaks.

Vandeginste et al. obtained test vectors for the concentration profiles using principal components decomposition followed by VARIMAX rotation [42]. They have further shown that ITTFA with concentration profiles as test vectors gives better results than if spectra are used for that purpose [66]. In a particular application, this will, of course, depend upon the relative selectivity in the spectral and chromatographic dimensions.

Although ITTFA has proved to be a good method for resolving data from hyphenated chromatography, it should be pointed out that the method cannot guarantee unique resolution in a strict mathematical and chemical sense, since the iteratively refined concentration profiles result from successive projections onto a possible non-unique solution space.

5.1.2.2. Evolving methods.
The methods labelled evolving methods take advantage of the continuous character of the elution profiles, by revealing and using the sequential appearance and disappearance of the concentration profile of each chemical species. The first method of this kind was evolving factor analysis (EFA), developed by Maeder and coworkers [68,69]. It was originally developed for the determination of equilibrium constants from spectrophotometric titration curves, but was then extended to chromatography using the evolving information in the forward and backward directions in order to obtain concentration windows for the eluting chemical species. This information was subsequently used to obtain estimates of the concentration profiles of the pure chemical species by using the 'zero-concentration window' for each eluting species to estimate a rotation vector for transforming the

principal components decomposition of the whole matrix into concentration profiles and spectra of the pure species. Unfortunately, the zero-concentration window for a species defines a rank-deficient estimation problem, with the result that the resolution is indeterminate [69].

A second major problem with the EFA method is the problem of minor chemical species coeluting with major ones. Keller and Massart [70] developed a fixed size moving window EFA with improved resolution for minor components compared to the original method.

A combined visual/numeric method based upon latent-projective graphs and local rank analysis has been developed by two of the present authors [47]. In recent work fixed size moving window EFA is extended to an 'eigenstructure tracking analysis' procedure in order to obtain a rank map showing the number of chemical species eluting at every retention time [71]. These developments constitute a new method, called heuristic evolving latent projections (HELP). The name implies an inductive, stepwise and graphic character, i.e. it acts as a datascope penetrating into local structures of the data [47,48,71].

Also using the information of the sequential appearance and disappearance of the concentration profiles of the pure chemical components, Schostack and Malinowski [72,73] developed a method called evolutionary factor analysis (EVOLU). This method provides upper and lower bounds for spectra and concentration profiles of the pure chemical species.

5.2. Three-way resolution methods

As discussed above, most methods utilizing two-way data cannot provide unique resolution for black systems. Thus, more efficient methods are still needed. Two resolution methods using three-way data have been developed recently [74,75]. Burdick et al. [74] essentially extended the target transformation technique to three-way factor analysis. A multiscaling technique is used to find the common row and column space for all the 'slabs' in three-way data. Rotation matrices that transfer the common vector spaces into spectra of pure components are subsequently ob-

tained by analysis of the projection matrix. Unique resolution into spectra of pure components can be obtained by rotation of the common principal components basis. The method may be called projective rotation factor analysis (PRFA).

The generalized rank annihilation factor analysis (GRAFA) can also be extended to trilinear decomposition [75]. The only difference between two-way and three-way GRAFA is that the latter uses linear combinations of all the slabs instead of single ones.

Both PRFA and three-way GRAFA have been successfully applied to emission-excitation fluorescence data [74,75].

6. CONCLUSIONS

Quantification of multicomponent systems by means of multivariate calibration and resolution methods has released the chemist from constraints imposed by classic analytical methods. On the other hand, the application of these methods has placed new burdens upon the shoulders of the cautious analytical chemist. He must not only be a knowledgeable chemist, but also needs competence in applied mathematics, computer science and statistics in order to be at the forefront of his profession. These requirements reflect the development of chemical instrumentation during the last decade and the penetration of chemical measurement techniques, to an ever-increasing extent, into new areas.

Despite the prophesy of August Comte, cited at the beginning of this tutorial, analytical chemistry has not deteriorated or lost its spirit because of this. On the contrary, analytical chemistry has revived and expanded to become the science of measurements [76] and will surely continue to do so in the future.

APPENDIX

Common abbreviations for available multivariate calibration and resolution methods and corresponding references.

AITTFA	Additional iterative target testing factor analysis	[41]
AKF	Adaptive Kalman filtering	[39]
AR	Alternating regression	[55-57]
BCM	Background correction method	[40]
CBBL	Constrained background bilinearization	[11]
EFA	Evolving factor analysis	[68,69]
EVOLU	Evolutionary factor analysis	[72,73]
GIRM	Generalized internal reference method	[26]
GRAFA	Generalized rank annihilation factor analysis	[6–8,43,75]
GSAM	Generalized standard additional method	[25]
HELP	Heuristic evolving latent projections	[47,48]
IKSFA	Iterative key-set factor analysis	[58]
ISMA	Interactive self-modeling multivariate analysis	[62]
ITTFA	Iterative target testing factor analysis	[42,65–67]
KF	Kalman filtering method	[15]
K-M	K-matrix method	[17]
LCM	Local curve fitting	[5]
MLR	Multiple linear regression	[14]
PCR	Principle component regression	[1,20]
PLS	Partial least squares	[1,21,22]
P-M	P-matrix method	[19]
PRFA	Projective rotation factor analysis	[74]
RBL	Residual bilinearization	[9]
SIMPLISMA	Interactive self-modeling multivariate analysis	[63]
SMCR	Self-modeling curve resolution	[4]
WLSR	Weighted least squares regression	[16]

REFERENCES

1 H. Martens and T. Næs, *Multivariate Calibration*, Wiley, Chichester, 1989.

2 E.R. Malinowski and D. Howery, *Factor Analysis in Chemistry*, Wiley, New York, 1980.

3 H. Mark, Catalogue of the effects of nonlinearity on multivariate calibrations, *Applied Spectroscopy*, 42 (1988) 832–844.

4 W.E. Lawton and E.A. Sylvestre, Self modeling curve resolution, *Technometrics*, 13 (1971) 617–633.

5 T.V. Karstang and O.M. Kvalheim, Multivariate prediction and background correction using local modeling and derivative spectroscopy, *Analytical Chemistry*, 63 (1991) 767–772.

6 C.-N. Ho, G.D. Christian and E.R. Davidson, Application of the method of rank annihilation to quantitative analysis of multicomponent fluorescence data from the video fluorometer, *Analytical Chemistry*, 50 (1978) 1108–1113.

7 C.-N. Ho, G.D. Christian and E.R. Davidson, Simultaneous multicomponent rank annihilation and application to multicomponent fluorescent data acquired by the video fluorometer, *Analytical Chemistry*, 53 (1981) 92–98.

8 A. Lorber, Quantifying chemical composition from two-dimensional data arrays, *Analytica Chimica Acta*, 164 (1984) 293–297.

9 J. Öhman, P. Geladi and S. Wold, Residual bilinearization. Part 1: Theory and algorithms, *Journal of Chemometrics*, 4 (1990) 79–90.

10 J. Öhman, P. Geladi and S. Wold, Residual bilinearization. Part 2: Application to HPLC–diode array data and comparison with rank annihilation factor analysis, *Journal of Chemometrics*, 4 (1990) 135–146.

11 Y.-Z. Liang, R. Manne and O.M. Kvalheim, Constrained background bilinearization, *Chemometrics and Intelligent Laboratory Systems*, 14 (1992) 175–184.

12 Y.-Z. Liang and O.M. Kvalheim, Component selectivity

and identification of the collinearity for multivariate calibration, to be submitted.

13 L.S. Ramos, K.R. Beebe, W.P. Carey, E. Sánchez, M.B.C. Erickson, B.E. Wilson, L.E. Wangen and B.R. Kowalski, Chemometrics, *Analytical Chemistry*, 58 (1986) 294R–315R.

14 R.H. Liu, S. Ramesh, J.Y. Liu and S. Kim, Qualitative and quantitative analyses of commercial polychlorinated biphenyl formulation mixtures by single ion monitoring gas chromatography/mass spectrometry and multiple regression, *Analytical Chemistry*, 56 (1984) 1808–1816.

15 D.M. Haaland and R.G. Easterling, Application of new least-squares for the quantitative infrared analysis of multicomponent samples, *Applied Spectroscopy*, 36 (1982) 665–673.

16 H.N.J. Poulisse, Multicomponent analysis computations based on Kalman filtering, *Analytica Chimica Acta*, 112 (1979) 361–374.

17 C.W. Brown, P.F. Lynch, R.J. Obremski and D.S. Lavery, Matrix representations and criteria for selecting analytical wavelength for multicomponent spectroscopic analysis, *Analytical Chemistry*, 54 (1982) 1472–1479.

18 H.J. Kisner, C.W. Brown and G.J. Kavernos, Simultaneous determination of triglycerides phospholipids and cholesteryl esters by infrared spectrometry, *Analytical Chemistry*, 54 (1982) 1479–1485.

19 C.W. Brown, P.F. Lynch, R.J. Obremski, D.S. Lavery, H.J. Kisner, C.W. Brown and G.J. Kavarnos, Multiple analytical frequencies and standards for the least-squares spectrometric analysis of serum lipids, *Analytical Chemistry*, 55 (1983) 1703–1707.

20 I.A. Cowe and J.W. McNicol, The use of principal component in analysis of near-infrared spectra, *Applied Spectroscopy*, 39 (1985) 257–266.

21 S. Wold, A. Ruhe, H. Wold and W.J. Dunn III, The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses, *SIAM Journal of Scientific and Statistical Computations*, 5 (1984) 735–743.

22 R. Manne, Analysis of two partial-least-squares algorithms for multivariate calibration, *Chemometrics and Intelligent Laboratory Systems*, 2 (1987) 187–197.

23 S. de Jong, Principal covariates regression. Part 1. Theory, *Chemometrics and Intelligent Laboratory Systems*, 14 (1992) 155–164.

24 E.V. Thomas and D.M. Haaland, Comparison of multivariate calibration methods for quantitative spectral analysis, *Analytical Chemistry*, 62 (1990) 1091–1099.

25 B.E.H. Saxberg and B.R. Kowalski, Generalized standard addition method, *Analytical Chemistry*, 51 (1979) 1031–1038.

26 A. Lorber and Z. Goldbart, Generalized internal reference method for simultaneous multichannel analysis, *Analytical Chemistry*, 56 (1984) 37–43.

27 J.R. Long, V.G. Gregoriou and P.J. Gemperline, Spectroscopic calibration and quantitation using artificial neural networks, *Analytical Chemistry*, 62 (1990) 1791–1797.

28 S. Clementi, G. Cruciani, C. Curt and B. Skagerberg, PLS

response surface methodology: the CARSO procedure, *Journal of Chemometrics*, 3 (1989) 499–510.

29 O.M. Kvalheim, Model building in chemistry, *Analytica Chimica Acta*, 223 (1989) 53–73.

30 N.B. Vogt and K. Kolset, Composition activity relationships – CARE. Part III. Polynomial principal component regression and response surface analysis of mutagenicity in air samples, *Chemometrics and Intelligent Laboratory Systems*, 6 (1989) 221–237.

31 T. Næs, T. Isaksson and B.R. Kowalski, Locally weighted regression and scatter correction for near-infrared reflectance data, *Analytical Chemistry*, 62 (1990) 664–673.

32 R. Danielson and G. Malmquist, Multi-dimensional simplex interpolation. An approach to local models for prediction, *Chemometrics and Intelligent Laboratory Systems*, 14 (1992) 115–128.

33 T.V. Karstang and R. Manne, Optimized scaling. A novel approach to linear calibration with closed data sets. *Chemometrics and Intelligent Laboratory Systems*, 14 (1992) 165–173.

34 S. Wold, Nonlinear partial least squares modeling II. Spline inner relation, *Chemometrics and Intelligent Laboratory Systems*, 14 (1992) 71–84.

35 V. Taavitsainen and P. Korhonen, Non-linear data analysis with latent variables, *Chemometrics and Intelligent Laboratory Systems*, 14 (1992) 185–194.

36 A. Lorber, Error propagation and figures of merit for quantification by solving matrix equations, *Analytical Chemistry*, 58 (1986) 1167–1172.

37 J.H. Kalivas, Variance-decomposition of pure-component spectra as a measurement of selectivity, *Journal of Chemometrics*, 3 (1989) 409–418.

38 P.J. Gemperline, Mixture analysis using factor analysis, I. Calibration and quantitation, *Journal of Chemometrics*, 3 (1989) 549–568.

39 S.C. Rutan and S.D. Brown, Adaptive Kalman filtering used to compensate for model errors in multicomponent methods, *Analytica Chimica Acta*, 160 (1984) 99–108.

40 D.W. Osten and B.R. Kowalski, Background detection and correction in multicomponent analysis, *Analytical Chemistry*, 57 (1985) 908–917.

41 Y.-Z. Liang, Y. Xie and R. Yu, Multivariate spectral analysis with background interference detection and correction, *Chinese Science Bulletin (English Edition)*, 34 (1989) 1533–1538.

42 E. Sanchez and B.R. Kowalski, Generalized rank annihilation factor analysis, *Analytical Chemistry*, 58 (1986) 496–501.

43 E. Sanchez and B.R. Kowalski, Tensor calibration: II second-order calibration, *Journal of Chemometrics*, 2 (1988) 265–272.

44 B.E. Wilson, E. Sanchez and B.R. Kowalski, An improved algorithm for the generalized rank annihilation method, *Journal of Chemometrics*, 3 (1989) 493–498.

45 B. Vandeginste, R. Essers, T. Bosman, J. Reijnen and G. Kateman, Three-component curve resolution in liquid chromatography with multiwavelength diode array detection, *Analytical Chemistry*, 57 (1985) 971–985.

46 M. Kubista, A new method for the analysis of correlated data using Procrustes rotation which is suitable for spectral analysis, *Chemometrics and Intelligent Laboratory Systems*, 7 (1990) 273–279.

47 O.M. Kvalheim and Y.-Z. Liang, Heuristic evolving latent projections – resolving two-way multicomponent data. Part 1. Selectivity, latent-projective graph, datascope, local rank, and unique resolution, *Analytical Chemistry*, 64 (1992) 936–946.

48 Y.-Z. Liang, O.M. Kvalheim, H.R. Keller, D.L. Massart, P. Kiechle and F. Erni, Heuristic evolving latent projections – resolving two-way multicomponent data. Part 2. Detection and resolution of minor components, *Analytical Chemistry*, 64 (1992) 946–953.

49 N. Ohta, Estimating absorption bands of component dyes by means of principal component analysis, *Analytical Chemistry*, 45 (1972) 553–557.

50 A. Meister, Estimation of component spectra based on principal component analysis, *Analytica Chimica Acta*, 161 (1984) 149–161.

51 S. Kawata, H. Komeda, K. Sasaki and S. Minami, Advanced algorithm for determining component spectra based on principal component analysis, *Applied Spectroscopy*, 39 (1985) 610–614.

52 B. Vandeginste, R. Essers, T. Bosman, J. Reijnen and G. Kateman, Three-component curve resolution in liquid chromatography with multiwavelength diode array detection, *Analytical Chemistry*, 57 (1985) 971–985.

53 O.S. Borgen and B.R. Kowalski, An extension of the multivariate component-resolution method to three components, *Analytica Chimica Acta*, 174 (1985) 1–26.

54 O.S. Borgen, N. Davidson, Z. Mingyang and O. Øyen, The multivariate N-component resolution problem with minimum assumptions, *Mikrochimica Acta*, 2 (1986) 63–73.

55 E.J. Karjalainen and U.P. Karjalainen, Mathematical chromatography – resolution of overlapping spectra in GC-MS, in F.H. Roger, P. Grönroos, R. Tervo-Pellikka and R. O'Moore (Editors), *Medical Informatics Europe 85*, Proceedings, Helsinki, Finland, Springer-Verlag, Berlin, 1985, pp. 572–578.

56 E.J. Karjalainen, Isolation of pure spectra in GC/MS by mathematical chromatography: Entropy considerations, in H.L.C. Meuzelaar (Editor), *Computer-Enhanced Analytical Spectroscopy*, Vol. 2, Plenum Press, New York, 1989, pp. 49–70.

57 E.J. Karjalainen, The spectrum reconstruction problem. Use of alternating regression for unexpected spectral components in two-dimensional spectroscopies, *Chemometrics and Intelligent Laboratory Systems*, 7 (1989) 31–38.

58 K.J. Schostack and E.R. Malinowski, Prefered set selection by iterative key set factor analysis, *Chemometrics and Intelligent Laboratory Systems*, 6 (1989) 21–29.

59 W. Windig, J. Haverkamp and P.G. Kistemaker, Interpretation of sets of pyrolysis mass spectra by discriminant analysis and graphical rotation, *Analytical Chemistry*, 55 (1983) 81–88.

60 H.H. Harman, *Modern Factor Analysis*, University of Chicago Press, Chicago, IL, 2nd edn., 1967, pp. 269–270.

61 W. Windig and H.L.C. Meuzelaar, Nonsupervised numerical component extraction from pyrolysis mass spectra of complex mixtures, *Analytical Chemistry*, 56 (1984) 2297–2303.

62 W. Windig, J.L. Lippert, M.J. Robbins, K.R. Kresinske, J.P. Twist and A.P. Snyder, Interactive self-modeling multivariate analysis, *Chemometrics and Intelligent Laboratory Systems*, 9 (1990) 7–30.

63 W. Windig, C.E. Heckler, F.A. Agblevor and R.J. Evans, Self-modeling mixture analysis of categorized pyrolysis mass spectral data with the SIMPLISMA approach, *Chemometrics and Intelligent Laboratory Systems*, 14 (1992) 195–207.

64 O.M. Kvalheim, Latent-structure decomposition (projections) of multivariate data, *Chemometrics and Intelligent Laboratory Systems*, 4 (1987) 283–290.

65 P.J. Gemperline, A priori estimates of the elution profiles of pure components of overlapped liquid chromatography peaks using target factor analysis, *Journal of Chemical Information and Computer Science*, 24 (1984) 206–212.

66 B.G.M. Vandeginste, T. Esser, T. Bosman, J. Reijnen and G. Kateman, Three-component curve resolution in liquid chromatography with multiwavelength diode array detection, *Analytical Chemistry*, 57 (1985) 971–985.

67 B.G.M. Vandeginste, F. Leyten, M. Gerritsen, J.W. Noore, G. Kateman and J. Frank, Evaluation of curve resolution and iterative target transformation factor analysis in quantitative analysis by liquid chromatography, *Journal of Chemometrics*, 1 (1987) 57–71.

68 M. Maeder and A.D. Zuberbuhler, The resolution of overlapping chromatographic peaks by evolving factor analysis, *Analytica Chimica Acta*, 181 (1986) 287–291.

69 M. Maeder, Evolving factor analysis for the resolution of overlapping chromatographic peaks, *Analytical Chemistry*, 59 (1987) 527–530.

70 H.R. Keller and D.L. Massart, Peak purity control in liquid chromatography with photodiode-array detection by a fixed size moving window evolving factor analysis, *Analytica Chimica Acta*, 246 (1991) 279–290.

71 Y.-Z. Liang, O.M. Kvalheim, A. Rahmani and R.G. Brereton, Resolution of strongly overlapping two-way multicomponent data by means of heuristic evolving latent projections, *Journal of Chemometrics*, in press.

72 K.J. Schostack and E.R. Malinowski, Theory of evolutionary factor analysis for resolution of ternary mixtures, *Chemometrics and Intelligent Laboratory Systems*, 8 (1990) 121–141.

73 K.J. Schostack and E.R. Malinowski, Theory of evolutionary factor analysis for resolution of quaternary-plus mixtures, *Chemometrics and Intelligent Laboratory Systems*, 10 (1991) 303–324.

74 E. Sanchez and B.R. Kowalski, Tensorial resolution: a direct trilinear decomposition, *Journal of Chemometrics*, 4 (1990) 29–46.

75 D.S. Burdick, X.M. Tu, L.B. McGown and D.W. Millican, Resolution of multicomponent fluorescent mixtures by analysis of excitation-emission–frequency array, *Journal of Chemometrics*, 4 (1990) 15–28.

76 A. Murray, Analytical chemistry: the science of chemical measurements (Editorial), *Analytical Chemistry*, 63 (1991) 271A.