# Strategy To Obtain Accurate Analytical Solutions in Second-Order Multivariate Calibration with Curve Resolution Methods

Mahdiyeh Ghaffari,[†] Alejandro C. Olivieri,[‡] and Hamid Abdollahi[†,*]

[†]Department of Chemistry, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137-66731, Iran

[‡]Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Instituto de Química de Rosario (IQUIR-CONICET), Suipacha 531, Rosario S2002LRK, Argentina

ABSTRACT: A novel procedure is described for processing the second-order data matrices with multivariate curve resolution-alternating least-squares; while the data set is nontrilinear and severe profile overlapping occurs in the instrumental data modes. The area of feasible solutions can be reduced to a unique solution by including/considering the area correlation constraint, besides the traditional constraints (i.e., non-negativity, unimodality, species correspondence, etc.). The latter is implemented not only for the unknown samples but also for all calibration samples, regardless of their interferent content. The area of correlation constraint was specially designed to remove rotational ambiguity in the chemical data sets when information about calibration samples is at hand. In this contribution a comprehensive strategy is developed to uniquely unravel nontrilinear data sets or data sets with severely overlapped profiles in the instrumental data modes. The approach is illustrated with simulated and experimental data sets. Borgen plots are employed to adequately visualize the extent of rotational ambiguity under non-negativity constraint.
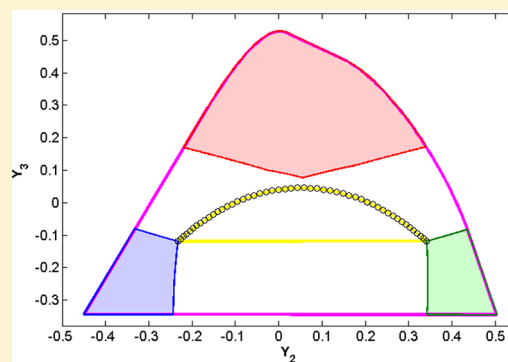
Multivariate curve resolution is a promising tool which has the ability of solving the mixture analysis problem, providing chemically meaningful pure component contributions from experimental data sets. Multivariate curve resolution—alternating least-squares (MCR-ALS) is an iterative algorithm that solves the bilinear model based on Bouguer−Beer−Lambert's law. The output consists of pure concentration and spectral profiles resulting from suitable alternating least-squares optimization subjected to different constraints. The latter is imposed based on the chemical knowledge of the studied systems.[1−4]

When data sets are analyzed by MCR-ALS, the result might be challenging due to lack of unique solutions, which is an intrinsic characteristic of bilinear matrix decompositions if incomplete information is available about the system.[5] Rotational ambiguities are an important source of uncertainty in MCR results and probably the most difficult problem to solve. Nonunique MCR solutions derived from rotational ambiguity are less reliable and difficult to evaluate. The effect of rotational ambiguity on the accuracy of quantitative results obtained from soft-modeling methods has been investigated in detail, even in the presence of constraints. The resulting uncertainty may be dramatically large in certain cases, e.g., when extensive or complete profile overlapping occurs in one of the data modes.[6]

Although nonuniqueness is omnipresent in MCR methods, it can be alleviated and in some cases completely avoided by means of a judicious use of the data structure and well-suited constraints. For example, imposing different constraints such as non-negativity, unimodality, selectivity,, and hard modeling

may dramatically decrease the extent of rotational ambiguity by adding more information to the MCR analysis of the system under study.[7−12] Area correlation is a recently developed additional constraint which includes pseudounivariate local regressions using the area of resolved profiles against reference concentration values during the ALS optimization.[13]

In the usual extended MCR-ALS analysis of multiple data sets with an appropriate design of the simultaneously analyzed experiments, including the use of proper information as constraints, the solutions can be significantly improved through/by reducing or even eliminating rotational ambiguities and achieving uniqueness.

It is important to notice that a careful design of the augmented data set, should include, if possible, samples containing single pure analytes or pure interferents. This might allow the decomposition to reach unique solutions using the species correspondence constraint, also called the sample selectivity constraint, which forces resolved concentration profiles to have zero area for the known absent species in certain samples as a consequence of the well-known Manne theorems.[14−16]

Finally, the trilinearity constraint can guarantee accurate unique profiles for constituents with trilinear structure.[17−21] To fulfill the trilinearity condition, the augmented data matrix should contain some similar factors in its submatrices, which

should not be affected by experimental conditions; so that they share a common shape. Only their areas (and vertical heights) should proportionately change according to the constituent concentration. It is a very strong constraint that enforces the decompositions to give unique solutions under mild conditions. However, trilinearity is seldom achieved in chromatographic experiments due to run-to-run differences in peak shapes and positions of pure concentration profiles, which can commonly be found in practice. As a consequence, nontrilinear data sets might always be subjected to rotational ambiguity to some extent generating uncertainty in analyte prediction. The problem is especially severe in cases in which selectivity is compromised in the direction of the augmented data, due to extensive or total profile overlapping among analyte and interferents.

In the present report, the possibility of achieving uniqueness in nontrilinear data sets has been studied by applying an area correlation constraint and including proper calibration samples in the augmented data matrix, even in cases with total profile overlapping in the augmented direction. The proper calibration samples should contain analyte(s) at known nominal concentrations and all interferents that might be expected in the following unknown samples. It has already been shown that in the absence of these samples, MCR is unable to generate uniqueness, especially under severe or total profile overlapping in the concentration profiles. To indicate how an area correlation constraint can lead to uniqueness of MCR results in the presence of proper calibration samples, the extent of rotational ambiguity was evaluated by analysis of different simulated and experimental data sets. The latter involves the quantification of ciprofloxacin in a second-order luminescence excitation-time decay data of human serum.

■ THEORETICAL BACKGROUND

Multivariate Curve Resolution-Alternating Least-Squares (MCR-ALS). In the extended version of MCR-ALS, data matrices can be augmented row-wise, column-wise, or row- and column-wise to form a multiset structure by appending individual matrices in a proper direction and respectively having/keeping the same number of rows, columns, or both. Individual data matrices should share information with the remaining ones. A column-wise augmented matrix $\mathbf{D}_{aug}$ can be written as $[\mathbf{D}_1; \mathbf{D}_2; \mathbf{D}_3; ...; \mathbf{D}_K]$, where the semicolon ';' MATLAB notation is used to indicate that the different data matrices are appended column-wise and k indicates the total number of samples.

From the mathematical point of view, $\mathbf{D}_{aug}$ can be described by a bilinear MCR model, a factor decomposition which can be written as

$$\mathbf{D}_{aug} = \mathbf{C}_{aug}\mathbf{S}_{aug}^{\mathrm{T}} + \mathbf{E}_{aug} = [\mathbf{D}_1; \mathbf{D}_2; \mathbf{D}_3; ...; \mathbf{D}_K]$$
$$= [\mathbf{C}_1; \mathbf{C}_2; \mathbf{C}_3; ...; \mathbf{C}_K]\mathbf{S}_{aug}^{\mathrm{T}} + [\mathbf{E}_1; \mathbf{E}_2; \mathbf{E}_3; ...; \mathbf{E}_K]$$
$$(1)$$

where $\mathbf{C}_{aug}$ (augmented concentration profiles) and $\mathbf{S}_{aug}$ (spectra) are the factor matrices obtained by the bilinear decomposition of the experimental data matrix $\mathbf{D}_{aug}$ and $\mathbf{E}_{aug}$ collects the model errors. This decomposition is performed for a number of components which are contributing to the observed data variance in $\mathbf{D}_{aug}$. Alternating least-squares (ALS) optimization is a possible way to solve eq 1. The main advantage of this algorithm is the amount of information that can be included in the optimization process and the ability for

working with either single data matrices or multiset data structures. Drawbacks include permutation, intensity, and rotational ambiguities. While the first two can be easily resolved, rotational ambiguity is the most dangerous one, because it may eventually lead to inaccurate analytical results. Including chemical information in the decomposition process, via the above-mentioned constraints, can significantly reduce or even eliminate the extent of rotational ambiguity.

Usually MCR-ALS optimization starts with the initial estimates of either concentration or spectral profiles. Consequently, different methods can be used to find suitable initial estimates to start the MCR-ALS. The so-called purest variables method was used to calculate the initial spectral estimates.

Borgen Plots. Borgen plots[22,23] can be considered as a microscope to observe the details of abstract space geometry in a three-component system. The subspace of the pure components in a three-component system is three-dimensional. However, the geometry of the bilinear model in this system can be shown in a 2D plane using proper normalization. In a three-component system, the shape of the area of feasible solutions which fulfill the non-negativity constraint and construct the whole data matrix is a triangle enclosing the inner polygon. Every vertex of triangles belongs to one pure component and locates somewhere between the inner and outer polygons or on the edges of the outer polygon. In the reduced abstract space, the projected rows or columns of the data span the special parts of space around the origin. This section is enclosed by a convex polygon named inner. In addition, the outer polygon is the boundary between positive and negative parts of the abstract space.

Figure 1 shows an arbitrary Borgen plot of typical data in U-space for better understanding. After proper normalization of
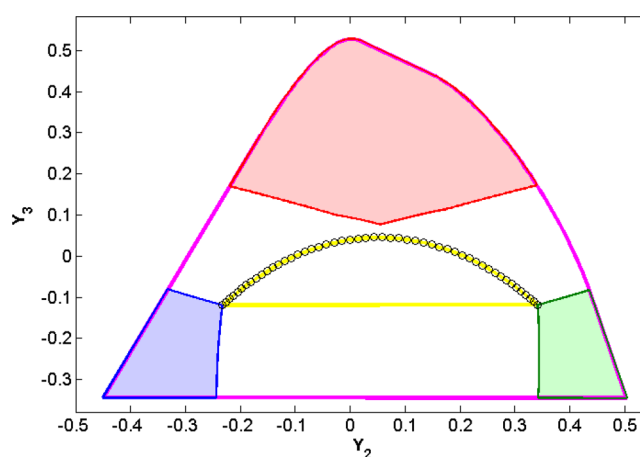

Figure 1. Visualization of the geometry of abstract space with feasible regions.

the columns of the three-component data matrix, the original three-dimensional problem is reduced and visualized in a two-dimensional abstract space; this is the reason why only two principal components are required for plotting. Singular value decomposition (SVD) can be employed to define the basis row (**V**) and column vectors (**U**) of a data matrix:

$$\mathbf{D} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} + \mathbf{E}' = \mathbf{X}\mathbf{V}^{\mathrm{T}} + \mathbf{E}' = \mathbf{U}\mathbf{Y}^{\mathrm{T}} + \mathbf{E}' \qquad (2)$$

where **X** and **Y** are coordinates of the rows and columns of **D** in its row and column space, and **E**′ is the matrix of residuals or
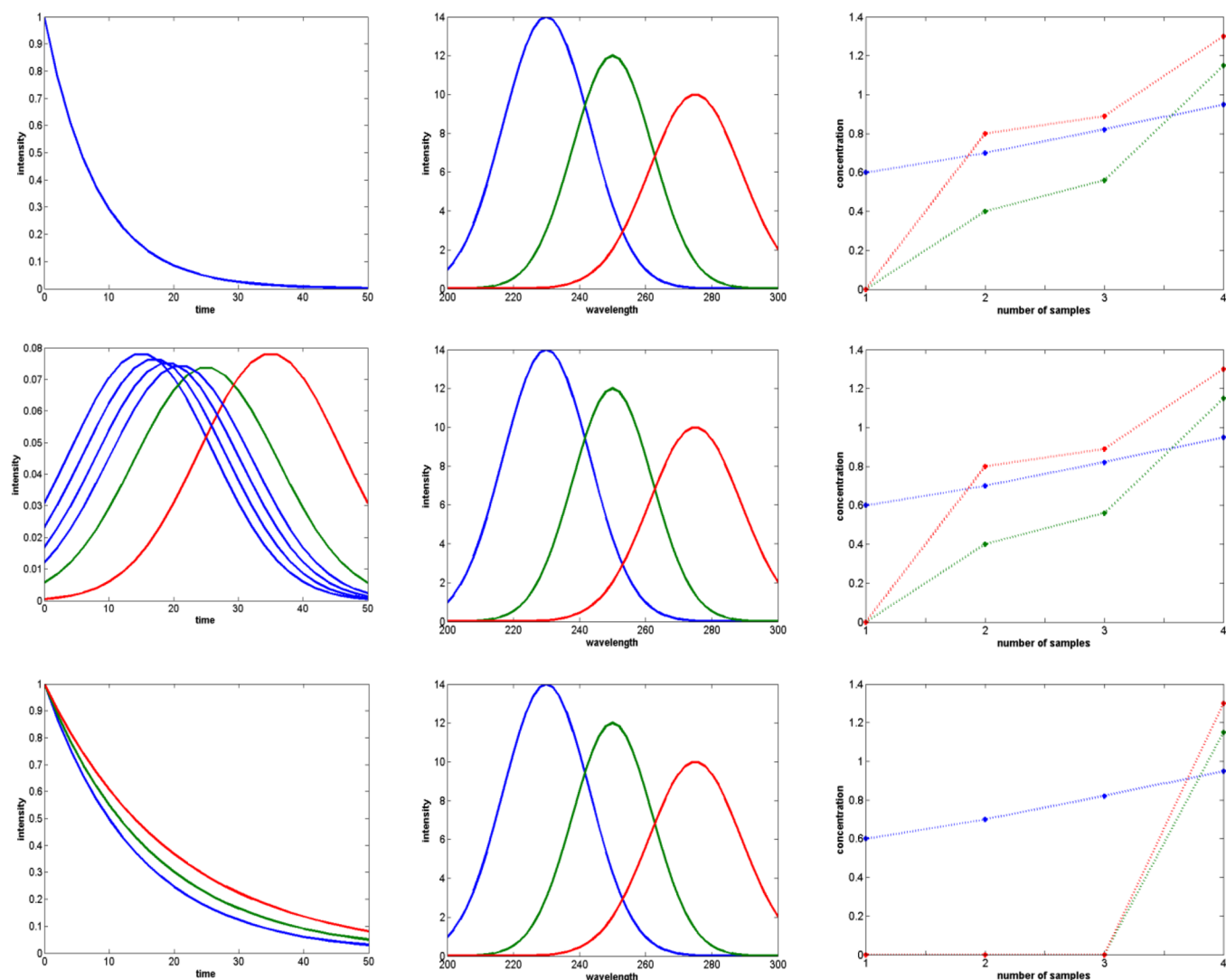
Figure 2. Summary of three different simulated cases. Each row represents a designed case (the first, second, and third row correspond to three simulated cases, simultaneously). Each panel represents real profiles. From left to right: pure concentration profiles, pure spectral profiles, and initial concentration of three components. The analyte is specified by blue and the interferents by green and red.

nonmodeled parts of **D**. A common procedure for normalization in a Borgen plot of U-space is to force the elements of the first eigenvector to be one.

The whole abstract space is the area of feasible solutions (AFS) for all components in the absence of constraints. However, considering the non-negative property of a data set, the abstract space is delimited by the outer polygon indicated in magenta in Figure 1. The coordinates of the abstract representation of the data columns are presented as black circles. Three sections shown in blue, green, and red are the three feasible regions which belong to the three components. In addition, the inner polygon is shown in yellow in Figure 1.

◼ USING/APPLYING/IMPOSING AREA CORRELATION CONSTRAINT IN THE PRESENCE OF PROPER CALIBRATION SAMPLES

In the extended MCR-ALS analysis of second-order data, relative quantitative estimations of a constituent in the different simultaneously analyzed samples (different data matrices) can be easily derived from the relative areas of the concentration profiles of this resolved component. A

calibration model can be built in this case if the concentration of this constituent (analyte) is known in some samples. It can usually be done for the pure calibration samples, but here we have proposed to extend the concept to proper calibration samples by expecting interferents in future/following samples.

To apply the area correlation constraint, the integrated concentration subprofiles for the analyte are regressed to known analyte concentrations at each iteration step of the ALS process, by means of the following linear model:

$$a_{\text{known},k} = b_0 + b_1 c_k \tag{3}$$

where $a_{\text{known},k}$ represents the area of the analyte concentration profile in the kth calibration sample, c_k is the corresponding nominal analyte concentration and b_0 and b_1 are the intercept and slope of the regression line, respectively. The linear model parameters b_0 and b_1 are estimated for k known calibration samples and employed to rescale the elements of the analyte profiles in all the calibration samples; that is, the vectors $\mathbf{c}_{\text{known},k}$ are given by
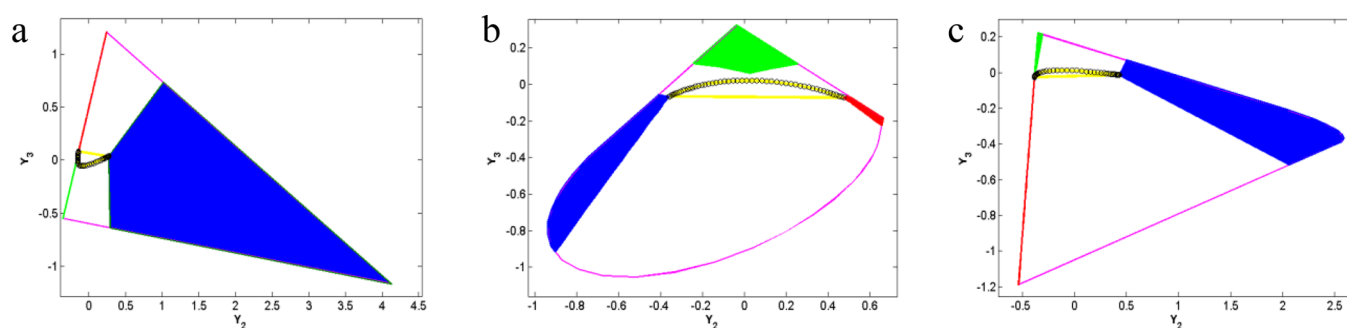
Figure 3. a, b, and c are the Borgen plots in U-space of three different simulated cases, respectively. The AFS of the analyte is specified in blue and the interferents in green and red.

$$\mathbf{c}_{\text{known},k,\text{new}} = \mathbf{c}_{\text{known},k} \frac{b_0 + b_1 c_k}{a_{\text{known},k}} \tag{4}$$

The vector $\mathbf{c}_{\text{known},k,\text{new}}$ is the rescaled analyte concentration profile in the kth calibration sample. For the unknown sample, an expression analogous to eq 4 applies: the analyte concentration profile in the unknown sample is also scaled:

$$\mathbf{c}_{\text{unknown},\text{new}} = \mathbf{c}_{\text{unknown}} \frac{b_0 + b_1 c_{\text{unk}}}{a_{\text{unk}}} \tag{5}$$

where a_{unk} is the area of the resolved analyte concentration profile in the unknown sample and the analyte concentration in the unknown mixture c_{unk} is estimated as

$$c_{\text{unk}} = (a_{\text{unk}} - b_0)/b_1 \tag{6}$$

The above procedure is implemented in every ALS step; therefore, at convergence to the optimal fit, the estimation of the concentration and spectral profiles will also be optimal as well as the analyte concentration estimates. Proper calibration samples, which have similar compositions as unknown mixtures, lead to unique solutions in trilinear and nontrilinear data sets. The number of these samples is an important factor in application of this constraint; it should be one unit less than the chemical rank of the unknown sample. As an example, for a three-component unknown mixture, two proper calibration samples with similar qualitative compositions to the unknown mixture are necessary. It is clear from the literature that analyzing full rank second-order chemical data sets with trilinear structure by second-order calibration methods can generate uniqueness. In these cases, it is not necessary to employ the proposed procedure. On the other hand, if the data are definitely nontrilinear or its trilinearity is unsure, the present approach would be beneficial for generating a unique solution.

■ DATA SIMULATIONS

In this section, three different examples are described to demonstrate the power of extended MCR to resolve diverse chemical problems through applying the area correlation constraint. The examples involve the simultaneous analysis of (1) several kinetic-spectroscopic data sets involving three parallel first-order kinetic processes with equal rate constant for all reactions, (2) different nontrilinear chromatographic runs of a complex mixture, and (3) several kinetic-spectroscopic data sets without calibration samples containing interferents.

In the first two cases, four data matrices were simulated. The first data matrix consists of only pure analyte data, whereas the remaining three ones contain the analyte and all interferents.

The analyte concentration was known not only in the first data matrix but also in the second and third, which were used as proper calibration samples. The composition of the last sample is qualitatively similar to the proper calibration samples; however the analyte concentration is regarded as unknown. The last simulation consists of only pure analyte as standard data sets and a mixture of three components as unknown sample.

Data Set 1: Kinetic-Spectrophotometric Data with Identical Concentration Profiles. The first example is generated from artificially constructed UV−vis spectra and kinetic profiles. The simulated kinetic models are three simple parallel first-order processes with equal rate constants monitored by spectrophotometric measurements (the products are supposed to be spectrally inactive). The first row of Figure 2 shows pure concentration profiles, pure spectral profiles, and initial concentration of three components from left to right, respectively. Analyte is specified by blue and interferents by green and red. In addition, the Borgen plot in the U-space of this case is shown in Figure 3a. The coordinates of the abstract representation of the data columns are presented as black circles. Three sections shown in blue, green, and red are the three feasible regions which belong to the three components (blue is used for the analyte). In addition, the inner and outer polygons are shown in yellow and magenta, respectively.

Data Set 2: Three-Component Mixture in a Chromatographic System. As a second example, four chromatographic-spectrophotometric data sets were simulated. In this case, the analyte elution profile has similar widths with different retention times in the various chromatographic runs, causing the system to be nontrilinear. The second row of Figure 2 shows pure concentration profiles, pure spectral profiles, and initial concentration of three components from left to right, respectively. Analyte is specified by blue and interferents by green and red. In addition, the Borgen plot in the U-space of this case is shown in Figure 3b. The coordinates of the abstract representation of the data columns are presented as black circles. Three sections shown in blue, green, and red are the three feasible regions which belong to the three components (blue has been used for the analyte). In addition, the inner and outer polygons are shown in yellow and magenta, respectively.

Data Set 3: Three-Component Mixture and Kinetic-Spectrophotometric Data (No proper calibration samples). The third example is generated from artificially constructed UV−vis spectra and kinetic profiles. The simulated kinetic models are three simple parallel first-order processes monitored by spectrophotometric measurements (the products
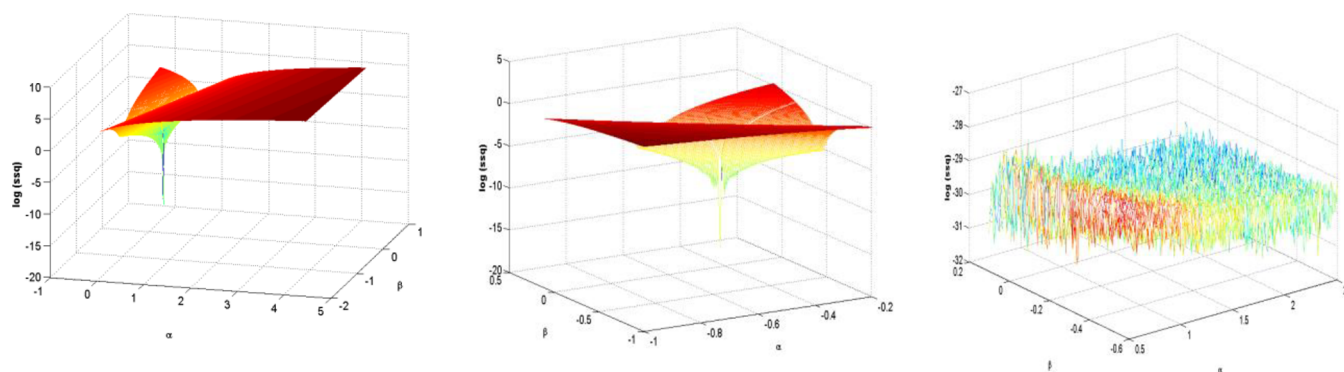
Figure 4. Values of log(ssq) vs α and β in the column space of simulated cases 1, 2, and 3, respectively.
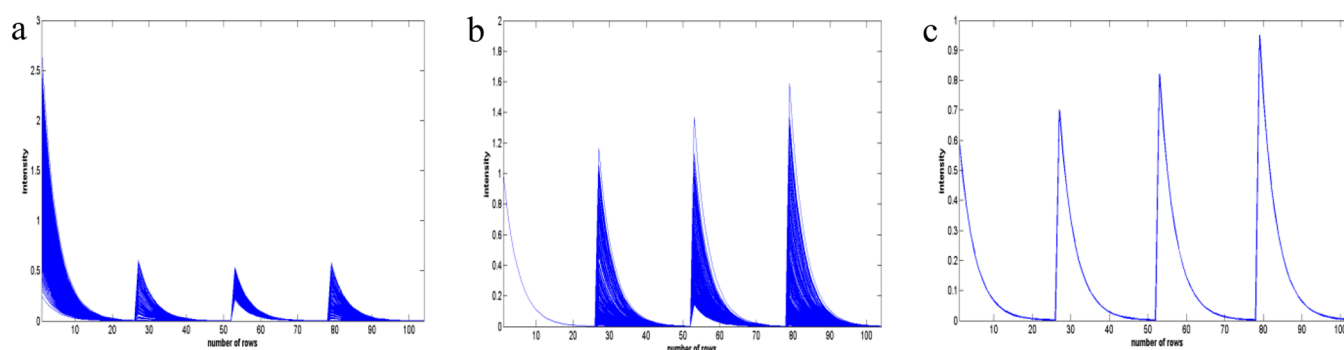

Figure 5. a, b, and c are the concentration profiles of the analyte in data set 1, under non-negativity constraint, under non-negativity and normalized to the maximum value of the standard part, and under non-negativity and area correlation constraint, respectively.

are supposed to be spectrally inactive). The third row of Figure 2 shows the pure concentration profiles, the pure spectral profiles, and the initial concentration of the three components from left to right, respectively. The analyte is specified by blue and the interferents by green and red. In addition, the Borgen plot in the U-space for this case is shown in Figure 3c. The coordinates of the abstract representation of the data columns are presented by black circles. Three sections shown in blue, green, and red are the three feasible regions which belong to the three components (blue is used for the analyte). In addition, the inner and outer polygons are shown in yellow and magenta, respectively.

■ EXPERIMENTAL DATA

This data set consists of terbium(III)-sensitized luminescence excitation−time decay matrices, measured for human serum samples spiked with the fluoroquinolone antibiotic ciprofloxacin, in the presence of the potential interferent salicylate. The analyte concentration ranges were all within the therapeutic range, i.e., 0−6 mg L^{-1} in serum, with final concentrations in the measuring cell in the order of 0.2 mg L^{-1}. For additional experimental details, see ref 24.

Software. Data analysis was run under MATLAB 8.3, and to construct the Borgen plots, the generalized Borgen plots module of the FACPACK software was used.[25]

■ VISUALIZATION

Analysis and visualization of multivariate data sets are important aspects of chemometrics. Thus, in this section we discuss the visualization process in the simulated data sets, to reach a better understanding of the effect of the area correlation constraint in the presence and absence of proper

calibration samples. In order to visualize this concept, it is necessary to prove that in the presence of proper calibration samples, a single point exists in the U-space of the data which fulfills the area correlation constraint. Consequently, in the absence of the latter it is not possible to reach uniqueness.

Each point in the U-space of the three simulated cases in the Borgen plots has two coordinates, α and β. Based on the coordinates of points and column eigenvectors, it is possible to generate the profile corresponding to the point. The values of first coordinates are ones because of normalization.

$$\boldsymbol{p} = [1\alpha\beta]\mathbf{U}' \tag{7}$$

$\boldsymbol{p}$ is the translated profile belonging to each α and β. Applying the area correlation constraint on $\boldsymbol{p}$ generates $\tilde{\boldsymbol{p}}$. Based on the explained procedure in part 3 the area correlation constraint must be applied not only on pure and mixture calibration samples but also the unknown samples will be scaled. Pure analyte standards and mixture standards can enter into the internal calibration model as long as no sample matrix effect exists. Otherwise, only mixture standards act to define the calibration model.

Consequently, it is possible to calculate the residual between $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$. If the residual shows a clear minimum, $\boldsymbol{p}$ fulfills the area correlation constraint.

$$\mathbf{r} = \boldsymbol{p} - \tilde{\boldsymbol{p}}; \ \text{ssq} = \sum r_i^2 \tag{8}$$

Figure 4 displays the value of log(ssq) vs α and β in the column space of the three simulated cases from left to right.

The two left panels of Figure 4 demonstrate that applying the area correlation constraint in the presence of two proper calibration samples in a three-component system will generate a unique solution, as is clear in cases 1 and 2. In addition, in
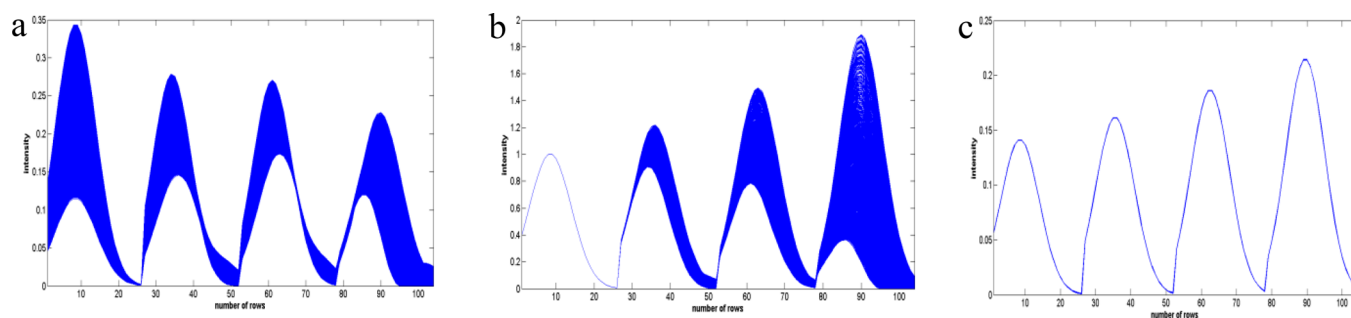
Figure 6. a, b, and c are the concentration profiles of the analyte in data set 2, under non-negativity constraint, under non-negativity and normalized to the maximum value of the standard part, and under non-negativity and area correlation constraint, respectively.
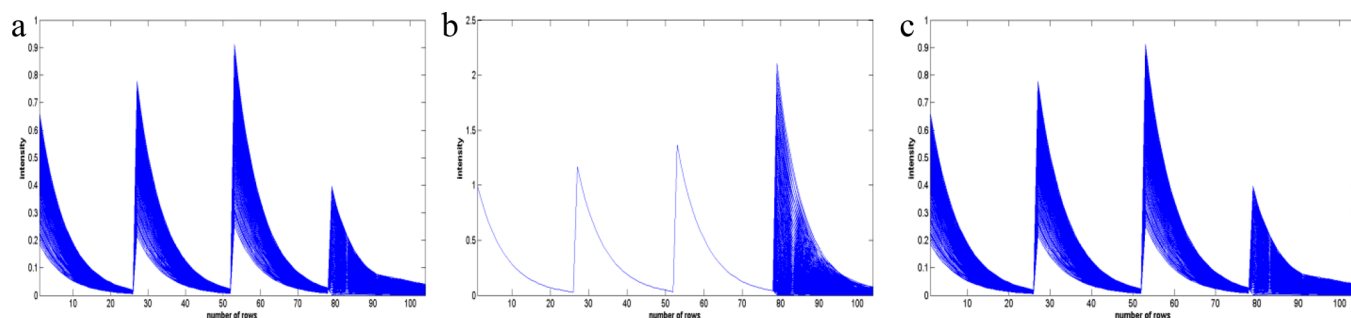

Figure 7. a, b, and c are the concentration profiles of the analyte in data set 3, under non-negativity constraint, under non-negativity and normalized to the maximum value of the first standard part, and under non-negativity and area correlation constraint, respectively.

case 3, which does not contain proper calibration samples, it is not possible to get a unique solution, because all profiles in the space fulfill the constraint as well.

■ RESULTS AND DISCUSSION

In order to investigate the effect of area correlation constraint on the accuracy of MCR-ALS analyte quantification results in the presence and absence of proper calibration samples, the simulated and experimental examples have been studied and the results are discussed below.

Data Set 1: Kinetic-Spectrophotometry. In this case, the four simulated data matrices were augmented column-wise, i.e. in the time direction. As explained above, the first matrix corresponds to a standard analyte solution with known concentration and the two next ones are corresponding to proper calibration samples, also with known analyte concentrations, but in the presence of interferents. The last sample contains interferents and analyte at an unknown concentration with the aim of its quantification in this mixture.

First, the area of feasible solutions for the analyte was computed in the concentration space of the augmented multivariate data matrix, by only applying the non-negativity constraint. Since every point in this AFS can be transformed to a concentration profile, the feasible band for the analyte profile is shown in Figure 5a. In the next step all of the analyte's concentration profiles which were calculated under non-negativity constraint were normalized to the maximum value of the standard part as illustrated in Figure 5b. It is clear that the result is nonunique, except for the pure analyte sample (the first one). To reach uniqueness, additional information should be provided during the ALS procedure. Thus, in the next step, the area correlation constraint was applied to all concentration profiles from the area of feasible solutions of the analyte under non-negativity. In all the profiles a single one occurs which

fulfills the area correlation constraint, as seen in the visualization section. The corresponding concentration profile is shown in Figure 5c, illustrating the unique resolution of the analyte profile in this case. The extent of rotational ambiguity on the accuracy of quantitative results obtained from soft-modeling methods has been investigated in detail in the presence of non-negativity and species correspondence. The resulting uncertainty was dramatically large in the case with complete profile overlapping in one of the data modes.[26] However, the proposed procedure can guarantee the accuracy of results for this case in the presence of proper calibration samples and area correlation constraint as shown in simulation.

Data Set 2: Nontrilinear Chromatography with Spectral Detection. Nontrilinearity is an intrinsic property of chromatographic-spectral data sets because total reproducibility in the concentration profile of compounds in different experimental runs can seldom be achieved. A data structure of this type was employed to show the ability of area correlation constraint in dealing with nontrilinear data sets. The above-described four simulated data matrices were augmented column-wise, i.e. in the direction of the elution time. The first sample is taken as a standard analyte solution with known concentration, and the next two are proper calibration samples with similar qualitative composition as unknown mixtures and with known analyte concentration. The last sample is a truly unknown mixture with the aim of analyte quantitation.

Similar to the previous case, the area of feasible solutions for the analyte was computed in the concentration space of the augmented multivariate data matrix, by only applying the non-negativity constraint. Since every point in this AFS can be transformed to a concentration profile, the feasible band for the analyte profile is shown in Figure 6a. In the next step, all of the concentration profiles of analyte which were calculated under non-negativity constraint were normalized to the maximum value of the standard part as illustrated in Figure

6b. It is clear that the result is nonunique except for the pure analyte sample (the first one). In the next step, the area correlation constraint was applied to all concentration profiles from the area of feasible solutions of the analyte under non-negativity. In all profiles a single one occurs which fulfills the area correlation constraint. The corresponding concentration profile is shown in Figure 6c, illustrating the unique resolution of the analyte profile in this case. Again, the range of feasible solution of analyte in concentration mode is reduced to a single point, owing to useful information added to the analysis.

Data Set 3: Three-Component Mixture and Kinetic-Spectrophotometric Data (No proper calibration). In this case, the four simulated kinetic-spectrophotometric data matrices were augmented column-wise, i.e. in the time direction. As explained above, the first three matrices correspond to a standard analyte solution with known concentration and in the absence of interferents. The last sample contains interferents, and the analyte at an unknown concentration, with the aim of its quantification in this mixture. Similar to other cases, in the first step the analyte area of feasible solutions was computed in the concentration space of the augmented multivariate data matrix, by only applying the non-negativity constraint. Since every point in this AFS of analyte is transformed to a concentration profile, the feasible band for the analyte profile is shown in Figure 7a. In the next step, all of the analyte's concentration profiles which were calculated under non-negativity constraint were normalized to the maximum value of the first standard part as indicated in Figure 7b. It is clear that the result is nonunique, except for the pure analyte samples (the first one).

In the next step, the area of correlation constraint was applied to all the solutions in the AFS of analyte under non-negativity. All profiles in the AFS fulfill the constraint as well, so the extent of rotational ambiguity did not change (Figure 7c).

Effect of Noise in the Performance of the Proposed Procedure. Analysis of a simulated noisy data set helps to provide an insight about the problem in real experimental systems. Data set 1 is used to show the efficiency of the proposed method in nonideal situations. Hence a homoscedastic noise equivalent to 0.7% of the maximum values of the signals was added to data set 1. Figure 8 displays of the values of $\log(ssq)$ vs α and β in the column space of data set 1 in the
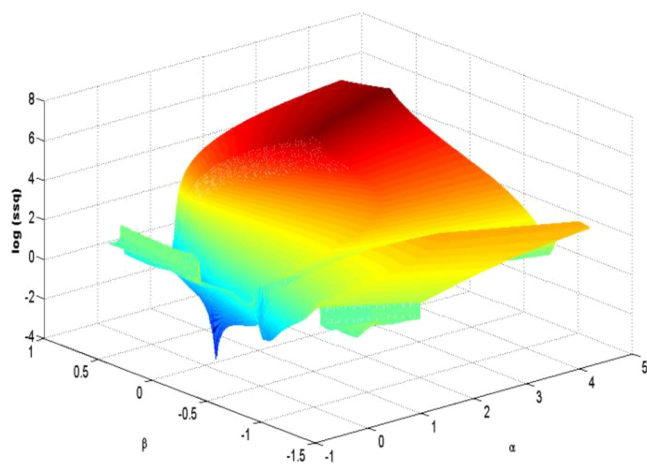
presence of noise. It is clear that a single point exists in the U-space of the data which fulfills the area correlation constraint in the presence of proper calibration samples.

Experimental Data Set. In order to discuss the ability of area correlation constraint in the presence of noise, a real example is discussed. In this case, 12 kinetic-fluorescence data matrices were augmented column-wise in the direction of decay time. Two pure standards are available, three proper calibration samples which contain interferent and the remaining seven ones were considered as unknown mixtures. Similar to the simulated cases, the analyte area of feasible solutions was first computed in the score space of the augmented data matrix, only under non-negativity. All points in the AFS were transformed to analyte concentration profiles as shown in Figure 9a. More information is available about the studied system, i.e. known concentration of analyte in the first five samples. Thus, in the next step, besides non-negativity, the area correlation constraint was applied as additional information to obtain uniqueness. Only one profile fulfill the area correlation constraint as well which is shown in Figure 9b, that illustrates the achieved uniqueness.

The specific quantitation results are displayed in Table 1 in both conditions for application of constraint. In the presence of only non-negativity, a range of analyte concentrations can be predicted. Table 1 provides the upper and lower analyte predictions in this particular case. The smaller concentration value corresponds to the lower analyte profile of its AFS, while the larger corresponds to the upper AFS profile. Large relative uncertainties are derived from these lower and upper analyte predictions, and from Table 1 its range can be estimated to be from −40% to +80%. However, in the case with both non-negativity and area correlation constraints, the predicted values are unique (Table 1). Moreover, Figure 10 shows the satisfactory regression plot of predicted ciprofloxacin concentration versus nominal values in the unknown samples under both constraints, leading to a reasonable average error of 0.03 mg L^{-1} (ca. 10% of relative error with respect to the mean calibration concentration).

■ CONCLUSIONS AND OUTLOOK

Noise free simulated data sets were analyzed by MCR-ALS optimization under proper constraints during the least-squares optimization to retrieve the final profiles, namely non-negativity in all profiles in both data modes and area correlation constraint for the analyte concentration profile. In the two first simulated data sets, the resolved analyte profiles were virtually the same as the simulated ones in the simulated data sets. Area correlation constraint and non-negativity for both concentration and spectral profiles were imposed as suitable constraints which lead to the acceptable analytical results.

The uniqueness property, which ensures the accuracy of qualitative and quantitative application of soft modeling results, has immense consequences in second-order analytical calibration. Uniqueness provides the possibility of quantitating analytes in a sample containing unexpected constituents. The main conclusion of the present study is that the area correlation constraint in the presence of proper calibration samples which have qualitatively similar composition as unknown data sets can guarantee accurate solutions in multivariate curve resolution methods even in data sets with nontrilinear structure. The composition of proper calibration samples is a key factor in using this constraint. The proper
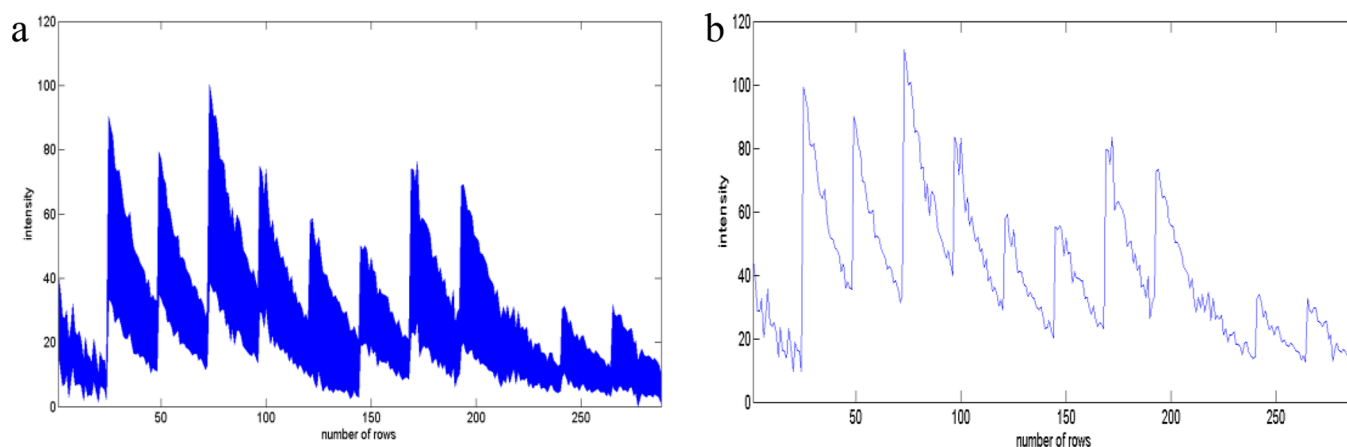

Figure 8. Values of $\log(ssq)$ vs α and β in the column space of the nonideal data set 1.

Figure 9. Concentration profiles of the analyte in the experimental data, under non-negativity constraint (a) and under non-negativity and area correlation constraint (b).

Table 1. Quantification Result of Experimental Data Set in Two Different Conditions

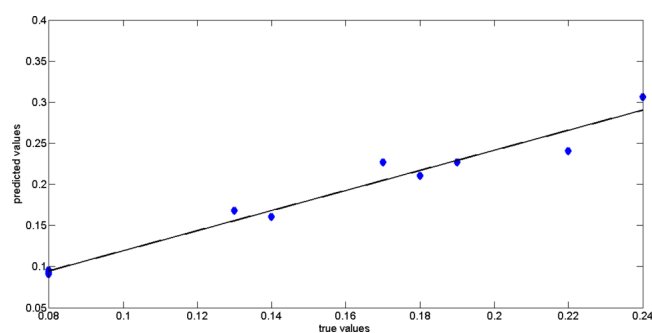| | | Predictions under constraint(s) | | |
| | | Only non-negativity | | |
Sample	Nominal	Upper	Lower	Non-negativity and area correlation
1	0.22	0.27	0.12	0.24
2	0.24	0.35	0.11	0.3
3	0.19	0.26	0.4	0.22
4	0.14	0.21	0.03	0.16
5	0.13	0.19	0.08	0.16
6	0.17	0.25	0.10	0.22
7	0.18	0.21	0.02	0.21
8	0.08	0.13	0.05	0.10
9	0.08	0.14	0.04	0.09
10	0.08	0.14	0.03	0.09

Figure 10. Regression plot of predicted ciprofloxacin concentration vs nominal values in the unknown samples under non-negativity and area correlation constraint.

calibration data sets should contain all interferents which are expected in future/later unknown samples. The number of calibration samples which should contain all interferents must be one unit less than the chemical rank of the unknown sample. The presently described procedure is a new possibility for general cases of second-order multivariate calibration data in the presence of unknown interferents or in more difficult cases.

■ **AUTHOR INFORMATION**

Corresponding Author

*E-mail: abd@iasbs.ac.ir.

ORCID ⊙

Alejandro C. Olivieri: 0000-0003-4276-0369

Hamid Abdollahi: 0000-0002-5994-6365

Notes

The authors declare no competing financial interest.

■ **REFERENCES**

(1) Tauler, R.; Kowalski, B.; Fleming, S. *Anal. Chem.* **1993**, *65*, 2040−2047.

(2) Tauler, R.; Barceló, D. *TrAC, Trends Anal. Chem.* **1993**, *12*, 319.

(3) Tauler, R. *Chemom. Intell. Lab. Syst.* **1995**, *30*, 133−146.

(4) Tauler, R.; Smilde, A.; Kowalski, B. *J. Chemom.* **1995**, *9*, 31−58.

(5) Golshan, A.; Abdollahi, H.; Beyramysoltan, S.; Maeder, M.; Neymeyr, K.; Rajkó, R.; Sawall, M.; Tauler, R. *Anal. Chim. Acta* **2016**, *911*, 1−13.

(6) Ahmadi, G.; Abdollahi, H. *Chemom. Intell. Lab. Syst.* **2013**, *120*, 59−70.

(7) Ahmadi, G.; Tauler, R.; Abdollahi, H. *Chemom. Intell. Lab. Syst.* **2015**, *142*, 143−150.

(8) Rajkó, R.; Abdollahi, H.; Beyramysoltan, S.; Omidikia, N. *Anal. Chim. Acta* **2015**, *855*, 21−33.

(9) Beyramysoltan, S.; Abdollahi, H.; Rajkó, R. *Anal. Chim. Acta* **2014**, *827*, 1−14.

(10) Beyramysoltan, S.; Rajkó, R.; Abdollahi, H. *Anal. Chim. Acta* **2013**, *791*, 25−35.

(11) Olivieri, A. C.; Tauler, R. *J. Chemom.* **2017**, *31*, e2875.

(12) Omidikia, N.; Abdollahi, H.; Kompany-Zareh, M.; Rajkó, R. *Anal. Chim. Acta* **2016**, *939*, 42−53.

(13) de Oliveira Neves, A. C.; Tauler, R.; de Lima, K. M. G. *Anal. Chim. Acta* **2016**, *937*, 21−28.

(14) Manne, R. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 89−94.

(15) Alinaghi, M.; Rajkó, R.; Abdollahi, H. *Chemom. Intell. Lab. Syst.* **2016**, *153*, 22−32.

(16) Tavakkoli, E.; Rajkó, R.; Abdollahi, H. *Talanta* **2018**, *184*, 557−564.

(17) Olivieri, A. C.; Arancibia, J. A.; Muñoz de la Peña, A.; Duran-Meras, I.; Espinosa Mansilla, A. *Anal. Chem.* **2004**, *76*, 5657−5666.

(18) Cattell, R. B. *Psychometrika* **1944**, *9*, 267−283.

(19) Alier, M.; Felipe, M.; Hernández, I.; Tauler, R. *Anal. Bioanal. Chem.* **2011**, *399*, 2015−2029.

(20) Sanchez, E.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 496−499.

(21) Ghaffari, M.; Abdollahi, H. *Chemom. Intell. Lab. Syst.* **2018**, *177*, 17−25.

(22) Borgen, O. S.; Kowalski, B. R. *Anal. Chim. Acta* **1985**, *174*, 1−26.

(23) Rajkó, R.; István, K. *J. Chemom.* **2005**, *19*, 448−463.

(24) Lozano, V. A.; Ibañez, G. A.; Olivieri, A. C. *Anal. Chem.* **2010**, *82*, 4510−4519.

(25) Sawall, M.; Jürß, A.; Neymeyr, K.: FACPACK, 2014.

(26) Lozano, V. A.; Tauler, R.; Ibañez, G. A.; Olivieri, A. C. *Talanta* **2009**, *77*, 1715−1723.