Multivariate Curve Resolution for Signal Isolation from Fast-Scan Cyclic Voltammetric Data

Justin A. Johnson,[†] Josh H. Gray,[†] Nathan T. Rodeberg,[†] and R. Mark Wightman[*,†,‡]

[†]Department of Chemistry and [‡]Neuroscience Center and Neurobiology Curriculum, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3290, United States

[S] *Supporting Information*

ABSTRACT: The use of multivariate analysis techniques, such as principal component analysis−inverse least-squares (PCA−ILS), has become standard for signal isolation from in vivo fast-scan cyclic voltammetric (FSCV) data due to its superior noise removal and interferent-detection capabilities. However, the requirement of collecting separate training data for PCA−ILS model construction increases experimental complexity and, as such, has been the source of recent controversy. Here, we explore an alternative method, multivariate curve resolution− alternating least-squares (MCR−ALS), to circumvent this issue while retaining the advantages of multivariate analysis. As compared to PCA−ILS, which relies on explicit user definition of component number and profiles, MCR−ALS relies on the unique temporal signatures of individual chemical components for analyte-profile determination. However, due to increased model freedom, proper deployment of MCR−ALS requires careful consideration of the model parameters and the imposition of constraints on possible model solutions. As such, approaches to achieve meaningful MCR−ALS models are characterized. It is shown, through use of previously reported techniques, that MCR−ALS can produce similar results to PCA−ILS and may serve as a useful supplement or replacement to PCA−ILS for signal isolation from FSCV data.
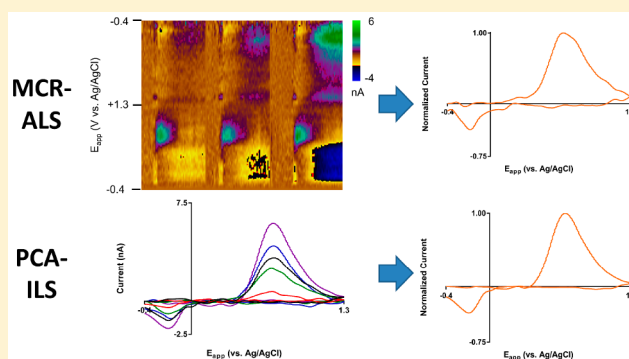
Fast-scan cyclic voltammetry (FSCV) has several advantages over other electrochemical techniques for studying in vivo extracellular neurotransmitter dynamics, particularly the selectivity afforded by analyte voltammetric profiles. Full realization of this, however, demands multivariate data analysis.[1−3] Several methods have been reported, including partial least-squares and elastic net regression.[4,5] Among these, the factor analysis-based principal component analysis−inverse least-squares regression (PCA−ILS, elsewhere referred to as principal component regression, or PCR, but written in full here for consistency in abbreviation), introduced for FSCV analysis by our lab, has been shown as a reliable approach for in vitro and in vivo analyte-signal isolation from multicomponent data.[6−8] Further, its development for FSCV analysis has resulted in implementation of model-validation procedures, giving confidence in model-generated estimates.[9,10] However, a primary drawback has been the necessity of separate training-set construction for model generation. Chemometrics and previous research suggest training sets must be generated under the experimental conditions for proper model validation and confidence in the concentration estimates.[11−14] This requirement adds to experimental complexity, and the degree to which unrepresentative training data affects model predictions has been the subject of recent debate.[15−17] Thus, a method that relaxes this requirement should be of interest to the field.

Here, we explore an alternative method, multivariate curve resolution by alternating least-squares (MCR−ALS), to resolve overlapping FSCV signals. Like PCA−ILS, the data are modeled as a linear combination of appropriately scaled component signals. However, whereas PCA−ILS model construction typically relies on user-isolated, single-component training data for spectral definition, MCR−ALS can use raw experimental data itself to define component spectral and concentration profiles, requiring only definition of the number of components. For this, the unique temporal signature of components is used for their identification and isolation. MCR−ALS has been successfully used in the analysis of data derived from a number of analytical techniques, including mass spectrometry,[18−20] spectroscopic techniques,[21−23] and slow-scan voltammetry.[24−27] Thus, this opens the possibility of circumventing the need for explicit training-set construction.

However, due to increased model freedom, important concerns must be addressed before MCR−ALS deployment. First, the aforementioned applications relied on generation of second-order data (e.g., data separated along two variables). Limited success is seen with poorly resolved signals.[28] Resolution is often achieved using separation techniques (e.g., liquid chromatography) or controlled independent-variable manipulation (e.g., concentration). For FSCV, second-order

data (i.e., current as a function of potential and time) is typically collected. However, temporal-signal separation, and thus the technique's potential, must rely on naturally occurring processes.

Second, MCR−ALS solutions are susceptible to a number of ambiguities, resulting in mathematical solutions that may not have correspondence to meaningful chemical information.[29,30] The most relevant of these are intensity and rotational ambiguities. Intensity ambiguity refers to the fact that MCR−ALS only provides the shapes, and not absolute scales, of spectral and concentration profiles. However, in the analysis of FSCV data, PCA−ILS suffers from similar ambiguities, and this problem can be addressed through normalization of obtained voltammetric profiles and their subsequent scaling by previously determined calibration factors. Rotational ambiguity, referring to the fact that the data can be fit by an infinite number of combinations of voltammetric and concentration profiles, is a more serious issue. This requires the imposition of constraints, derived from prior knowledge of meaningful solution characteristics, on the MCR−ALS fits. Commonly employed constraints include non-negativity of spectral- or concentration-profile values, peak unimodality, and hard-modeling approaches using known equations that govern the experimental system.[29,31] For instance, in the previous study of electrochemical data, parametric equations for peak-shape definition and closure (i.e., constant total species concentrations) constraints, as well as non-negativity and unimodality constraints, were used.[24] Background-subtracted FSCV data, however, is more limited in the information that can define the subset of meaningful fits. For instance, due to the relative nature of measurements, negative values can be found in spectral and concentration profiles. Further, the equations governing the observed voltammetric behavior are not sufficiently well-defined to use as strict constraints. Thus, a characterization of the subset of reported constraints that may be used is needed to ensure their sufficiency for robust MCR−ALS deployment for FSCV data analysis.[31−36]

In this study, the potential of MCR−ALS for the analysis of FSCV data is explored and compared to the performance of PCA−ILS. First, the basic implementation of the method is described. Next, the method is characterized in vitro to determine the conditions that enable successful signal isolation. This is followed by in vivo comparison of PCA−ILS and MCR−ALS. It is shown that the method, given appropriate constraints, is capable of producing similar results to PCA−ILS without the need for separate training data. Further, methods to extend the utility of the technique and for model validation (i.e., residual analysis) are explored.

EXPERIMENTAL SECTION

Instrumentation and Software. T-650 type, cylindrical carbon-fiber microelectrodes (Thornel, Amoco Corporation, Greenville, SC; pulled in glass capillaries and cut to 75−125 μm exposed lengths) were used in experimentation. After the microelectrodes were pulled, the seals of the electrodes were dipped in epoxy (EPON Resin 828, Miller-Stephenson, Danbury, Connecticut) mixed with 14% w/w *m*-phenylenedi-amine (Sigma-Aldrich, St. Louis, MO) at 80 °C, briefly washed with acetone, and heated at 100 °C (5 h) and then 150 °C (at least 12 h).

Data were acquired with a commercial interface (PCI-6052, 16 bit, National instruments, Austin, TX) with a personal home computer and analyzed using locally constructed hardware and software written in LabVIEW (HDCV, National Instruments, Austin, TX).[37] Unless otherwise noted, triangular excursions of the working electrode potential (−0.4 to 1.3 V vs Ag/AgCl) were made at a scan rate of 400 V/s and repeated at a frequency of 10 Hz. Measurements were conducted inside a grounded Faraday cage to minimize electrical noise.

Electrochemical Experiments. Flow-injection analysis experiments were performed using a syringe pump (Harvard Apparatus, Holliston, MA) operated at 0.8 mL/min using PEEK tubing (Sigma-Aldrich) connected to a pneumatically controlled six-port injection valve (Rheodyne, Rohnert Park, CA). All solutions were prepared in TRIS (2.0 mM Na_2SO_4, 1.25 mM $NaH_2PO_4 \cdot H_2O$, 140 mM NaCl, 3.25 KCl, 1.2 mM $CaCl_2 \cdot 2H_2O$, 1.2 mM $MgCl_2 \cdot 6H_2O$, and 15 mM Trizma HCl) and adjusted to pH 7.4 with NaOH as necessary.

In Vivo Measurements. Male Sprague−Dawley rats from Charles River (Wilmington, MA, USA) were housed individually on a 12/12 h light/dark cycle. Animal procedures were approved by the UNC-Chapel Hill Institutional Animal Care and Use Committee (IACUC). Surgery was performed on the animals in the manner described previously for intracranial self-stimulation (ICSS) experiments and given a minimum of 3 days of recovery prior to training.[38] Rats were trained in ICSS using a fixed-ratio 1 or fixed-interval 5 schedule.[39]

Data Analysis. Data and statistical analyses were performed in GraphPad Prism 6 (GraphPad Software Inc., La Jolla, CA), LabView (National Instruments, Austin, TX), and MATLAB (Mathwork, Natick, MA). Convergence was defined as achievement of differences in consecutive residual values of 0.1% or the performance of 200 iterations, whichever was achieved first.

THEORY

Multivariate Curve Resolution−Alternating Least Squares. The theory of PCA−ILS has been previously discussed.[16] Here, we address the general theory behind MCR−ALS. As in PCA−ILS, the "bilinear" model is used.[28,32,36] That is, each measurement (i.e., individual current measurements and entire voltammograms) is assumed to be a linear combination of the independent contributions of analytes and noise

$$\mathbf{D} = \mathbf{CS}^T + \mathbf{E} \tag{1}$$

where **D** is the $(r \times c)$ data matrix containing c spectra consisting of r individual measurements (e.g., current measurements in a voltammogram), **C** and **S** are $(r \times l)$ and $(c \times l)$ matrices containing the l pure concentration profiles and spectra, respectively, and **E** is the $(r \times c)$ error matrix. This equation is visually shown in Figure 1.

First, the model parameters and inputs, namely, the number of expected components and the initial estimates of the either the voltammograms or concentrations, must be defined.[32,40] There exist many ways to achieve the former, and a comparison of the effectiveness of a large number for LC−NMR data analysis has been reported.[41] One class of techniques relies on factor-based analysis of the experimental data, including evaluation of PCA-generated singular values (e.g., Malinowski's F-test) or other PCA-based methods (e.g., evolving-factor analysis, or EFA, and target-factor analysis).[42−48] Additionally, orthogonalization methods (e.g., orthogonal-projection approach, or OPA), which select the most dissimilar spectra from the data, may be used, and have been shown to perform favorably relative to Malinowski's F-test in HPLC-DAD data
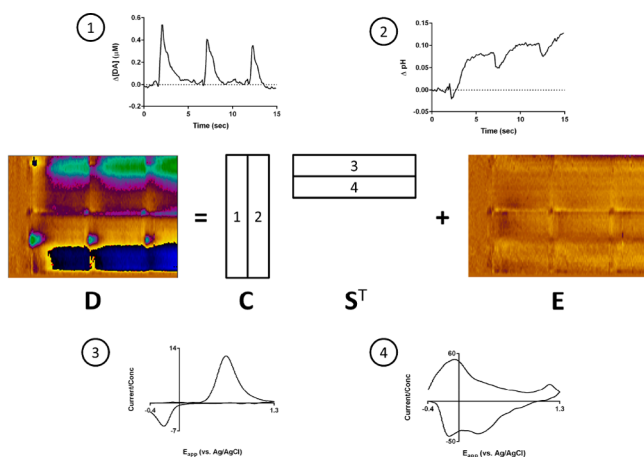
Figure 1. Graphical representation of the bilinear calibration model (eq 1) with background-subtracted FSCV data. Above are shown the dopamine (1) and pH (2) concentration traces. Below are shown the dopamine (3) and pH (4) voltammograms.

analysis.[49,50] Alternatively, this can be set through a priori knowledge or multiple fits with varying number of components to yield meaningful solutions.[31,32,40,51,52] For initial estimates of either the voltammetric or concentration profiles, many of the same methods apply. Otherwise, expected concentration or spectral profiles or even the data themselves may be used as inputs.[36,40] Here, Malinowski's F-test, which has been used previously for FSCV, and the EFA and OPA techniques are explored, the latter described briefly below.

With the model defined, eq 1 is solved using the alternating least-squares (ALS) approach. This method iterates between generating concentration or spectral estimates, given spectral or concentration estimates, respectively. In its unconstrained form, the following equations are used

$$\mathbf{C}_{est} = \mathbf{D}\mathbf{S}_{est}^{+} \tag{2}$$

$$\mathbf{S}_{est} = \mathbf{C}_{est}^{+}\mathbf{D} \tag{3}$$

where the superscript + indicates the matrix pseudoinverse. This process continues until a predefined threshold of convergence, or a set number of iteration cycles, is reached. If constraints are to be applied, this is done either through direct alteration of the obtained estimates,[40] the use of penalty functions,[36] or alternative means of regression.[53,54] Additionally, the experimental-data matrix is often pretreated by factor analysis (e.g., PCA) itself prior to fitting, as this reduces the effect of noise.[32,34] If this is done, the reconstructed data set using only significant principal components ($\mathbf{D}^*$) is used in lieu of the original data matrix ($\mathbf{D}$) in eqs 2 and 3. Such an approach is particularly advantageous for low signal-to-noise data and is used here.

Implementation of Soft Constraints Using Penalty Functions. As mentioned before, FSCV data fails to meet the criteria for the application of many commonly used constraints. However, reference data may be incorporated to help define the fits through equality constraints (i.e., spectral or concentration estimates are forced to equal to values of reference data).[55] Further, constraints can be enforced in a "soft" manner, allowing deviations from the reference values through incorporation of weighted penalty functions into the model. Here, we use the P−ALS algorithm introduced by Gemperline and Cash to realize this.[36] While illustrated in detail in reference

34, in short, "soft" equality constraints with a complete set of spectral reference data are implemented through modification of the system of equations represented by (1) by addition of the equivalent of the following

$$w\mathbf{S}_{ref} = w\mathbf{H} \tag{4}$$

where $\mathbf{S}_{ref}$ is the reference spectral matrix, $\mathbf{H}$ is diagonal matrix of ones, and w is a scalar weighting factor that determines the relative importance of this equation during the fitting procedure. Note that the symbol for the weighting factor here (w) is changed from that (λ) used by Gemperline and Cash to avoid confusion with its use to represent eigenvalues, used below and in previous work from our lab.[47] The power of this equation lies in its flexibility. Incomplete reference data (e.g., one spectrum for a multicomponent system) can be used by appropriately adjusting the $\mathbf{H}$ matrix. Further, the weighting factor w can be used to tune how strictly this constraint is enforced. Small values of w allow strong deviations from the reference spectra, while very large values force strict adherence. Of note, use of the k vectors, PCA model estimates of spectral shape from training-set analysis, as reference spectra with this P−ALS method and a high w (approaching infinity) would produce the results obtained from PCA−ILS.

Methods for Model Initialization. Here, the methods described above that were explored in this study for model definition will be briefly covered with the exception of the Malinowski's F-test, which has been explored in detail in the context of FSCV elsewhere.[47]
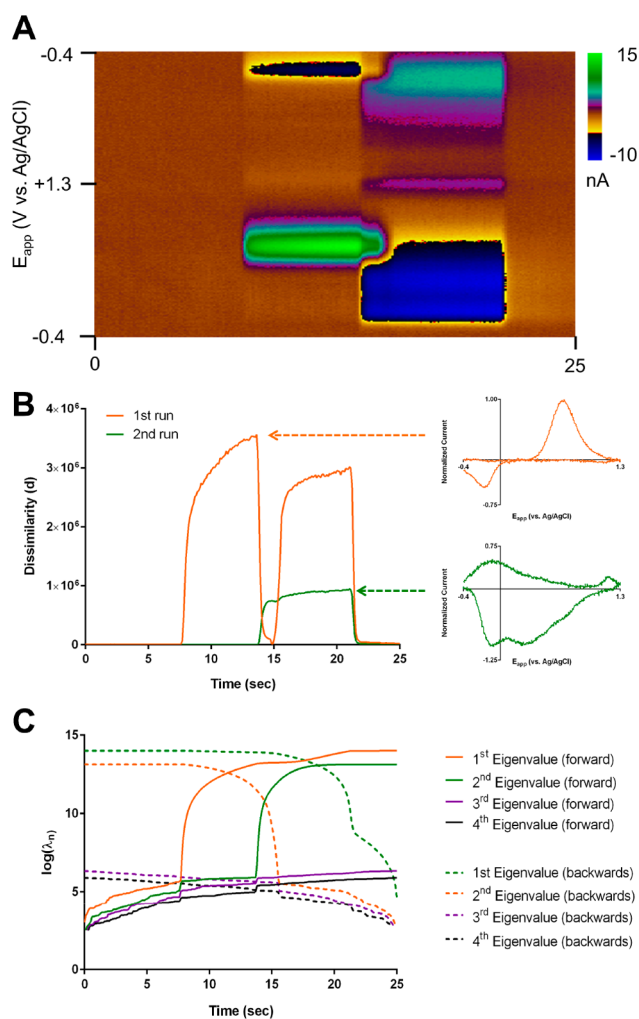
Orthogonal Projection Approach. The orthogonal-projection approach (OPA) relies on the iterative determination of most dissimilar spectra in the data. For each iteration, every spectrum ($\mathbf{s}_i$) is compared to a normalized reference spectra set ($\mathbf{s}_{ref}$), first defined as only the mean data spectrum ($\bar{s}$). The dissimilarity (d_i) is calculated as the determinant of the square matrix formed by the product of a matrix $\mathbf{Y}_i$, which has $\mathbf{s}_{ref}$ and $\mathbf{s}_i$ as its rows, and its transpose

$$d_i = \det(\mathbf{Y}_i\mathbf{Y}_i^{T}) \tag{5}$$

On the first iteration, the spectra ($\mathbf{s}_{ref,1}$) with the highest dissimilarity replaces $\bar{s}$ as the reference spectrum in $\mathbf{Y}_i$. Dissimilarities are calculated again; however, the most dissimilar spectrum with $\mathbf{s}_{ref,1}$ is now added to $\mathbf{Y}_i$. This process continues until a plot of the dissimilarity versus time (Figure 2B) shows no distinct peak or contains only random noise, or there is redundancy in the reference shapes.

Evolving-Factor Analysis. For evolving-factor analysis (EFA), PCA is performed on successively larger portions of the data window, typically in the forward and backward direction along the relevant variable (e.g., time). For identification of the number of components (N_C), an EFA plot (logarithm of the eigenvalues vs time) is constructed (Figure 2C) for the forward and backward analyses. In the forward direction, an increase (moving from time 1 to T) in the nth eigenvalue suggests the appearance of the nth analyte (e.g., the rise in the orange solid line at $t \approx 7$ indicates the appearance of dopamine in Figure 2C). In the backward direction, an increase (moving from time T to 1) in the nth eigenvalue is interpreted as the disappearance of the ($N_C - n + 1$)th analyte, provided the analytes appear and disappear in successive order. A rough estimate of the concentration profile can be obtained by combining the forward and backward analysis, taking the lower value of the two at any given time.

Figure 2. Evaluation of color plot with orthogonal-projection approach (OPA) and evolving-factor analysis (EFA). (A) Color plot with 8 s dopamine and pH injections, with onset separated by 6 s and dopamine appearing first. (B) Dissimilarity plots determined from (A) for the first (orange) and second (green) runs of OPA. The spectra shown to the right are those selected by OPA for a given run (i.e., the voltammograms at the time corresponding to the maximum dissimilarity value). (C) EFA plot of logarithm of eigenvalues shown for forward and backward analysis. Note the colors for the first and second eigenvalues are swapped between the forward and backward direction to aid in the interpretation detailed in the text (i.e., the same colored lines for the forward and backward direction indicate the appearance and disappearance of a given analyte).

RESULTS AND DISCUSSION

In Vitro Evaluation of Dopamine-pH (DA-pH) Mixtures with MCR−ALS. First, the utility and limitations of MCR−ALS for FSCV signal were assessed using data from in vitro flow-injection analysis of the previously studied system of dopamine and pH changes.[15,16] Note that no attempt at determining "true" concentration values will be made, as signal isolation, not quantitation, is the focus here.

It was first verified that the technique successfully isolated the signals from single-component solutions, which could be done through unconstrained MCR−ALS. Various initialization methods (current at the dopamine oxidation potential for concentration, data spectra, and an OPA-generated spectra) and use of PCA pretreatment were tried, and nearly identical solutions were obtained (data not shown). Figure 3A shows an
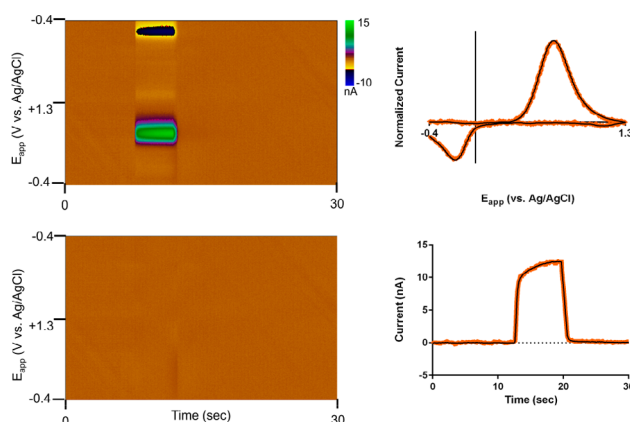
Figure 3. Results of unconstrained MCR−ALS analysis of FSCV data from a flow-injection analysis of a bolus of dopamine. Clockwise from top left: Color plot representation of the background-subtracted data, with time as the abscissa, the applied potential as the ordinate, and the current in false color; dopamine CV during injection (orange) and MCR−ALS estimate (black); current at the dopamine oxidation potential as a function of time (orange) and MCR−ALS estimate (black); color plot of the residual current after MCR−ALS analysis.

example fit to data from flow-cell analysis of a 1.0 μM dopamine bolus (initialized with the dopamine oxidation current and untreated with PCA) compared to an experimental dopamine CV and the current−time trace. It can be seen that the MCR−ALS estimates are nearly identical to these references, but have lower noise levels.

Next, the potential for MCR−ALS for separating DA-pH mixtures was evaluated. As noted before, the success of MCR−ALS is anticipated to be dependent on the temporal signal separation. Thus, simulated mixture data were created from independent injections (8 s duration) of dopamine and pH by adding these data together with differing time delays between the appearances of analyte signal. The performance of the methods described in the Theory section for model definition was evaluated using this data (Table S-1). Malinowski's F-test has been used in FSCV analysis in the context of rank selection in PCA−ILS; however, abnormally high rank estimates were obtained when large portions of the data matrix were analyzed, similar to the results reported by Vivo-Truyols et al. in the analysis of HPLC-DAD data.[50] Visual inspection of the PCs and PC-reconstructed data confirmed that these extra components consisted of random noise (data not shown). This may be due to the large number of voltammograms that carry no chemical information in the in vitro data or issues related to well-documented limitations and criticisms of the method (e.g., the small number of degrees of freedom used in calculation of the F-statistic and assumption of homoscedastic and uncorrelated noise).[47,50,56−58] Data-matrix truncation, and also confidence-level increases, decreased the number of predicted components; however, the results were inconsistent and always greater than two. While not as readily automated, the OPA and EFA approaches proved reliable indicators of the number of components, correctly predicting two components regardless of time separation. With regard to OPA, the most dissimilar spectra identified in all analyses, except with no separation, matched the pure DA and pH CVs present in the data. With no separation, only DA-pH mixture voltammograms were selected. However, the dissimilarity plots only degraded to random noise after three CVs. EFA analysis correctly indicated the time of appearance and disappearance for all cases, except

again with no temporal separation. However, in that case, the technique was able to detect subtle differences in the rate of disappearance after injection to suggest two components.

These approaches are also advantageous, as they provide information about the expected spectra (OPA) and concentration traces (OPA and EFA, the latter more reliably) that can be used for fit initialization. Since OPA is computationally inexpensive, this was chosen as the initialization method. For temporal separation greater than 8 s (complete injection separation), OPA spectral initialization and unconstrained MCR−ALS provided excellent agreement with the results expected from the individual single-component runs. With smaller separations, significant distortions appeared in the solutions. For example, at separations of 2−7 s, the concentration estimates had sudden artificial changes during the periods of signal overlap, despite having similar spectral profiles to those obtained from the single-component runs. At separations of 1.5 s or less, DA-pH mixture voltammograms were primarily obtained (as shown in Figure 4A for a 1 s separation), likely due to rotational ambiguity. Thus, constraints that could provide meaningful solutions were explored.

One approach is the use of reference data, obtained separately from the experimental data being analyzed. It is worth noting that this library approach has been used with cross-validated elastic net regression using in vitro data. However, given that this reference data is not expected to be perfectly descriptive of the experimental data, we sought to explore the use of library reference data (here, the average of 10 DA and of 10 pH CVs obtained from separate experiments using separate T-650 carbon fibers) with "soft" penalty functions (P−ALS) for imposing loose equality constraints to guide solutions. First, the effects of the "weighting" parameter w on obtained solutions for single-component runs was characterized. MCR−ALS fits to these were obtained with various w values ($w = 0$ to 8), and the sum of the squares of the residuals were determined. For both DA and pH data (Figure S-1A), a smooth transition can be seen between two different solutions, the unconstrained ($w = 0$) and that defined entirely by the average library CV (large w). As expected, fits with large w values had higher residual values, as the average CV is unrepresentative of the data. The largest changes occurred around w values of 1 for both data sets, as evidenced by the derivative plot of the sum of squares of the residual with w (Figure S-1B). The use of a pH library CV for fitting also lead to considerably higher error, due to the larger variability seen between pH CVs at carbon-fiber electrodes as compared to DA CVs.[15,47] Thus, the use of DA library reference CVs is preferred to use of pH CVs. Additionally, the approach should not be used with large weighting parameters, as these introduce considerable error into the fits. However, ideally, these constraints can be removed before the final fit is obtained.

This library P−ALS approach was used to analyze the DA-pH mixture data. During the initial fit, soft constraints (w_{pH}, $w_{DA} = 1$) were imposed on both analytes using the library CVs as reference data until convergence was achieved. Then, since the pH library is less reliable, the pH equality constraint was lifted ($w_{pH} = 0$, $w_{DA} = 1$). The solution, using the previous fit as the initialization, was again allowed to converge. Finally, the DA equality constrained was lifted (w_{pH}, $w_{DA} = 0$), and the final solution was obtained. The process was then repeated to ensure overall convergence. The approach proved successful in mitigating the issues seen with unconstrained MCR−ALS, as
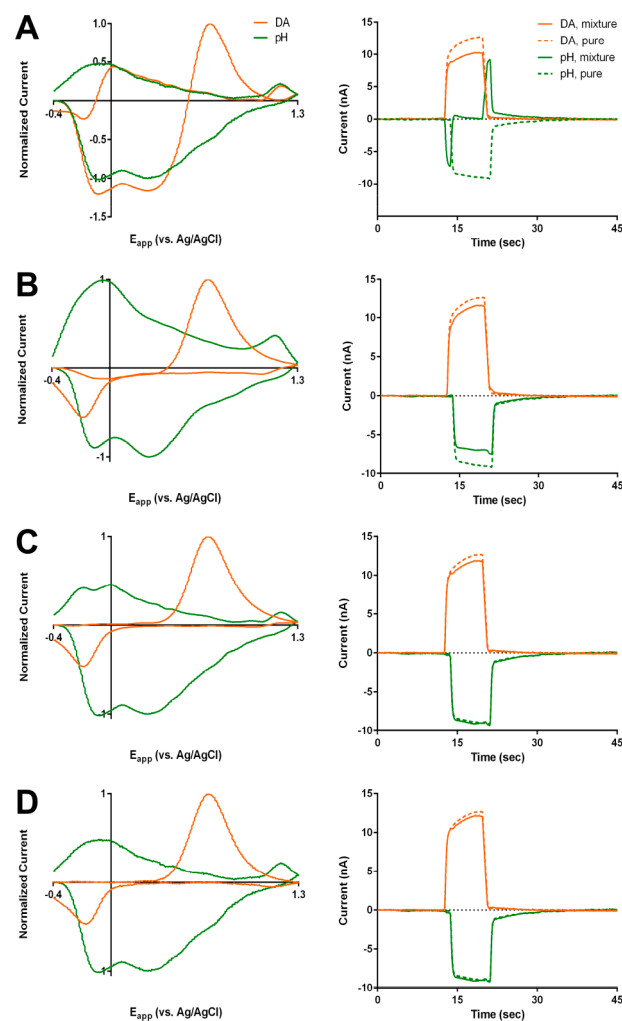

Figure 4. Successive fitting using P−ALS "soft" equality constraints for analysis of simulated in vitro DA-pH mixtures (temporal separation of 1 s). (A−D) MCR−ALS spectral (left) and (concentration) estimates for the initial unconstrained model (A), the DA/pH "soft" equality constrained model (B), the DA-only "soft" equality constrained model (C), and the final unconstrained model after successive iterations of A−C. (D). The dashed lines indicate the concentration estimates for the isolated runs for comparison.

highlighted for the 1 s separation data in Figure 4. The unconstrained solution (Figure 4A) shows distortions in both the concentration and spectral profiles. After imposition of soft constraints on both analytes (Figure 4B), the spectral and concentration profiles significantly improve; however, the sum of squares of the residual values is over 9 times higher for this constrained fit than the unconstrained fit. Removal of the pH constraint (Figure 4C) leads to a better fit at the cost of fidelity of the spectral shapes. Finally, the removal of both constraints (Figure 4D) leads to improved spectral and concentration profiles with an identical residual value as that of the original unconstrained solutions.

Evaluation of In Vivo FSCV Data. MCR−ALS was then evaluated using in vivo FSCV data obtained during intracranial self-stimulation (ICSS) sessions ($n = 25$ rats, 1 session per rat containing multiple, typically greater than 50, electrically evoked dopamine transients). These data mainly contain contributions from pH and DA, but are less well-defined and noisier than the in vitro data. However, within a given

experiment, the presence of multiple electrically evoked transients opens the possibility of using multiple data sections for model definition. These data were originally analyzed using PCA−ILS models built from training data collected during the experiment, which served as a reference point for comparison.

First, the advantages of using multiple transients for model definition were explored. Separate background-subtracted color plots were obtained from a given experimental run and concatenated together to form the data matrix for MCR−ALS analysis. Figure 5A shows an example using three separate
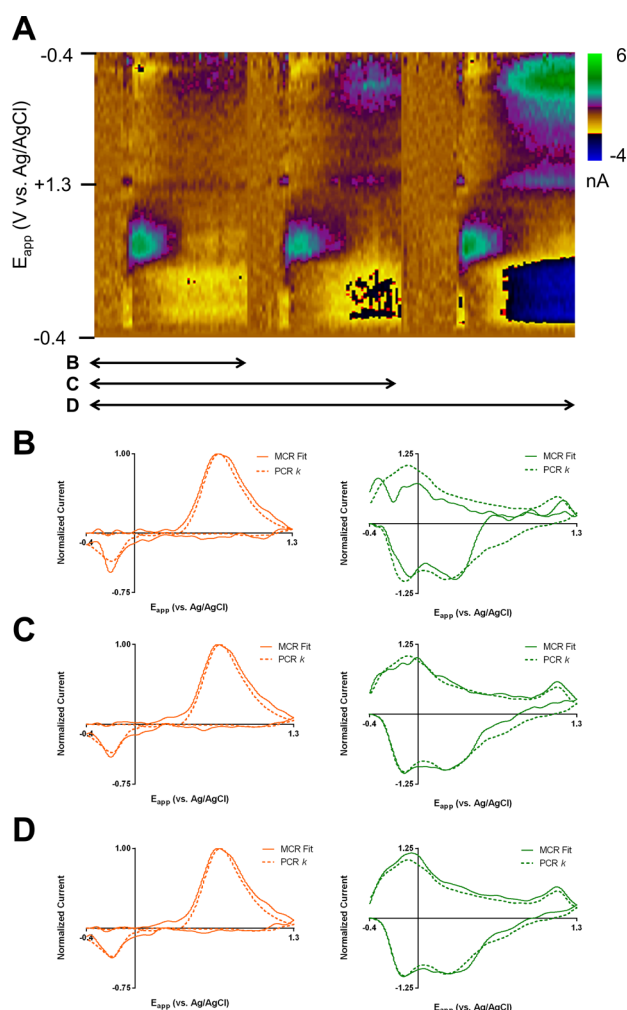
Figure 5. Spectral fits for MCR−ALS analysis of increasing numbers of electrically evoked dopamine transients. (A) Color plot showing three separate transients joined together, with the arrows underneath indicating the data windows used in subsequent MCR−ALS analysis. (B−D) Spectral fits (solid lines) for dopamine (left) and pH (right) for analysis of one (B), two (C), and three transients (D). PCA−ILS k vectors for the two analytes are shown as dashed lines for comparison.

transient windows. Unconstrained MCR−ALS spectral fits were determined for increasing numbers of snippets, and the PCA−ILS k vectors are shown for comparison. Analysis of only one snippet provided a moderately accurate estimate of the DA spectrum; however, the pH spectrum is only weakly determined and both are considerably noisy (Figure 5B). Increasing the number of snippets (Figure 5C,D) provided increasingly good estimations of the underlying component spectra, with improvements in the spectral shapes and noise.

The most computationally inexpensive application of MCR−ALS relies on determining a subset of the experimental data for model definition that is used to analyze the entirety. Voltammetric profiles are anticipated to remain constant throughout a given experimental session, and thus, the spectral estimates were determined. Ideally, the training subset should contain considerable contributions from all components expected throughout the experimental data. Additionally, signals should be resolved from one another, which may be evaluated using analysis of EFA time-course estimates of the analytes present in a given window. Single-analyte data subsets can be used; however, care must be taken to ensure that all analytes are represented in the training subset. Under-representation of a given analyte can lead to poor estimates of its spectral profile. Further, depending on data quality, constraints, like the P−ALS equality approach, may be needed, using either using single-analyte experimental CVs or a library approach using separately collected data.

To test MCR−ALS performance, each experimental-data set was analyzed to select a training submatrix to generate dopamine and pH spectral-profile estimates, which were subsequently used to analyze the other experimental data. These results were compared to those obtained from PCA−ILS using separate training data (Table 1). For each fit, the correlation coefficient between the MCR−ALS and the PCA−ILS estimates and the signal power of the difference between the MCR−ALS and PCA−ILS estimates (their lack of agreement, normalized to signal power of the PCA−ILS estimates) were determined. Overall, there was good agreement between PCA−ILS and MCR−ALS. In general, the estimation of pH profiles differed more between the two techniques than those for DA, and there was greater variability in the performance of MCR−ALS for determining the pH profiles. This was generally due to the difficulty in finding isolated pH spectra within the ICSS data, given that electrical stimulations were typically closely spaced, to use for model training. However, no separate training data was used in this analysis, and the collection and inclusion of even minimal amounts of reference data collected separately from the experiment could be expected to improve this performance.

Residual Analysis. One of the advantages of using higher-order calibration models is the use of residual analysis for interferent detection and evaluation of model applicability. Thus, we sought to explore adoption of the residual-analysis procedure introduced by Jackson and Mudholkar,[59] currently used in FSCV with PCA−ILS analysis, as a first step toward establishing a means of model validation during the application of a constructed model to other experimental data. During the training phase, significant interferences can typically be detected, through methods such as EFA and OPA or fit distortion.

This residual-analysis procedure relies on calculation of an experiment-specific Q_α value, a threshold for residual-value evaluation for each voltammetric measurement to determining model suitability ($Q_t > Q_\alpha$ leads to model rejection for analysis of that data) that is characteristic of the noise level.[9] This is done using PCA, identification of the significant components, and error eigenvalue analysis (i.e., those associated with the nonsignificant principal components). In PCA−ILS analysis, this step is performed during spectral-profile definition through training-data analysis. However, we sought to evaluate whether the use of the experimental data itself could generate a suitable estimate. For each experimental-data set, random sets of 5 s

Table 1. Summary of Correlation and Lack of Agreement between PCA−ILS and MCR−ALS Estimates of Spectral and Concentration Profiles[§]

		correlation coefficient (R^2)			lack of agreement[a] (%)		
		average	min	max	average	min	max
spectra	DA	0.985	0.971	0.995	3.76	2.16	5.77
	pH	0.984	0.963	0.996	4.35	1.21	12.1
concentration	DA	0.994	0.983	0.999	1.82	0.37	4.92
	pH	0.980	0.942	0.999	4.77	0.48	15.8

[a]Lack of agreement = $100\% * \frac{[\sum_{i=0}^{N}(x_{i,\text{MCR−ALS}} - x_{i,\text{PCA−ILS}})^2]}{[\sum_{i=0}^{N}(x_{i,\text{PCA−ILS}})^2]}$. [§]For the in vivo FSCV data collected during the intracranial self-stimulation trials ($n = 25$ rats).

windows ($n = 6$ windows/data set, 50 CVs per window, 25 data sets) were obtained and analyzed with PCA to estimate Q_α using the eigenvalues of the nonsignificant components identified by Malinowski's F-test. Table 2 shows a summary

Table 2. Percent Difference between Q_α Values[a]

	average	minimum	maximum
percent difference (%) between training-set- and experimentally-derived Q_α values	(+)10.4	(−)0.8	(+)132.9
relative standard deviation of experimental Q_α values	16.6	5.5	48.3

[a]Determined from Malinowski F-test analysis of independently collected training set (10 CVs × 1) and experimental data (50 CVs × 6) for 25 intracranial self-stimulation data sets and relative standard deviation of the latter.

of these estimates, as well as those from PCA analysis of the associated training-set data. Overall, moderate agreement between the values obtained from analysis of training set and experimental data was observed. Although there was a large spread of the percent differences, the Q_α values obtained from the different experimental windows were consistent (average relative standard deviation of 17.5%). The majority of cases (60%) had lower Q_α values obtained from experimental data than from the training-set data, meaning that for these data, the former approach generates more conservative Q_α values.

Finally, one advantage of MCR−ALS as an exploratory technique should be highlighted. In PCA−ILS, the number of components is defined using a priori knowledge, and data that fails residual analysis is thrown out. Further, residual analysis reveals the presence of interferents but does not necessarily provide robust information about the interferent spectra. With MCR−ALS, should a set of data fail residual analysis, a component can be added to the model, and MCR−ALS can be performed to attempt to gather information on the nature of the interferent. This advantage is highlighted in Figure 6. Here, a DA-pH mixture (Figure 6A) is analyzed with both PCA−ILS (using only a dopamine training set) and MCR−ALS with two components. The PCA−ILS residual spectrum (Figure 6B) does have general features resembling the reference pH k vector; however, some of the current has been assigned as arising from dopamine, resulting in a deviation in the residual spectrum from a "pure" pH signal. Alternatively, the MCR−ALS estimate (Figure 6C) gives a more robust estimate of the interferent spectra, giving greater confidence in component identification.
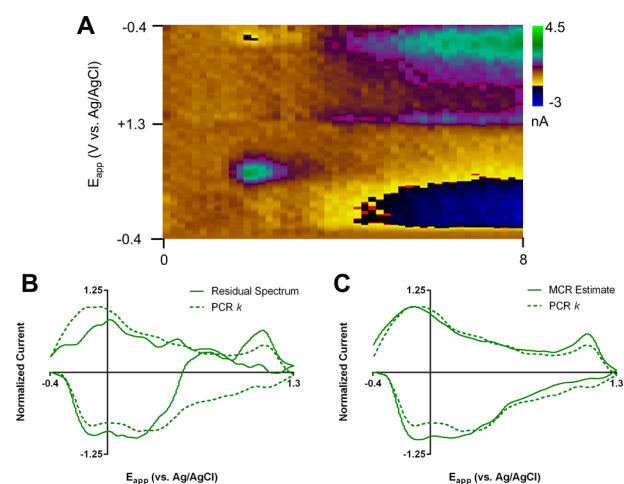
Figure 6. Interferent identification using MCR−ALS. (A) In vivo color plot containing both DA and pH signals. (B) Residual spectrum (solid line) after PCA−ILS analysis with a DA-only training set. (C) MCR−ALS spectral estimate (solid line) using a two-component model. The dashed line shows the PCA−ILS pH k vector estimated for this experiment.

■ CONCLUSIONS

The MCR−ALS approach has several advantages over PCA−ILS in the analysis of FSCV data, including more flexibility in model definition, decreased experimental requirements (i.e., relaxation of the need to collect separate training data), and more robust handling of interferents. However, due to this increased freedom in model definition, considerably more caution must be employed, and the methods explored here for its deployment (OPA, EFA, and P−ALS) require more user input than the currently established PCA−ILS protocols. Regardless, the two techniques generated highly similar spectral and concentration estimates under the conditions studied here, and MCR−ALS demonstrates considerable potential as a complementary or alternative analysis method to PCA−ILS or other reported methods like elastic net regression. However, further characterization of the technique for FSCV will greatly help in understanding its potential and limitations. In particular, the data studied here contained relatively high signal-to-noise ratios, allowing ready discrimination of analyte and noise contributions. Preliminary studies suggest that MCR−ALS may be suitable for analysis of noisier data; however, pretreatment of the data (PCA data reconstruction), the use of larger number of signal-containing spectra, and the strength of constraint imposition become important considerations.

ASSOCIATED CONTENT

S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.anal-chem.7b02771.

> Summary of rank estimated by Malinowski's *F*-test, OPA, and EFA as a function of temporal separation of dopamine and pH signals (Table S-1); effect of weighting parameter on MCR−ALS fits (Figure S-1) (PDF)

AUTHOR INFORMATION

Corresponding Author

*Phone: 919-962-1472; E-mail: rmw@email.unc.edu.

ORCID

R. Mark Wightman: 0000-0003-2198-139X

Author Contributions

All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Booksh, K. S.; Kowalski, B. R. *Anal. Chem.* **1994**, *66*, 782A−791A.
(2) Lavine, B. K.; Workman, J. *Anal. Chem.* **2013**, *85*, 705−714.
(3) Olivieri, A. C. *Anal. Chem.* **2008**, *80*, 5713−5720.
(4) Yorgason, J. T.; Espana, R. A.; Jones, S. R. *J. Neurosci. Methods* **2011**, *202*, 158−164.
(5) Kishida, K. T.; Saez, I.; Lohrenz, T.; Witcher, M. R.; Laxton, A. W.; Tatter, S. B.; White, J. P.; Ellis, T. L.; Phillips, P. E.; Montague, P. R. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 200−205.
(6) Heien, M. L. A. V.; Johnson, M. A.; Wightman, R. M. *Anal. Chem.* **2004**, *76*, 5697−5704.
(7) Heien, M. L. A. V.; Khan, A. S.; Ariansen, J. L.; Cheer, J. F.; Phillips, P. E. M.; Wassum, K. M.; Wightman, R. M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10023−10028.
(8) Kramer, R. *Chemometric Techniques for Quantitative Analysis*; Marcel Dekker, Inc.: New York, NY, 1998.
(9) Keithley, R. B.; Heien, M. L.; Wightman, R. M. *TrAC, Trends Anal. Chem.* **2009**, *28*, 1127−1136.
(10) Keithley, R. B.; Wightman, R. M. *ACS Chem. Neurosci.* **2011**, *2*, 514−525.
(11) Booksh, K. S. In *Encyclopedia of Analytical Chemistry*; John Wiley & Sons, Ltd: New York, 2006.
(12) Wang, Y. D.; Kowalski, B. R. *Appl. Spectrosc.* **1992**, *46*, 764−771.
(13) Woody, N. A.; Feudale, R. N.; Myles, A. J.; Brown, S. D. *Anal. Chem.* **2004**, *76*, 2595−2600.
(14) Feudale, R. N.; Woody, N. A.; Tan, H. W.; Myles, A. J.; Brown, S. D.; Ferre, J. *Chemom. Intell. Lab. Syst.* **2002**, *64*, 181−192.
(15) Rodeberg, N. T.; Johnson, J. A.; Cameron, C. M.; Saddoris, M. P.; Carelli, R. M.; Wightman, R. M. *Anal. Chem.* **2015**, *87*, 11484−11491.
(16) Johnson, J. A.; Rodeberg, N. T.; Wightman, R. M. *ACS Chem. Neurosci.* **2016**, *7*, 349−359.
(17) Rodeberg, N. T.; Sandberg, S. G.; Johnson, J. A.; Phillips, P. E. M.; Wightman, R. M. *ACS Chem. Neurosci.* **2017**, *8*, 221−234.
(18) Sinanian, M. M.; Cook, D. W.; Rutan, S. C.; Wijesinghe, D. S. *Anal. Chem.* **2016**, *88*, 11092−11099.
(19) Pere-Trepat, E.; Lacorte, S.; Tauler, R. *J. Chromatogr A* **2005**, *1096*, 111−122.
(20) Dantas, C.; Tauler, R.; Castro Ferreira, M. M. *Anal. Bioanal. Chem.* **2013**, *405*, 1293−1302.
(21) Pere-Trepat, E.; Hildebrandt, A.; Barcelo, D.; Lacorte, S.; Tauler, R. *Chemom. Intell. Lab. Syst.* **2004**, *74*, 293−303.
(22) Gargallo, R.; Tauler, R.; CuestaSanchez, F.; Massart, D. L. *TrAC, Trends Anal. Chem.* **1996**, *15*, 279−286.
(23) Bortolato, S. A.; Olivieri, A. C. *Anal. Chim. Acta* **2014**, *842*, 11−19.
(24) Esteban, M.; Arino, C.; Díaz-Cruz, J. M.; Díaz-Cruz, M. S.; Tauler, R. *TrAC, Trends Anal. Chem.* **2000**, *19*, 49−61.
(25) Díaz-Cruz, M. S.; Mendieta, J.; Tauler, R.; Esteban, M. *Anal. Chem.* **1999**, *71*, 4629−4636.
(26) Torres, M.; Díaz-Cruz, J. M.; Arino, C.; Grabaric, B. S.; Tauler, R.; Esteban, M. *Anal. Chim. Acta* **1998**, *371*, 23−37.
(27) Grabaric, B. S.; Grabaric, Z.; Tauler, R.; Esteban, M.; Casassas, E. *Anal. Chim. Acta* **1997**, *341*, 105−120.
(28) Olivieri, A. C. *Chem. Rev.* **2014**, *114*, 5358−5378.
(29) Malik, A.; Tauler, R. In *Resolving Spectral Mixtures: With Applications from Ultrafast Time-Resolved Spectroscopy to Super-Resolution Imaging*; Ruckebusch, C., Ed.; Elsevier: Amsterdam, The Netherlands, 2016.
(30) Smilde, A. K.; Hoefsloot, H. C. J.; Kiers, H. A. L.; Bijlsma, S.; Boelens, H. F. M. *J. Chemom.* **2001**, *15*, 405−411.
(31) Tauler, R.; Izquierdoridorsa, A.; Casassas, E. *Chemom. Intell. Lab. Syst.* **1993**, *18*, 293−300.
(32) Tauler, R.; Smilde, A.; Kowalski, B. *J. Chemom.* **1995**, *9*, 31−58.
(33) Manne, R. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 89−94.
(34) Tauler, R.; Kowalski, B.; Fleming, S. *Anal. Chem.* **1993**, *65*, 2040−2047.
(35) Tauler, R.; Izquierdoridorsa, A.; Gargallo, R.; Casassas, E. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 163−174.
(36) Gemperline, P. J.; Cash, E. *Anal. Chem.* **2003**, *75*, 4236−4243.
(37) Bucher, E. S.; Brooks, K.; Verber, M. D.; Keithley, R. B.; Owesson-White, C.; Carroll, S.; Takmakov, P.; McKinney, C. J.; Wightman, R. M. *Anal. Chem.* **2013**, *85*, 10344−10353.
(38) Johnson, J. A.; Rodeberg, N. T.; Wightman, R. M. *ACS Chem. Neurosci.* **2016**, *7*, 349−359.
(39) Garris, P. A.; Kilpatrick, M.; Bunin, M. A.; Michael, D.; Walker, Q. D.; Wightman, R. M. *Nature* **1999**, *398*, 67−69.
(40) de Juan, A.; Tauler, R. In *Resolving Spectral Mixtures: With Applications from Ultrafast Time-Resolved Spectroscopy to Super-Resolution Imaging*, Ruckebusch, C., Ed.; Elsevier: Amsterdam, The Netherlands, 2016.
(41) Wasim, M.; Brereton, R. G. *Chemom. Intell. Lab. Syst.* **2004**, *72*, 133−151.
(42) Gemperline, P. J. *J. Chem. Inf. Model.* **1984**, *24*, 206−212.
(43) Maeder, M. *Anal. Chem.* **1987**, *59*, 527−530.
(44) Gampp, H.; Maeder, M.; Meyer, C. J.; Zuberbuehler, A. D. *Anal. Chim. Acta* **1987**, *193*, 287−293.
(45) Borgen, O. S.; Davidsen, N.; Mingyang, Z.; Oyen, O. *Microchim. Acta* **1986**, *89*, 63−73.
(46) Gemperline, P. J. *Anal. Chem.* **1986**, *58*, 2656−2663.
(47) Keithley, R. B.; Carelli, R. M.; Wightman, R. M. *Anal. Chem.* **2010**, *82*, 5541−5551.
(48) Malinowski, E. R. *J. Chemom.* **1989**, *3*, 49−60.
(49) Sanchez, F. C.; Toft, J.; vandenBogaert, B.; Massart, D. L. *Anal. Chem.* **1996**, *68*, 79−85.
(50) Vivo-Truyols, G.; Torres-Lapasio, J. R.; Garcia-Alvarez-Coque, M. C.; Schoenmakers, P. J. *J. Chromatogr A* **2007**, *1158*, 258−272.
(51) Tauler, P.; Casassas, E. *Anal. Chim. Acta* **1989**, *223*, 257−268.
(52) Tauler, R.; Casassas, E.; Izquierdoridorsa, A. *Anal. Chim. Acta* **1991**, *248*, 447−458.
(53) Bro, R.; Sidiropoulos, N. D. *J. Chemom.* **1998**, *12*, 223−247.
(54) Bro, R.; DeJong, S. *J. Chemom.* **1997**, *11*, 393−401.
(55) Van Benthem, M. H.; Keenan, M. R.; Haaland, D. M. *J. Chemom.* **2002**, *16*, 613−622.
(56) Faber, K.; Kowalski, B. R. *Anal. Chim. Acta* **1997**, *337*, 57−71.
(57) Faber, K.; Kowalski, B. R. *J. Chemom.* **1997**, *11*, 53−72.
(58) Malinowski, E. R. *J. Chemom.* **1999**, *13*, 69−81.

(59) Jackson, J. E.; Mudholkar, G. S. *Technometrics* **1979**, *21*, 341−349.