We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

High-Order Calibration and Data Analysis in Chromatography

Hai-Long Wu, Xiao-Dong Sun, Huan Fang and
Ru-Qing Yu

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.78624

Abstract

Multiway data analysis and tensorial calibration are gaining widespread acceptance with the rapid development of multichannel chromatographic instruments. By combining chromatographic techniques with chemometrics based on high-order calibration methods, some traditional problems in analysis, such as complicated pretreatment steps, long elution times, or even worse analysis results, can be avoided/improved. This chapter presents an overview from second-order to third-order data that cover theories and applications together with corresponding data processing in chromatography.

Keywords: chemometrics, second-order advantages, trilinear model, high-performance liquid chromatography-diode array detection, liquid chromatography-mass spectrometry

1. Introduction

Chromatography, first employed in Russia by the Italian-born scientist Mikhail Tsvet in 1900, is a laboratory technique for the separation of a mixture. In the first decade of the twentieth century, scientists continued to work with chromatography primarily for the purpose of separating plant pigments such as chlorophyll, carotenes, and xanthophylls. Since these pigments separated showed various colors (green, orange, and yellow, respectively), they gave the technique its name. After various types of chromatography had sprung up in the 1930s and 1940s, it became useful for many separation processes. Up to now, many chromatographic techniques have been developed, and they can be classified according to different properties. Based on chromatographic bed shape techniques, they can be divided into column chromatography and planar chromatography. Also, gas chromatography and liquid chromatography are

classified by the physical state of mobile phase. In addition, there are also many other categories classified by other properties (i.e. separation mechanism, special techniques); but the chromatographic classification is out of scope of this chapter.

The purpose of chromatography is to separate the components of a mixture for later use. The mixture is dissolved in a fluid called the mobile phase, which carries it through a structure holding another material called the stationary phase. The separation is based on differential partitioning between the mobile and stationary phases. Subtle differences in a compound's partition coefficient result in differential retention on the stationary phase and thus affect the separation.

Nowadays, due to its prominent separation properties, chromatography techniques have become an indispensable tool for the routine analysis and research in pharmaceutical, biomedical, food, and environmental industries [1]. However, there are two main drawbacks needed to be solved/improved. The first one is about the sample itself; when complex matrix samples are analyzed, some proper tedious pretreatment procedures, such as extraction and purification, are necessary to remove the potential interferences contained in complex matrices. Optimizing these procedures is rather tedious and large sum of solvents' consumption are inevitable, making this method become uneconomical and environmentally unfriendly. What's more, in traditional chromatography analysis, when a complex sample is analyzed, the overlap between the analytes and matrix constituents is frequently observed; consequently, a long time or much more complex chromatography condition is required for the separation. In general, the elution time for each sample often costs 30–50 min, which is quite time consuming and inefficient. In the same time, some other problems such as baseline drift, changes in the shape of the peaks, incomplete extraction of the analytes, and shifts in the elution times may also decrease the quality of the final result of the analysis. Another problem with chromatography is due to its universal aspect. There are now hundreds of different chromatographic columns, which can be obtained from the market and new ones are being developed constantly [2, 3]. However, when faced with the large number of possible columns, it is hard for analysts to select which could be the most appropriate one for a given condition. Meanwhile, many laboratories and public institutions may not possess all available stationary phases and the column performance may become worse during long-term storage and/or usage in LC analysis [4]. Thus, the analysts often waste a lot of time in search of the most appropriate one from several different stationary phases for analysis. All of the above shortcomes may hinder the further development for chromatographic applications.

A current trend in quantitative analysis is to avoid tedious sample preprocessing steps and long chromatographic elution, exploiting the ability of modern data processing tools for mathematical resolution of coeluting components [5]. The combination of suitable chemometric tools along with chromatographic-spectral data or chromatographic-mass data may solve/improve the problem. With less time and solvent consumption, better quantitative results can be obtained. The multiway (second- and third-order) calibration based on "mathematical separation" is a dazzling pearl in the field of chemical analysis and can calibrate the potential interferences and resolve coeluting peaks successfully in real samples with minimum sample preparation steps. Accurately concentration profiles of individual components of interest can also be obtained. This property generally

refers to the prominent "second-order advantage," which has enormous potential in multiway analysis and becomes a recent focus of theoretical research and practical uses. Combining chromatography with multiway calibration has some distinct advantages because it can simplify the tedious multistep pretreatment and exploration of complicated chromatography separate conditions, showing the potential abilities for the analysis of various samples with different interferences at a time. Tedious pretreatment or purification procedures can be discarded by using prominent "mathematic separation" instead of tradition "physical/chemical separation." HPLC coupled with second-order calibration methods is especially popular for it can rapidly and simultaneously determine multiple compounds in complex backgrounds with unknown interferences, resolve coeluted peaks, and remove baselines drifts [6–10].

So far, a lot of algorithms for decomposition of multiway data arrays have already been proposed and genuinely provided alternative tools to analytical chemists for the convenient study of the body of multiway data arrays. Several methodologies have also been expounded in "Encyclopedia of Analytical Chemistry" [11] and "Factor Analysis in Chemistry" [12] at some length. To help readers systematically and intensively understand about concerning algorithms, a detailed description including multilinear models, the multiway cyclic symmetry property, the algorithms for multiway calibration, the estimation of the chemical rank, the toolbox for multiway calibration, and other fundamental issues and applications in chromatography has been presented in this paper.

2. Terminology and nomenclature in multiway data

To facilitate understanding for readers when dealing with multivariate analysis on multiway data arrays, it is necessary to introduce the terminology and nomenclature used in multiway data in the following.

2.1. Terminology

The relationship and difference between the concepts of "data order" and "data way" should be investigated firstly. The term "order" is the dimensions for data of a single sample and term "way" represents the data arrays stacked by all samples with similar properties. As shown in **Figure 1**, zeroth order corresponds to instruments producing a single response per sample, such as the reading of a pH meter or the absorbance at a single wavelength. First-order data are arranged as a vector or first-order tensor for a single sample, such as UV, fluorescence, infrared, and nuclear magnetic resonance spectra. At the same time, second-order data are formed when matrix data can be obtained for a single sample. There are two ways that second-order data can be obtained: (i) using a single instrument such as excitation-emission spectro-fluorimeter (EEMs) or diode-array spectrophotometer to monitor the kinetics of a chemical reaction and (ii) using the hyphenated instruments such as high performance liquid chromatography-photodiode array detection (HPLC-DAD) or liquid chromatography-mass spectrometry (LC-MS). When the second-order data that obtained from a series of samples
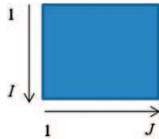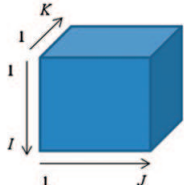
Figure 1. Relationships and differences between the concepts of "data order" and "data way" described with symbols.

(calibration and prediction samples) are stacked in one direction, three-dimensional array, which is also called as three-way array can be obtained, and the corresponding data are usually known as three-way data. Hence, when a series of samples are stacked into a single, zeroth-order (a scalar), first-order (a vector), second-order (a matrix), third-order (a three-way array), and higher order tensors can yield the corresponding one-way, two-way, three-way, four-way, and N-way data sets, respectively. The zeroth-order tensor calibration is also called as univariate calibration. This method has great restraint on its application as it needs full selectivity for the signals of target analytes. Except univariate calibration for the analysis of data, others are known as multivariate calibration, the analysis of second-order tensor and higher order tensor is denoted as multiway multivariate or multicomponent calibration.

Meanwhile, a detailed description for various sample types is also provided. Based on different functions, samples can be divided into calibration, prediction, and actual sets. Actual sets include predicted samples (the target analyte(s) is (are) unambiguously included) and/or real samples (whether the target analyte(s) is (are) included or unknown). Constituents present in the samples used for calibration and validation are regularly called "known" or "expected," which is expected in these sets as they are expected to be existed in actual samples. The expected constituents can be further divided into "calibrated" and "uncalibrated." The concentrations of former ones used for calibration are predesigned and known, while those of "calibrated" components in actual sets can also be available, involving the analyte(s) of interest. On the other hand, the constituents which are only included in actual sets are called "unknown" or "unexpected" and also potential interferences.

2.2. Nomenclature

In this chapter, lowercase italics represent scalars; two-way matrices are denoted by bold capitals; underlined bold capitals designate three-way arrays, the superscript T represents the transpose of a matrix, and the superscript + is the Moore-Penrose generalized inverse of a matrix. $||\cdot||_F$ designates the Frobenius matrix norm. To have a better understanding about

$$a_{(i)}^T = diagm(\mathbf{B}^+ \mathbf{X}_{i..}(\mathbf{C}^T)^+),$$
$$(i = 1, 2, ..., I)$$
$$\mathbf{X}_{I \times JK} = [\mathbf{X}_{..1}, \mathbf{X}_{..2}, ..., \mathbf{X}_{..K}]$$
$$\sigma(\mathbf{A}) = \sum_{i=1}^{I} \left\| \mathbf{X}_{i..} - \mathbf{B}diag(\mathbf{a}_{(i)})\mathbf{C}^T \right\|_F^2$$

$$b_{(j)}^T = diagm(\mathbf{C}^+ \mathbf{X}_{.j.}(\mathbf{A}^T)^+),$$
$$(j = 1, 2, ..., J)$$
$$\mathbf{X}_{J \times KI} = [\mathbf{X}_{1..}, \mathbf{X}_{2..}, ..., \mathbf{X}_{I..}]$$
$$\sigma(\mathbf{B}) = \sum_{j=1}^{J} \left\| \mathbf{X}_{.j.} - \mathbf{C}diag(\mathbf{b}_{(j)})\mathbf{A}^T \right\|_F^2$$

$$c_{(k)}^T = diagm(\mathbf{A}^+ \mathbf{X}_{..k}(\mathbf{B}^T)^+),$$
$$(k = 1, 2, ..., K)$$
$$\mathbf{X}_{K \times IJ} = [\mathbf{X}_{.1.}, \mathbf{X}_{.2.}, ..., \mathbf{X}_{.J.}]$$
$$\sigma(\mathbf{C}) = \sum_{k=1}^{K} \left\| \mathbf{X}_{..k} - \mathbf{A}diag(\mathbf{c}_{(k)})\mathbf{B}^T \right\|_F^2$$
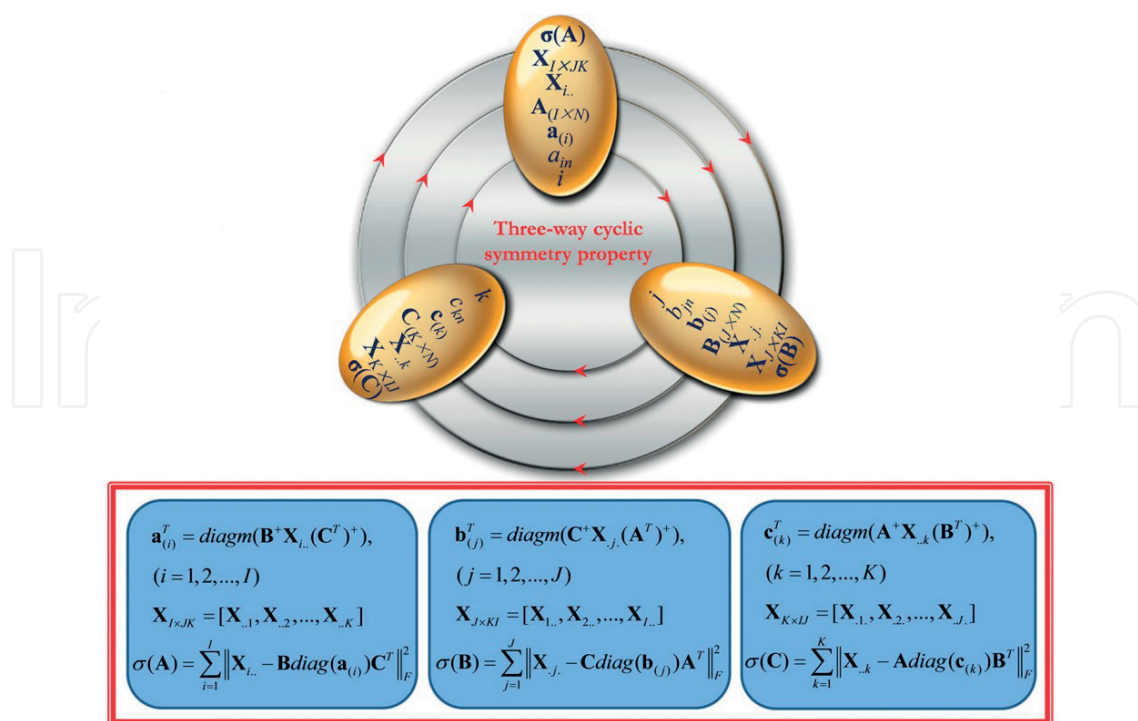
Figure 2. Schematic representation for three-way cyclic symmetry property.

the multiway calibration, readers are advised to comprehend an inner cyclic symmetry property of trilinear decomposition proposed by our laboratory in 1996 and also called as three-way cycle symmetry. As shown in **Figure 2**, elements, vectors, subscripts, and physical modes in resolved matrices, sliced matrices, and unfolded matrices, together with residue and resolution formulas, all obey the principle of inner cyclic symmetry property, circumrotating along the same way. **Table 1** provides the detailed information of the nomenclature mentioned.

X	Three-way data array
I, J, K	The three dimensions of three modes of **X**
x_{ijk}	The ijkth element of **X**
$\mathbf{A}_{I \times N}$, $\mathbf{B}_{J \times N}$, $\mathbf{C}_{K \times N}$	The three underlying profile matrices of **X** with $I \times N$, $J \times N$, and $K \times N$, respectively
a_{in}, b_{jn}, c_{kn}	The inth, jnth, and knth elements of the three underlying profile matrices **A**, **B**, and **C**, respectively
$\mathbf{a}_{(i)}$, $\mathbf{b}_{(j)}$, $\mathbf{c}_{(k)}$	The ith, jth, and kth row vectors of profile matrices **A**, **B,** and **C**, respectively
$diag(\mathbf{a}_{(i)})$, $diag(\mathbf{b}_{(j)})$, $diag(\mathbf{c}_{(k)})$	Diagonal matrices with elements equal to the elements of $\mathbf{a}_{(i)}$, $\mathbf{b}_{(j)}$, and $\mathbf{c}_{(k)}$, respectively
$\mathbf{X}_{i..}$, $\mathbf{X}_{.j.}$, $\mathbf{X}_{..k}$	The ith horizontal, jth lateral, and kth frontal slices of **X**, respectively
$\mathbf{E}_{i..}$, $\mathbf{E}_{.j.}$, $\mathbf{E}_{..k}$	The ith horizontal, jth lateral, and kth frontal slices of the three-way array residue **E,** respectively
$\mathbf{e}_{ijk}$	The ijkth element of the three-way residue array **E**

Table 1. Detailed information of the nomenclature mentioned.

Similar to the three-way cyclic symmetry, the four-way and five-way cyclic symmetries for quadrilinear and quinquelinear decomposition can be obtained easily by simple mathematical manipulation of exchanging the symbols similar to three-way cyclic symmetry. These regularities provide useful instructions for the standardization of symbol systems in multiway data analysis, for better understanding the essence of multiway multilinear decomposition, developing new multiway calibration algorithms, and exploring multilinear algebra in mathematics.

3. Theory

3.1. Multilinear models

According to the data type and its inner cyclic symmetry property, the multilinear models can be divided into trilinear, quadrilinear, quinquelinear, and even higher linear models. In chromatographic analysis combined with multiway calibration, the trilinear and quadrilinear models are commonly used.

3.1.1. Trilinear model

Harshman [13] together with Carroll and Chang [14] first proposed the PARAFAC (PARallel FACtor analysis) model with the name of CANDECOMP in the year of 1970. In this trilinear model, each element x_{ijk} of a three-way array $\mathbf{X}$ ($I{\times}J{\times}K$) can be reasonably fit to the following equation:

$$x_{ijk} = \sum_{n=1}^{N} a_{in}b_{jn}c_{kn} + e_{ijk}, i = 1, 2, …, I; j = 1, 2, …, J; k = 1, 2, …, K. \tag{1}$$

where N represents the total number of detectable components including the component(s) of interest, uncalibrated background(s), and unknown interferences. **Figure 3** illustrates the graphical representation of a trilinear model of three-way data array $\mathbf{X}$. $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are the three underlying profile matrices of X with $I \times N$, $J \times N$, and $K \times N$, respectively; $\mathbf{I}$ is the three-way diagonal core array of size $N \times N \times N$ with ones on the superdiagonal and zeros
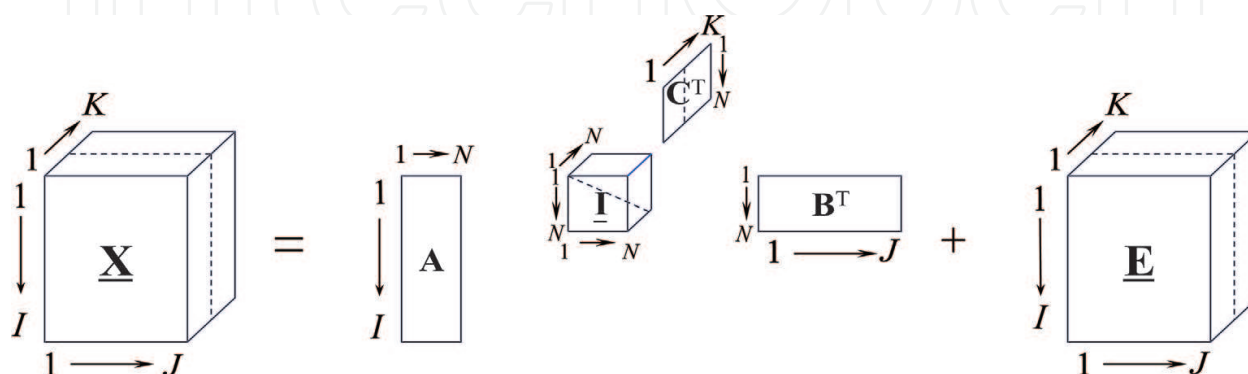

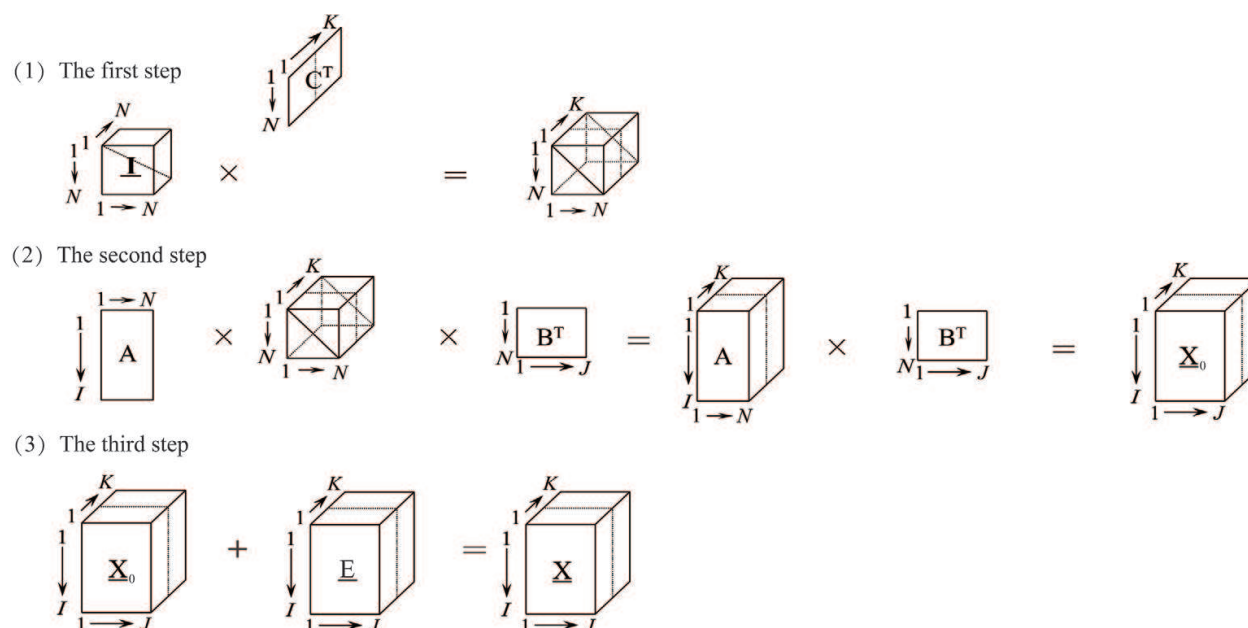**Figure 3.** Schematic representation of trilinear model.

Figure 4. The inverse procedures to return three-way data array X.

elsewhere; and E is the three-way residue data array of size $I \times J \times K$. To further comprehend the mathematic meaning of the trilinear model graphically expressed, the response data array $\mathbf{X}$ is returned with three inverse steps as described in **Figure 4**.

3.1.2. Quadrilinear model

Considering a model of the real-valued four-way data array $\mathbf{X}$ ($I \times J \times K \times L$), in which each element x_{ijk} can be expressed as [8, 15]:

$$x_{ijkl} = \sum_{n=1}^{N} a_{in} b_{jn} c_{kn} d_{ln} + e_{ijkl}, \quad i = 1, 2, \ldots, I; j = 1, 2, \ldots, J; k = 1, 2, \ldots, K; l = 1, 2, \ldots, L. \quad (2)$$

where a_{in}, b_{jn}, c_{kn}, and d_{ln} correspond to the underlying profile matrices $\mathbf{A}_{I \times N}$, $\mathbf{B}_{J \times N}$, $\mathbf{C}_{K \times N}$, and $\mathbf{D}_{L \times N}$ of $\mathbf{X}$ ($I \times J \times K \times L$), respectively. The term e_{ijkl} is the element of the four-way residual array $\mathbf{E}$ ($I \times J \times K \times L$). Then, the modeled part of x_{ijkl} is quadrilinear in the parameter sets a_{in}, b_{jn}, c_{kn}, and d_{ln}. The graphical representation of a quadrilinear model of four-way data array $\mathbf{X}$ is shown in **Figure 5**.

3.2. Data preprocessing

The correctness of decomposition of a multilinear model requires that the multilinear model holds multilinearity. However, there are some nonmultilinear factors which can cause a multilinear model to deviate the multilinearity. For example, in the chromatography type of trilinear model such as HPLC-DAD and LC-MS data, the time shift and baseline problem among different runs will cause the trilinear model to deviate the trilinearity. Thus, the data
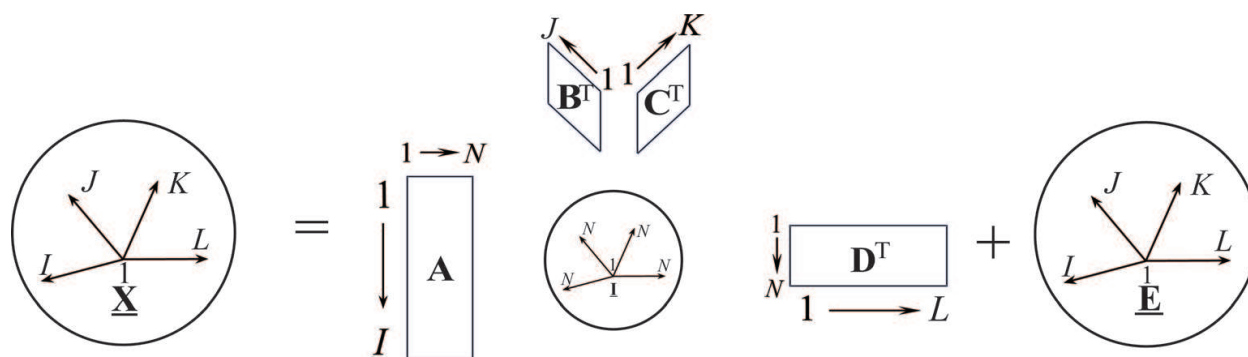
Figure 5. Schematic representation of quadrilinear model.

Step ①: Construction of three-way data

Step ②: Trilinear decomposition

Step ③: Pseudo-univariate calibration

$$x_{ijk} = \sum_{n=1}^{N} a_{in} b_{jn} c_{kn} + e_{ijk}, \quad i = 1,\ldots,I; j = 1,\ldots,J; k = 1,\ldots,K.$$
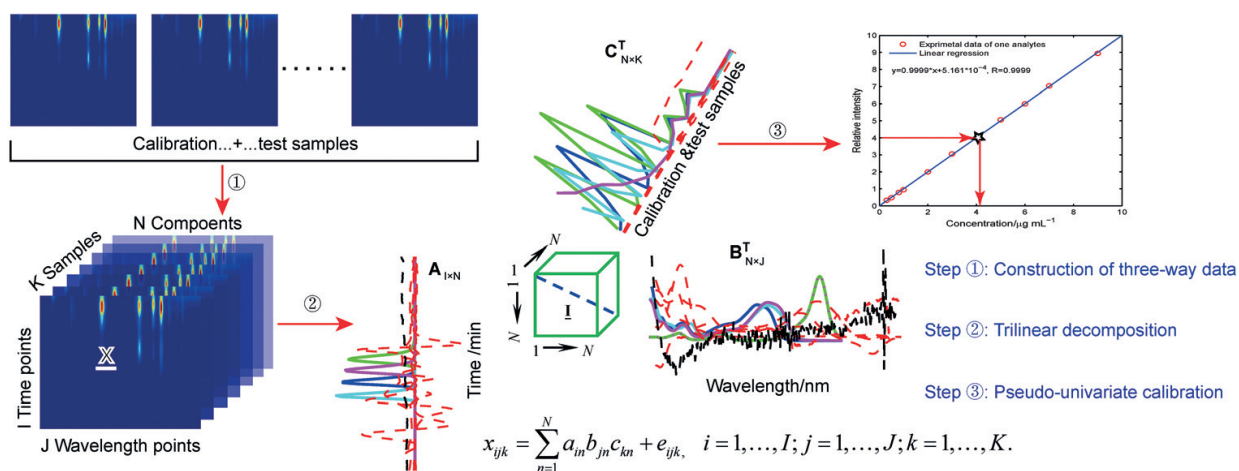
Figure 6. Schematic representation of entire chemometrics-assisted LC-DAD analytical strategy.

arrays in multivariate calibration must need appropriate data preprocessing procedures before a multilinear decomposition. The schematic representation of entire chemometrics-assisted LC-DAD and LC-MS analytical strategy is shown in **Figures 6** and **7**, respectively.

3.3. Algorithm

3.3.1. ATLD

The ATLD algorithm is a universal second-order calibration method for decomposition of three-way data arrays. It is based on an alternating least squares principle without any constrains and an improved iterative procedure that utilizes the Moore-Penrose generalized inverse based on singular value decomposition. It has been widely used in three-way data analysis due to the advantages of being insensitive to excessive component numbers and fast convergence.

According to its cyclic symmetry property, the trilinear model can also be expressed in matrix notation as follows:
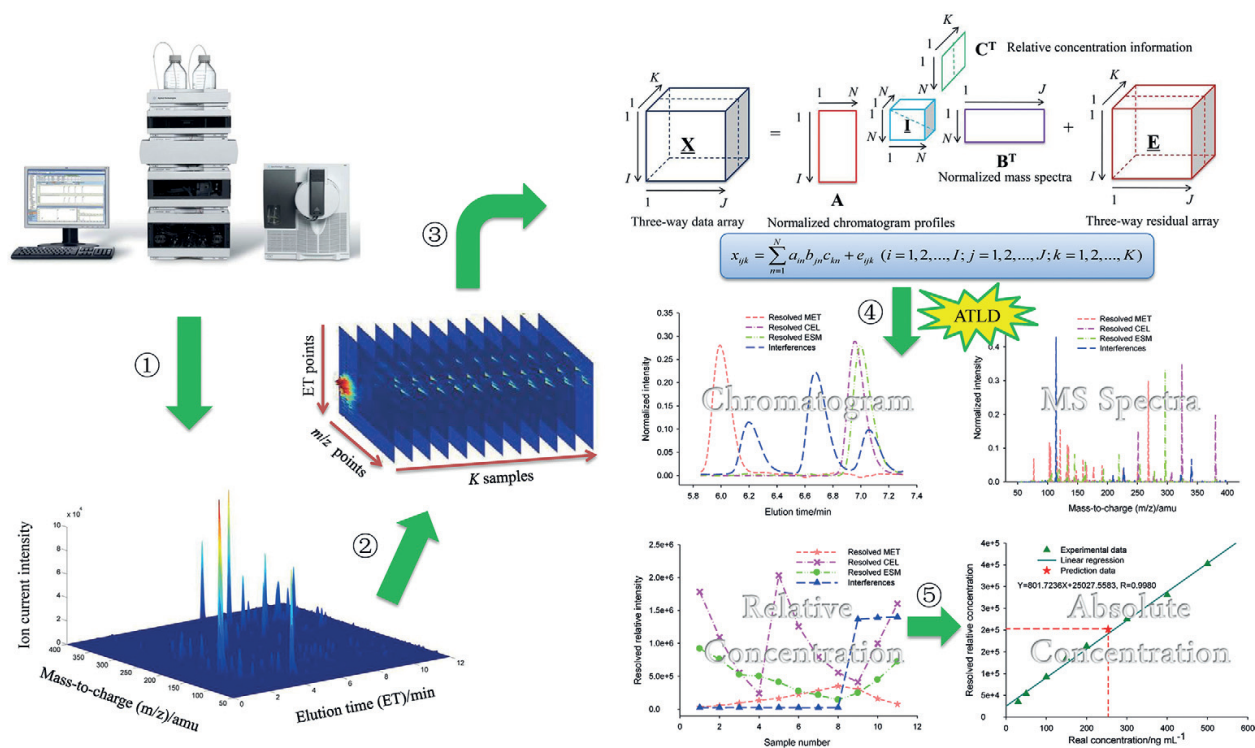
Figure 7. Schematic representation of entire chemometrics-assisted LC-MS analytical strategy.

$$\mathbf{X}_{i..} = \mathbf{B}diag\big(\mathbf{a}_{(i)}\big)\mathbf{C}^{\mathrm{T}} + \mathbf{E}_{i..}, \text{ for } i = 1, 2, \dots, I, \tag{3}$$

$$\mathbf{X}_{.j.} = \mathbf{C}diag\big(\mathbf{b}_{(j)}\big)\mathbf{A}^{\mathrm{T}} + \mathbf{E}_{.j.}, \text{ for } j = 1, 2, \dots, J, \tag{4}$$

$$\mathbf{X}_{..k} = \mathbf{A}diag\big(\mathbf{c}_{(k)}\big)\mathbf{B}^{\mathrm{T}} + \mathbf{E}_{..k}, \text{ for } k = 1, 2, \dots, K, \tag{5}$$

Due to the property called as cyclic symmetry of the trilinear model, the three expressions are equal to each other in mathematics. According to Eqs. (3)–(5), the loss function to be minimized is the sum of the squares of the elements of the residual matrices, which can be expressed as:

$$\sigma\big(\mathbf{a}_{(i)}\big) = \sum_{i=1}^{I} \left\| \mathbf{X}_{i..} - \mathbf{B}diag\big(\mathbf{a}_{(i)}\big)\mathbf{C}^{\mathrm{T}} \right\|_{\mathrm{F}}^{2}, \tag{6}$$

$$\sigma\big(\mathbf{b}_{(j)}\big) = \sum_{j=1}^{J} \left\| \mathbf{X}_{.j.} - \mathbf{C}diag\big(\mathbf{b}_{(j)}\big)\mathbf{A}^{\mathrm{T}} \right\|_{\mathrm{F}}^{2}, \tag{7}$$

$$\sigma\big(\mathbf{c}_{(k)}\big) = \sum_{k=1}^{K} \left\| \mathbf{X}_{..k} - \mathbf{A}diag\big(\mathbf{c}_{(k)}\big)\mathbf{B}^{\mathrm{T}} \right\|_{\mathrm{F}}^{2}, \tag{8}$$

By using the loss functions abovementioned, ATLD alternately minimizes the three objective functions over $\mathbf{C}$ on fixed $\mathbf{A}$ and $\mathbf{B}$, over $\mathbf{A}$ on fixed $\mathbf{B}$ and $\mathbf{C}$, and then over $\mathbf{B}$ on fixed $\mathbf{C}$ and $\mathbf{A}$.

The updates for the three profile matrices (A, B, and C) are based on the least squares principle and can be represented as follows:

$$\mathbf{a}_{(i)}^{\mathrm{T}} = diagm\left(\mathbf{B}^{+}\mathbf{X}_{i..}(\mathbf{C}^{\mathrm{T}})^{+}\right), \quad \text{for } i = 1, 2, \ldots, I, \tag{9}$$

$$\mathbf{b}_{(j)}^{\mathrm{T}} = diagm\left(\mathbf{C}^{+}\mathbf{X}_{.j.}(\mathbf{A}^{\mathrm{T}})^{+}\right), \text{for } j = 1, 2, \ldots, J, \tag{10}$$

$$\mathbf{c}_{(k)}^{\mathrm{T}} = diagm\left(\mathbf{A}^{+}\mathbf{X}_{..k}(\mathbf{B}^{\mathrm{T}})^{+}\right), \quad \text{for } k = 1, 2, \ldots, K, \tag{11}$$

herein $diagm(\cdot)$ stands for a column N-vector and its elements are diagonal elements in square matrix. In every iteration cycle, **A** and **B** are normalized column-wise with unit length. With the help of the resolved profile matrices **C**, we can get the concentrations of analytes of interests in actual samples via regression of the appropriate column of **C** corresponding to each analyte against its standard concentrations.

Due to the operation based on sliced matrices with less size and two other major strategies, ATLD holds the fastest convergence. The truncated least squares method employs the tolerance to truncate the small singular values in the singular value decomposition. In addition, selecting diagonal elements makes ATLD retain trilinearity property indeed and be insensitive to the excessive estimation of component numbers. The advantages have been reviewed by Fleming and Kowalski [16]. Based on the above advantages, it is very suitable to handle second-order data obtained from hyphenated instruments, such as HPLC-DAD, LC-MS, and GC-MS.

3.3.2. SWATLD

The SWATLD algorithm, as a derivative of ATLD, is also widely employed as it can yield better results in many cases. It alternately minimizes three objective functions with intrinsic relationship and also holds the characteristics of fast convergence and being insensitive to excessive component numbers. The detail explanations of these properties have been provided by authors in the original paper [17]. Three new residues can be expressed as:

$$\mathbf{B}^{+}\mathbf{X}_{i..} = diag\left(\mathbf{a}_{(i)}\right)\mathbf{C}^{\mathrm{T}} + \mathbf{B}^{+}\mathbf{E}_{i..}, \quad \mathbf{X}_{i..}(\mathbf{C}^{+})^{\mathrm{T}} = \mathbf{B}diag\left(\mathbf{a}_{(i)}\right) + \mathbf{E}_{i..}(\mathbf{C}^{+})^{\mathrm{T}}, \text{ for } i = 1, 2, \ldots, I, \tag{12}$$

$$\mathbf{C}^{+}\mathbf{X}_{.j.} = diag\left(\mathbf{b}_{(j)}\right)\mathbf{A}^{\mathrm{T}} + \mathbf{C}^{+}\mathbf{E}_{.j.}, \quad \mathbf{X}_{.j.}(\mathbf{A}^{+})^{\mathrm{T}} = \mathbf{C}diag\left(\mathbf{b}_{(j)}\right) + \mathbf{E}_{.j.}(\mathbf{A}^{+})^{\mathrm{T}}, \text{ for } j = 1, 2, \ldots, J, \tag{13}$$

$$\mathbf{A}^{+}\mathbf{X}_{..k} = diag\left(\mathbf{c}_{(k)}\right)\mathbf{B}^{\mathrm{T}} + \mathbf{A}^{+}\mathbf{E}_{..k}, \quad \mathbf{X}_{..k}(\mathbf{B}^{+})^{\mathrm{T}} = \mathbf{A}diag\left(\mathbf{c}_{(k)}\right) + \mathbf{E}_{..k}(\mathbf{B}^{+})^{\mathrm{T}}, \text{ for } k = 1, 2, \ldots, K, \tag{14}$$

By introducing some reasonable weight terms, three new objective functions are established and can be expressed as follows:

$$\begin{aligned} \mathrm{S}(\mathbf{A}) &= \sum_{i=1}^{I}\left\| \left(\mathbf{B}^{+}\mathbf{X}_{i..} - diag\left(\mathbf{a}_{(i)}\right)\mathbf{C}^{\mathrm{T}}\right)^{\mathrm{T}} \times diag\left(sqrt\left(\mathbf{1}./diagm\left(\mathbf{C}^{\mathrm{T}}\mathbf{C}\right)\right)\right) \right\|_{\mathrm{F}}^{2} \\ &+ \sum_{i=1}^{I}\left\| \left(\mathbf{X}_{i..}(\mathbf{C}^{\mathrm{T}})^{+} - \mathbf{B}diag\left(\mathbf{a}_{(i)}\right)\right) \times diag\left(sqrt\left(\mathbf{1}./diagm\left(\mathbf{B}^{\mathrm{T}}\mathbf{B}\right)\right)\right) \right\|_{\mathrm{F}}^{2}, \end{aligned} \tag{15}$$

$$S(\mathbf{B}) = \sum_{j=1}^{J}\left\|\left(\mathbf{C}^+\mathbf{X}_{.j.} - diag(\mathbf{b}_{(j)})\mathbf{A}^T\right)^T \times diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{A}^T\mathbf{A})\right)\right)\right\|_F^2$$

$$+ \sum_{j=1}^{J}\left\|\left(\mathbf{X}_{.j.}(\mathbf{A}^T)^+ - \mathbf{C}diag(\mathbf{b}_{(j)})\right) \times diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{C}^T\mathbf{C})\right)\right)\right\|_{F'}^2 \tag{16}$$

$$S(\mathbf{C}) = \sum_{k=1}^{K}\left\|\left(\mathbf{A}^+\mathbf{X}_{..k} - diag(\mathbf{c}_{(k)})\mathbf{B}^T\right)^T \times diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{B}^T\mathbf{B})\right)\right)\right\|_F^2$$

$$+ \sum_{k=1}^{K}\left\|\left(\mathbf{X}_{..k}(\mathbf{B}^+)^T - \mathbf{A}diag(\mathbf{c}_{(k)})\right) \times diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{A}^T\mathbf{A})\right)\right)\right\|_F^2. \tag{17}$$

Due to the unique optimizing strategy, this algorithm is more efficient than others. It can provide more satisfactory results than ATLD with moderate noise levels. Moreover, it can deal with the problem of moderate collinearity, but it is not so effective when data are collinear severely.

3.3.3. APTLD

The APTLD algorithm was developed by Xia et al. [18], and it can provide some improved properties. It alternately minimizes three new least squares-based objective functions by using the constraint functions as penalty terms of the PARAFAC error. Eqs. (12)–(14) are the new objective functions, which alternately used as the constraint terms. By introducing large penalty terms and combining them with residue functions (18)–(20) to establish three objective functions, APTLD transforms these constrained problems into non-constrained ones. Then, it alternately minimizes the following three objective functions to resolve the model:

$$S(\mathbf{A}) = \sum_{k=1}^{K}\left\|\mathbf{X}_{..k} - \mathbf{A}diag(\mathbf{c}_{(k)})\mathbf{B}^T\right\|_F^2 + q\left(\sum_{j=1}^{J}\left\|diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{B}^T\mathbf{B})\right)\right)\left(\mathbf{C}^+\mathbf{X}_{.j.} - diag(\mathbf{b}_{(j)})\mathbf{A}^T\right)\right\|_F^2\right.$$

$$+ \sum_{k=1}^{K}\left\|\left(\mathbf{X}_{..k}(\mathbf{B}^T)^+ - \mathbf{A}diag(\mathbf{c}_{(k)})\right)diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{C}^T\mathbf{C})\right)\right)\right\|_F^2\right), \tag{18}$$

$$S(\mathbf{B}) = \sum_{i=1}^{I}\left\|\mathbf{X}_{i..} - \mathbf{B}diag(\mathbf{a}_{(i)})\mathbf{C}^T\right\|_F^2 + r\left(\sum_{k=1}^{K}\left\|diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{C}^T\mathbf{C})\right)\right)\left(\mathbf{A}^+\mathbf{X}_{..k} - diag(\mathbf{c}_{(k)})\mathbf{B}^T\right)\right\|_F^2\right.$$

$$+ \sum_{i=1}^{I}\left\|\left(\mathbf{X}_{i..}(\mathbf{C}^T)^+ - \mathbf{B}diag(\mathbf{a}_{(i)})\right)diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{A}^T\mathbf{A})\right)\right)\right\|_F^2\right), \tag{19}$$

$$S(\mathbf{C}) = \sum_{j=1}^{J}\left\|\mathbf{X}_{.j.} - \mathbf{C}diag(\mathbf{b}_{(j)})\mathbf{A}^T\right\|_F^2 + p\left(\sum_{i=1}^{I}\left\|diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{A}^T\mathbf{A})\right)\right)\left(\mathbf{B}^+\mathbf{X}_{i..} - diag(\mathbf{a}_{(i)})\mathbf{C}^T\right)\right\|_F^2\right.$$

$$+ \sum_{j=1}^{J}\left\|\left(\mathbf{X}_{.j.}(\mathbf{A}^T)^+ - \mathbf{C}diag(\mathbf{b}_{(j)})\right)diag\left(sqrt\left(\mathbf{1}./diagm(\mathbf{B}^T\mathbf{B})\right)\right)\right\|_F^2\right), \tag{20}$$

where p, q, and r represent penalty factors. The performance of APTLD depends on the choice of the penalty factor values. When the values are very small, it will lead to a lot of iterations and sensitivity to excess factors, which is close to that of PARAFAC algorithm; particularly, when $p = q = r = 0$, APTLD can be regarded as a variant of PARAFAC. However, this algorithm will become insensitive to excess factors and speed up convergence when larger values of p, q, and r are selected. According to the variance among different trials and computational burdens, a further increase in p, q, and r values will make APTLD perform theoretically better. Therefore, its performance can be exquisitely improved by adjusting the penalty factors p, q, and r on the basis of particular circumstances and special needs.

3.3.4. APQLD

The APQLD algorithm [19] as an extension of APTLD for decomposition of quadrilinear data is applied to third-order calibration. Similar to APTLD, four objective functions can be obtained as:

$$
\begin{aligned}
\mathrm{S}(\mathbf{A}) = &\sum_{k=1}^{K}\sum_{l=1}^{L}\left\|\mathbf{X}_{..kl} - \mathbf{A}diag(\mathbf{d}_{(l)})diag(\mathbf{c}_{(k)})\mathbf{B}^{\mathrm{T}}\right\|_{\mathrm{F}}^{2} \\
&+ q(\sum_{k=1}^{K}\sum_{l=1}^{L}\left\|\left(\mathbf{X}_{..kl}(\mathbf{B}^{\mathrm{T}})^{+} - \mathbf{A}diag(\mathbf{d}_{(l)})diag(\mathbf{c}_{(k)})\right)sqrt(\mathbf{W}_{\mathrm{B}})\right\|_{\mathrm{F}}^{2} \\
&+ \sum_{j=1}^{J}\sum_{k=1}^{K}\left\|sqrt(\mathbf{W}_{\mathrm{D}})\left(\mathbf{D}^{+}\mathbf{X}_{jk..} - diag(\mathbf{c}_{(k)})diag(\mathbf{b}_{(j)})\mathbf{A}^{\mathrm{T}}\right)\right\|_{\mathrm{F}}^{2}),
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
\mathrm{S}(\mathbf{B}) = &\sum_{l=1}^{L}\sum_{i=1}^{I}\left\|\mathbf{X}_{i..l} - \mathbf{B}diag(\mathbf{a}_{(i)})diag(\mathbf{d}_{(l)})\mathbf{C}^{\mathrm{T}}\right\|_{\mathrm{F}}^{2} \\
&+ r(\sum_{l=1}^{L}\sum_{i=1}^{I}\left\|\left(\mathbf{X}_{i..l}(\mathbf{C}^{\mathrm{T}})^{+} - \mathbf{B}diag(\mathbf{a}_{(i)})diag(\mathbf{d}_{(l)})\right)sqrt(\mathbf{W}_{\mathrm{C}})\right\|_{\mathrm{F}}^{2} \\
&+ \sum_{k=1}^{K}\sum_{l=1}^{L}\left\|sqrt(\mathbf{W}_{\mathrm{A}})\left(\mathbf{A}^{+}\mathbf{X}_{..kl} - diag(\mathbf{d}_{(l)})diag(\mathbf{c}_{(k)})\mathbf{B}^{\mathrm{T}}\right)\right\|_{\mathrm{F}}^{2}),
\end{aligned}
\tag{22}
$$

$$
\begin{aligned}
\mathrm{S}(\mathbf{D}) = &\sum_{j=1}^{J}\sum_{k=1}^{K}\left\|\mathbf{X}_{jk.} - \mathbf{D}diag(\mathbf{c}_{(k)})diag(\mathbf{b}_{(j)})\mathbf{A}^{\mathrm{T}}\right\|_{\mathrm{F}}^{2} \\
&+ p(\sum_{j=1}^{J}\sum_{k=1}^{K}\left\|\left(\mathbf{X}_{jk.}(\mathbf{A}^{\mathrm{T}})^{+} - \mathbf{D}diag(\mathbf{c}_{(k)})diag(\mathbf{b}_{(j)})\right)sqrt(\mathbf{W}_{\mathrm{A}})\right\|_{\mathrm{F}}^{2} \\
&+ \sum_{i=1}^{I}\sum_{j=1}^{J}\left\|sqrt(\mathbf{W}_{\mathrm{C}})\left(\mathbf{C}^{+}\mathbf{X}_{ij..} - diag(\mathbf{b}_{(j)})diag(\mathbf{a}_{(i)})\mathbf{D}^{\mathrm{T}}\right)\right\|_{\mathrm{F}}^{2}),
\end{aligned}
\tag{23}
$$

where $\mathbf{W}_{\mathrm{A}} = diag(\mathbf{1}./diagm(\mathbf{A}^{\mathrm{T}}\mathbf{A}))$, $\mathbf{W}_{\mathrm{B}} = diag(\mathbf{1}./diagm(\mathbf{B}^{\mathrm{T}}\mathbf{B}))$, $\mathbf{W}_{\mathrm{C}} = diag(\mathbf{1}./diagm(\mathbf{C}^{\mathrm{T}}\mathbf{C}))$, and $\mathbf{W}_{\mathrm{D}} = diag(\mathbf{1}./diagm(\mathbf{D}^{\mathrm{T}}\mathbf{D}))$. APQLD algorithm decomposes the quadrilinear model by alternatively minimizing the four objective functions abovementioned. The performance of APQLD also

depends on the selection of the penalty factors p, q, r, and s. Obviously, it can be considered as a variant of the four-way PARAFAC when the four penalty factors equal to 0.

APQLD retains the second-order advantage possessed by second-order calibration and holds additional advantage. By introducing a new fourth mode, it can relieve the serious problem of collinearity, which cannot be solved by three-way algorithms.

3.4. Rank estimation

It is always an important and intractable problem to estimate chemical ranks (the number of factors or components) for the trilinear model before decomposing a three-way data array. Theoretically, it can be seemingly solved by selecting the appropriate algorithms, which are insensitive to the excessive component numbers (chemical ranks). Nevertheless, these algorithms also guarantee that the component number (chemical rank) chosen should be no fewer than the underlying one. As a matter of fact, when the component number selected is far more than the actual one, it may lead to a model fitting error and a large deviation for the predicted results. On the contrary, the performances of the algorithm on providing accurate solutions will be largely improved when the most appropriate factors are chosen in analytical system.

Based on this, a lot of methods have been developed for estimating the chemical ranks. In general, they can be roughly fallen into two main categories. The first one is on the basis of the trilinear model, which includes split-half analysis [20], Wu's maximum rank method [21], core consistency diagnostic (CORCONDIA) [22], ADD-ONE-UP [23], and self-weighted alternating trilinear decomposition and Monte Carlo simulation (SWATLD-MCS) [24]. The core of split-half analysis concerns a relatively complex splitting skill, and hence the result depends on splitting schemes greatly. CORCONDIA and ADD-ONE-UP are two of the most commonly used methods in determining the chemical ranks. However, they are quite time consuming sometimes. Furthermore, the severe collinearity data may also lead to a heavy computation burden and even get error results. Self-weighted alternating trilinear decomposition and Monte Carlo simulation (SWATLD-MCS) operate in two main steps. First of all, Monte Carlo simulation is applied to generated one pseudo three-way data array. Sorted mean relative concentration values can then be obtained by applying SWATLD to decompose the three-way data array created by MCS. By comparing the sorted mean relative concentration value, this method can determine the chemical rank. The other ones belong to nonmodel methods such as orthogonal projection approach (OPA) [25], two-mode subspace comparison (TMSC) [26], factor indicator function (IND) [27], subspace projection of pseudo high-way array (SPPH) [28], linear transform method incorporating Monte Carlo simulation (LTMC) [29], and region based on moving windows subspace projection technique (RMWSPT) [30]. Though all of the above methods can be applied to rank estimation, it is impossible to find one among them which can guarantee the correct results under all situations. Actually, more than one method is often utilized in analysis to ensure the accuracy of the analytical results [8, 15].

3.4.1. Maximum rank method

The maximum rank method was firstly proposed by Wu et al. [21] to estimate the chemical rank for ATLD and ATLD's variants, as the following form shows:

$$rank\ \underline{\mathbf{X}} = max\left\{rank\left(\mathbf{X}_{I \times JK}\right), rank\left(\mathbf{X}_{J \times KI}\right), rank\left(\mathbf{X}_{K \times IJ}\right)\right\}, \tag{24}$$

In practice, the number of factors will also be determined as follows:

$$rank\ \underline{\mathbf{X}} = max\left\{rank\left(\sum_{i=1}^{I}\mathbf{X}_{i..}\right), rank\left(\sum_{j=1}^{J}\mathbf{X}_{.j.}\right), rank\left(\sum_{k=1}^{K}\mathbf{X}_{..k}\right)\right\}, \tag{25}$$

$$rank\ \underline{\mathbf{X}} = max\left\{rank\left(\mathbf{X}_{I \times JK}\mathbf{X}_{I \times JK}^{\mathrm{T}}\right), rank\left(\mathbf{X}_{J \times KI}\mathbf{X}_{J \times KI}^{\mathrm{T}}\right), rank\left(\mathbf{X}_{K \times IJ}\mathbf{X}_{K \times IJ}^{\mathrm{T}}\right)\right\}, \tag{26}$$

where rank (.) denotes the numerical rank estimate of a matrix based on a singular value decomposition procedure with a default tolerance. This method is universal and suitable to be used in any instance and can get satisfactory results when estimating the chemical rank of the three-way data.

3.4.2. ADD-ONE-UP

ADD-ONE-UP was proposed by Chen et al. in [23] for determining the chemical rank. It operates by fitting two reconstructed three-way data arrays by PARAFAC with a gradually increasing component numbers and then determines the chemical rank by examining the residual sum of squares (SSR). The method is convenient and powerful, and some nonideal experimental conditions (such as slight collinearity and unknown backgrounds) can be handled.

1.   Unfold the obtained three-way data array X into a two-way data set $\mathbf{X}_{I \times JK}$.

2.   Decompose $\mathbf{X}_{I \times JK}$ by SVD, $\mathbf{X}_{I \times JK} = \mathbf{USV}^{\mathrm{T}}$.

3.   Define $\mathbf{X}_c = \mathbf{U}_c\mathbf{S}_c\mathbf{V}_c^{\mathrm{T}}$, $\mathbf{U}_c$ and $\mathbf{V}_c$ consist of the first c columns of $\mathbf{U}$ and $\mathbf{V}$, respectively; $\mathbf{S}_c$ is a diagonal matrix with diagonal elements equal to the first c diagonal elements of $\mathbf{S}$.

4.   Fold $\mathbf{X}_c$ into a three-way data array $\mathbf{X}_c$, then resolve it by PARAFAC with $N = c$ (c = 1, 2, 3, …,). The residual sum of squares is denoted by SSR_c.

5.   Repeat steps 3 and 4 until SSR_c reaches its minimum or satisfies the equations below: $SSR_{c1} < s_{c1}^2$ and $SSR_{c1 + 1} > s_{c1 + 1}^2$ and $SSR_{c1 + 2} > s_{c1 + 2}^2$ (s_i represents the ith diagonal element of matrix $\mathbf{S}$ and s_{c1}^2 denotes the variance obtained by the inclusion of c1th component in the truncating step).

6.   Unfold X in another dimension to obtain $\mathbf{X}_{IK \times J}$, then perform the same steps from 2 to 5 to get c2, which meets similar relationships like c1.

7.   The factor numbers applied in decomposing the trilinear data array X should be the smaller one between c1 and c2, i.e. F = min(c1, c2).

This method utilizes the eigenvalues of factor analysis and the residuals of trilinear decomposition. It can cope with nonideal experimental conditions like varying backgrounds and moderate collinearity. However, as it is based on the PARAFAC algorithm, ADD-ONE-UP has

some drawbacks. It is rather time consuming due to the need to run PARAFAC for many times. Furthermore, this method may suffer from a heavy computational burden by reason of two-factor degeneracies and may yield inaccurate results.

3.4.3. CORCONDIA

The principle of CORCONDIA is to assess the similarity between the superdiagonal array T and the least squares-fitted G with a gradually increasing number of components. CORCONDIA is defined as:

$$core\ consistency = 100 \times \left(1 - \frac{\sum_{d=1}^{N} \sum_{e=1}^{N} \sum_{f=1}^{N} \left(g_{def} - t_{def}\right)^2}{\sum_{d=1}^{N} \sum_{e=1}^{N} \sum_{f=1}^{N} t_{def}^2}\right), \qquad (27)$$

where g_{def} stands for the element of G, t_{def} represents the elements of T, and N denotes the number of factors in the model.

For an ideal trilinear model, g_{def} is equal to t_{def} and the value of core consistency will be equal to 100%. Usually, the model can be regarded as "very trilinear" as the value of the core consistency above 90%, whereas a value nearly 50% will indicate a problematic model, which contains both trilinear and non-trilinear variations. A value close to zero or even negative means that the model is not valid. Although it is an effective method, it suffers from the drawbacks of PARAFAC.

3.5. Some related fundamental issues

3.5.1. Chromatographic peak alignment procedure

Chromatographic peak alignment is a challenge in the field of complex system analysis by multiway calibration methods. Some methods for peak alignment have been developed based on the second-order instruments, which generate a matrix data for per sample. These methods [31], for example iterative target factor analysis coupled to COW (ITTFA-COW), rank minimization (RM), parallel factor analysis alignment, and other recently proposed methods based on multivariate curve resolution-alternating least squares, employ signals of two-way structure to align chromatographic peaks shifts. In theory, these methods are aimed at the alignment of local chromatographic regions and therefore satisfactory results can be obtained for the time shifts existed in the whole chromatogram. They can achieve accurate time alignment regardless of the presence of unknown interferences. Not long ago, Yu and co-workers developed a new algorithm for chromatographic peak alignment, derived from the famous rank minimization method. It aligns time shift among samples and then utilizes trilinear decomposition algorithm to interpret the overlapping chromatographic peaks to quantify target analytes [31].

Figure 8(A) depicts the graphical representation of the rank minimization method (RM). A significant advantage of this method is that alignment can be successfully carried out even when the potential interferences coeluted with the analyte of interest. To have a better view on
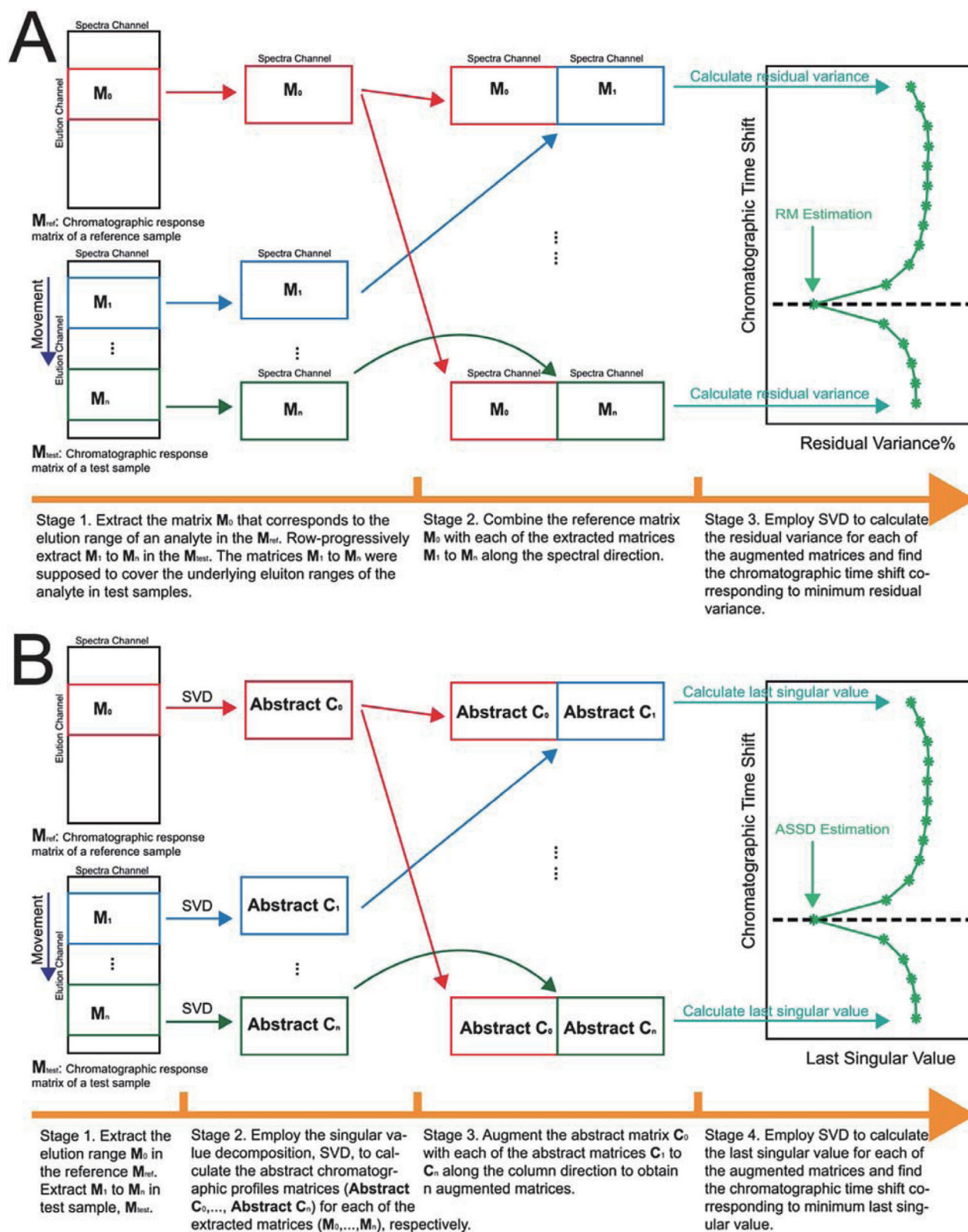
Figure 8. Graphical illustration of RM (A) and ASSD (B).

this method, a series of fixed-size time window (rectangles) along the retention time directions is applied in **Figure 8(A)**. In particular, the red rectangle $\mathbf{M_0}$ stands for the retention time range of analyte in the response of reference sample, and the retention time range between green and blue rectangles in the response of a test sample is the underlying time shift ranges of the analyte. By row-wisely moving the fixed-size time window on the test sample along the retention time direction, the rectangles from $\mathbf{M_1}$ to $\mathbf{M_n}$ can be extracted from the response of the test sample; then, augmented matrices, which are defined as $[\mathbf{M_0} \mid \mathbf{M_1}], \ldots, [\mathbf{M_0} \mid \mathbf{M_n}]$ [stage 2 in **Figure 8(A)**] in the retention time direction, can be obtained. Finally, the singular value decomposition is performed on these augmented matrices and results in a list of residual variance. Consequently, the percentage of the residual variance plotted against each chromatographic time shift will mark clearly the time shift point correction corresponding to the minimum residual variance.

The abstract subspace difference (ASSD) method uses abstract chromatographic profiles for alignment. Accordingly, the response matrix $\mathbf{X}$ can be expressed in the form of singular value decomposition (SVD) notations as follows:

$$\mathbf{X} = \mathbf{USV}^{\mathrm{T}} + \mathbf{E} \tag{28}$$

herein, the column vector $\mathbf{U}$ represents the abstract chromatographic profiles, while the V is the abstract spectra profiles; in the strict sense, all of them are not necessarily correspond to the real ones. Suppose that two data matrices have been collected: a reference data, $\mathbf{X_{ref}}$, which includes only one analyte, and a test data, $\mathbf{X_{test}}$, which collects the analyte together with other unknown interferences. Hence, based on the singular value decomposition, the abstract chromatographic profiles for reference and test samples can be acquired separately:

$$\mathbf{X}_{ref} = \mathbf{U}_{ref}\mathbf{S}_{ref}\mathbf{V}_{ref}{}^{\mathrm{T}} + \mathbf{E}_{ref}, \tag{29}$$

$$\mathbf{X}_{test} = \mathbf{U}_{test}\mathbf{S}_{test}\mathbf{V}_{test}{}^{\mathrm{T}} + \mathbf{E}_{test}. \tag{30}$$

In the ideal situation, no noise is present, and there is no time shift between reference and test samples. In this case, the mathematical rank of the augmented matrix $[\mathbf{U}_{ref} \mid \mathbf{U}_{test}]$ will be identical to that of $\mathbf{U}_{test}$. However, in the situations where the chromatographic retention time of the analyte is not the same for the reference and test samples, the mathematical rank of the augmented matrix, $[\mathbf{U}_{ref} \mid \mathbf{U}_{test}]$, will become larger than the actual ones. Therefore, the core of ASSD method is to look for the augmented matrix with minimum mathematical rank for alignment, which is the same as the rank minimization method, except that ASSD uses the abstract chromatographic profiles for alignment instead of the underlying ones.

Figure 8(B) shows the graphical illustration of the ASSD method. In order to calculate the abstract chromatographic profiles for each of the extracted matrices $\mathbf{M}_1$ to $\mathbf{M}_n$, an additional step, SVD, has been introduced in the Stage 1 of **Figure 8(B)**. Additionally, this new method uses the last singular value instead of the percentage of residual variance in the last stage to

represent time shift correction. In practical measurement, aligning time shift for target analyte between the reference and a test sample according to the critical criterion of the mathematical rank of the augmented matrix is impractical. However, the augmented matrix, $[\mathbf{U}_{ref} \mid \mathbf{U}_{test}]$, will become a seriously ill-conditioned matrix provided that the time shift has been successfully aligned. Hereby, chromatographic peak alignment can be transformed to find the most ill-conditioned augmented matrix among the augmented matrices as shown in the Stage 3 of **Figure 8(B)**. As the total variance is the sum of the squared elements of the augment matrix, $[\mathbf{U}_{ref} \mid \mathbf{U}_{test}]$, it will be a steady state value and equal to the column numbers. Hence, a smaller last singular value will definitely correspond to a more ill-conditioned matrix.

3.5.2. Background drift

Non-trilinear factors such as background drift is unavoidable sometimes in the chromatographic analysis due to the composition of gradient elution and/or nature of complicated matrices, which may lead to wrong analysis results by the aforementioned chemometric algorithms. Amigo and co-workers have summarized the intuitive graphics and mathematical models used in handling chromatographic data issues [32]. Multivariate curve resolution (MCR) methods are typical examples.

A chromatographic background drift correction strategy [33] was developed in 2007 by our group for LC × LC × DAD data. The core idea is to perform trilinear decomposition, which is based on the alternating trilinear decomposition (ATLD) algorithm for the instrumental response data. In analysis, the background drift can be eliminated by regarding it as an extra component or factor. This method uses trilinear decomposition to resolve the raw data, to extract, and subtract the background component from the raw data for acquisition of the signal of analytes with a flat baseline. A detailed schematic description on how to subtract the background drift from raw three-way chromatographic data is illustrated in **Figure 9**.

Recently, a method that uses orthogonal spectral signal projection (OSSP) to simultaneously solve various kinds of chromatographic background drift was studied [33]. The analytical results indicated that OSSP coupled with PARAFAC can be used for handling coelution and background drift problems in chromatographic analysis. It indicates that more accurate analysis results can be obtained, regardless of the presence of background drift and unknown interferences.

4. Application

Based on the "second-order or high-order advantages" provided by chemometrics methods, some actual applications have been developed for the analysis of pharmaceuticals, biological matrices, foods, cosmetics, environmental matrices, and others. Multiway calibration algorithms have been employed to enhance the selectivity and can obtain accurate predicted concentration of analyte(s) of interest free from interference of potential interfering matrix. These applications summarized in **Table 2** are reviewed in the following six aspects.

4.1. Pharmaceuticals

In this field, two or three drugs have been simultaneously detected in aqueous solution or Chinese traditional medicine. The data analyzed are second-order tensors, which are obtained by high performance liquid chromatography-photodiode array detection (HPLC-DAD).

Su et al. proposed a method for simultaneously quantifying the main effective constituents such as puerarin, daidzin, and daidzein in traditional Chinese medicine kudzuvine root by using HPLC-DAD with ATLD algorithm [34].

Nowadays, traditional Chinese medicine (TCM) plays an important role in the healthcare system. Thus, considerable attention has been paid to Chinese patent medicine (CPM), which generally consists of several TCMs and other ingredients. It is significantly important to quantify the constituents of CPM and plasma for pharmacological analysis. Liu et al. determined two effective constituents, costunolide and dehydrocostuslactone, in plasma sample and Chinese patent medicine Xiang Sha Yang Wei capsule by using HPLC-DAD coupled with alternating trilinear decomposition (ATLD) algorithm [35].

Besides, Ding et al. determined isoniazid and pyrazinamide by using HPLC-DAD coupled with three different second-order calibration algorithms including ATLD, alternating fitting residue (AFR), and self-weighted alternating trilinear decomposition (SWATLD). The results showed that all the three algorithms could be used for solving overlapped chromatograms and unknown interferences successfully, and the analysis results obtained from AFR were slightly better in this situation [36].
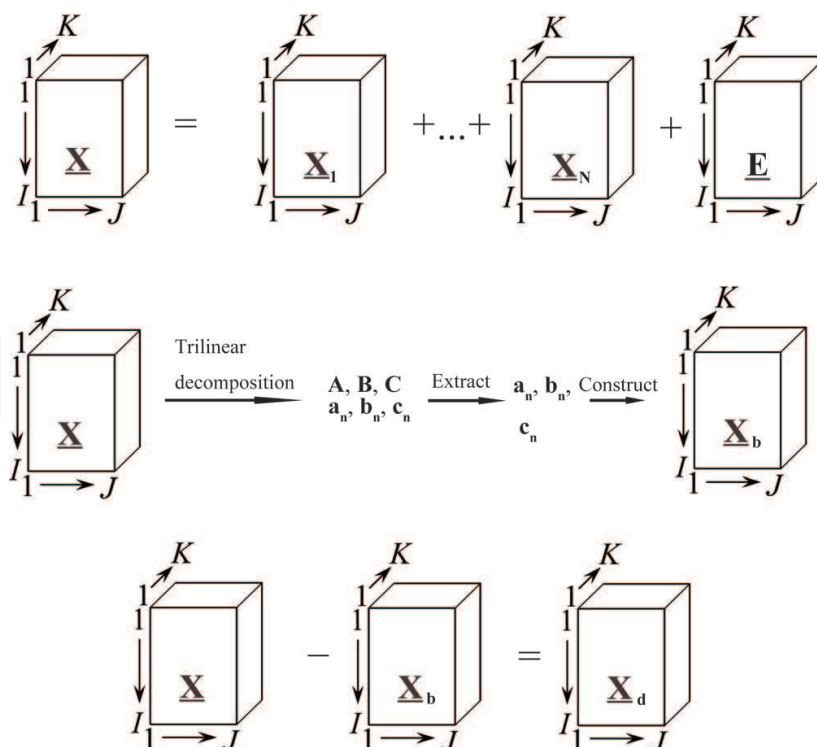

Figure 9. Schematic description on how to remove the background drift from three-dimensional instrumental data.

4.2. Biological matrices

Biological samples often contain various endogenous substances such as amino acids, hormones and neurotransmitters. Determining the concentrations of these molecules or metabolites is an integral part of clinical research and also helpful for understanding pathophysiology and mechanism of diseases. Human urine and plasma are commonly primary research systems.

High blood pressure, widely called hypertension, is a cardiac chronic disease with a symptom of sustaining rise in systemic arterial blood pressure. Zhao et al. carried out the simultaneous quantification of 11 antihypertensives, human serum, health product, and Chinese patent medicine samples by using HPLC-DAD with the aid of second-order calibration based on ATLD algorithm [37].

Tyrosine kinases are critical regulators of cell growth and differentiation growth and differentiation. The measurement of concentration of TKIs in different biofluids plays a significant role in optimizing the individual dosage regimen and reducing the risk of inapposite dosages. For the analysis of four tyrosine kinase inhibitors in different plasma samples, HPLC-DAD was utilized without absolutely chromatographic separations by resorting to ATLD algorithm. The contents of four tyrosine kinase inhibitors in different complex plasma samples can be accurately determined [38].

Liu et al. simultaneously determined vincristine, vinblastine, vindoline, catharanthine, and yohimbine in *Catharanthus roseus* and human serum samples utilizing ATLD algorithm to analyze the resulting three-way data array stacked by HPLC-DAD [39].

β-blockers are the first-line therapeutic agents for treating cardiovascular diseases and also a class of prohibited substances in athletic competitions. Therefore, rapid screening for multiple β-blockers in a single analysis has been of growing demand in clinical toxicology, forensic science, and antidoping control as well. Gu et al. proposed a smart strategy that combines three-way liquid chromatography-mass spectrometry (LC-MS) data with second-order calibration method based on alternating trilinear decomposition (ATLD) algorithm for simultaneous determination of 10 b-blockers in human urine and plasma samples [40]. The quantitative results were validated by the LC-MS/MS operated in multiple reaction monitoring (MRM) mode.

4.3. Foods

The applications in this field cover the analysis of contaminants, essential ingredients, and additives.

Synthetic phenolic antioxidants as food additives were successfully determined in edible vegetable oil by using HPLC-DAD and APTLD [9]. Some extraction procedures, in which the antioxidants of interest would be separated, is unnecessary and the 10 antioxidants can be eluted within 6 min.

Yin et al. proposed a smart strategy that combined HPLC-DAD with ATLD algorithm to solve varying interfering patterns from different chromatographic columns and sample matrices for the rapid simultaneous determination of six synthetic colorants in beverages with little sample pretreatment [1].

Type of data	Algorithm	Analytes	Ref.
Pharmaceuticals			
HPLC-DAD	ATLD	Puerarin, daidzin, and daidzein	[34]
HPLC-DAD	ATLD	Costunolide and dehydrocostuslactone	[35]
HPLC-DAD	ATLD, SWATLD, AFR	Isoniazid and pyrazinamide	[36]
Biological matrices			
HPLC-DAD	ATLD	Eleven antihypertensives	[37]
HPLC-DAD	ATLD	Four tyrosine kinase inhibitors	[38]
LC-MS	ATLD	Ten β-blockers	[40]
LC-MS	ATLD	Six antidiabetic agents	[48]
HPLC-DAD	ATLD	Five vinca alkaloids	[39]
Foods			
HPLC-DAD	PARAFAC, ATLD, SWATLD	Sudan I and Sudan II	[49]
HPLC-DAD	ATLD	Six synthetic colorants	[1]
HPLC-DAD	APTLD	Synthetic phenolic antioxidants	[9]
HPLC-DAD	ATLD,PCA	Eight coeluted compounds in tea	[7]
HPLC-DAD	ATLD	Eight flavonoids	[42]
HPLC-DAD	ATLD	nine polyphenols	[43]
HPLC-DAD	ATLD	Twelve quinolones	[44]
HPLC-DAD	ATLD, PCA-LDA	Thirteen phenolic compounds	[45]
HPLC-DAD	ATLD	Twelve polyphenols	[46]
LC-MS	ATLD	Ten mycotoxins	[47]
Environmental matrices			
HPLC-DAD	SWATLD	Three pre-emergence herbicides	[50]
HPLC	ATLD	1-Chloro-2,4-dinitrobenzene and 3,5-dinitrobenzoic acid	[51]
HPLC	ATLD	Five dimethylphenol isomers	[53]
HPLC	ATLD	Catechol, resorcinol and hydroquinone	[52]

Table 2. Reviewed applications.

Tea is one of the most widely consumed beverages in the world. The biological functions of tea have been reported in numerous studies, such as anti-inflammation, antiathero-sclerotic, antioxidant, anticarcinoma, antiobesity, and antiviral properties. These beneficial effects are related to the presence of purine alkaloids and polyphenols in tea. An attractive chemometrics-enhanced HPLC-DAD strategy was proposed by Yin et al. for simultaneous and fast determination of eight coeluted compounds including gallic acid, caffeine, and six catechins in 10 kinds of Chinese teas by using second-order calibration method based on ATLD algorithm [41]. Subsequently, based on the quantitative results, principal component analysis (PCA) was used to conduct a cluster analysis for these Chinese teas.

Propolis is a naturally occurring resinous hive product gathered by worker honeybees from buds and barks of different plant species. Sun et al. developed a fast analytical strategy by combining HPLC-DAD with ATLD algorithm for simultaneous determination of eight flavonoids in propolis capsule samples [42].

Honey is a wholesome natural food product well known for its high nutrition. The antioxidant ability of a number of honeys has been determined and found to be significantly correlated to the contents of polyphenols, which can affect the quality of honeys and their products beneficial for improving overall health and preventing some diseases. By using second-order calibration for development of HPLC-DAD method, Zhang et al. quantified nine polyphenols in five kinds of honey samples successfully [43]. Quinolones, a kind of antibacterial, which is widely used in agriculture for its high antimicrobial activity, were also detected by HPLC-DAD with ATLD algorithm in honey samples [44].

Wine phenolic compounds, as secondary metabolites and functional components, determine the important sensorial characteristics of wines, such as mouth-feel, fragrance, and color. The combination of HPLC-DAD and second-order calibration method based on ATLD has been used for the determination of 13 phenolic compounds in red wines, and linear discriminant analysis (PCA-LDA) was applied for distinguishing wines aged for years [45]. Similarly, the same strategy was carried out by Wang et al. for simultaneously quantify 12 polyphenols in different kinds of apple peel and pulp samples [46].

Mycotoxins are a class of highly carcinogenic substances often naturally occurring in the moldy foods, especially cereals. Liu et al. proposed a smart strategy that combines three-way LC-MS data with second-order calibration method based on ATLD algorithm for direct, fast, and interference-free determination of multiclass regulated mycotoxins in complex cereal samples [47]. Ten mycotoxins with different property could be fast eluted out and detected by full scanning MS with a segmented fragment program to enhance the sensitivity.

By using LC-MS in combination with second-order calibration method based on ATLD algorithm, Gu et al. simultaneously green determined six coeluted sulfonylurea-type oral antidiabetic agents in healthy herbal teas and human plasma samples [48]. The strategy proved to be a promising method for resolution and determination of coeluted multianalytes of interest in complex samples while avoiding elaborate sample pretreatment steps and complicated experimental conditions as well as more sophisticated high-cost instrumentations.

For the determination of Sudan dyes in hot chilli samples, HPLC-DAD was employed without completely chromatographic separations by using PARAFAC, ATLD, and SWATLD [49]. The low contents of Sudan I and Sudan II could be accurately determined in complex chilli mixtures.

4.4. Environmental matrices

In this field, we analyzed the analytes in aqueous solution, soil, tap water, river, and effluent water, mainly containing organic contaminants and pesticides.

Herbicides, which are chemicals often employed to kill weeds without causing injury to desirable vegetation, have been widely used. These may lead to their accumulation in the environment and cause continuous and serious pollution or even toxicity to crops and

humans. Qing et al. developed a novel strategy for analysis of three pre-emergence herbicides in environment samples using HPLC-DAD with SWATLD algorithm [50].

Chemometrics-assisted HPLC-DAD strategy has a great potential in analysis of target analytes in complex environmental matrices. So far, this strategy has been utilized for determination of 1-chloro-2,4-dinitrobenzene and 3,5-dinitrobenzoic acid [51], catechol, resorcinol, and hydroquinone [52] as well as five dimethylphenol isomers [53] in environment successfully.

5. Conclusion

This chapter scientifically describes in detail the various multiway chemometrics methodologies and applications in chromatography. We have built more canonical symbol systems, noted the inner mathematical cyclic symmetry property for multilinear decomposition, introduced several multiway calibration algorithms, explored the rank estimation of multiway data array, and analyzed numerous actual systems by homemade methods. Some fundamental issues related to chromatographic analysis such as peak alignment and background drift were also discussed and solved. By combining chromatographic techniques with chemometrics based on multiway calibration methods, complicated and tedious sample pretreatment can be greatly simplified and long chromatographic elution can be avoided. All the applications above-mentioned are universal, rapid, and sensitive for the determination of a variety of analytes in complex matrices.

Acknowledgements

Conflict of interest

There are no conflicts to declare.

Author details

Hai-Long Wu*, Xiao-Dong Sun, Huan Fang and Ru-Qing Yu

*Address all correspondence to: hlwu@hnu.edu.cn

State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha, China

References

[1] Yin XL, Wu HL, Gu HW, Hu Y, Wang L, Xia H, et al. Chemometrics-assisted high performance liquid chromatography-diode array detection strategy to solve varying interfering patterns from different chromatographic columns and sample matrices for beverage analysis. Journal of Chromatography A. 2016;**1435**:75-84

[2] Visky D, Haghedooren E, Dehouck P, Kovács Z, Kóczián K, Noszál B, et al. Facilitated column selection in pharmaceutical analyses using a simple column classification system. Journal of Chromatography A. 2006;**1101**(1):103-114

[3] Jandera P, Vyňuchalová K, Hájek T, Česla P, Vohralík G. Characterization of HPLC columns for two-dimensional LC× LC separations of phenolic acids and flavonoids. Journal of Chemometrics. 2008;**22**(3–4):203-217

[4] Haghedooren E, Farkas E, Kerner Á, Dragovic S, Noszál B, Hoogmartens J, et al. Effect of long-term storage and use on the properties of reversed-phase liquid chromatographic columns. Talanta. 2008;**76**(1):172-182

[5] Pérez RL, Escandar GM. Multivariate calibration-assisted high-performance liquid chromatography with dual UV and fluorimetric detection for the analysis of natural and synthetic sex hormones in environmental waters and sediments. Environmental Pollution. 2016;**209**:114-122

[6] Tan F, Tan C, Zhao A, Li M. Simultaneous determination of free amino acid content in tea infusions by using high-performance liquid chromatography with fluorescence detection coupled with alternating penalty trilinear decomposition algorithm. Journal of Agricultural and Food Chemistry. 2011;**59**(20):10839-10847

[7] Yin X-L, Wu H-L, Gu H-W, Zhang X-H, Sun Y-M, Hu Y, et al. Chemometrics-enhanced high performance liquid chromatography-diode array detection strategy for simultaneous determination of eight co-eluted compounds in ten kinds of Chinese teas using second-order calibration method based on alternating trilinear decomposition algorithm. Journal of Chromatography A. 2014;**1364**:151-162

[8] Wu HL, Li Y, Yu RQ. Recent developments of chemical multiway calibration methodologies with second-order or higher-order advantages. Journal of Chemometrics. 2014;**28**(5): 476-489

[9] Wang J-Y, Wu H-L, Chen Y, Sun Y-M, Yu Y-J, Zhang X-H, et al. Fast analysis of synthetic antioxidants in edible vegetable oil using trilinear component modeling of liquid chromatography–diode array detection data. Journal of Chromatography A. 2012;**1264**:63-71

[10] Braga JWB, Bottoli CB, Jardim IC, Goicoechea HC, Olivieri AC, Poppi RJ. Determination of pesticides and metabolites in wine by high performance liquid chromatography and second-order calibration methods. Journal of Chromatography A. 2007;**1148**(2): 200-210

[11] Fleming C, Kowalski B. Encyclopedia of Analytical Chemistry. New York: John Wiley& Sons Inc; 2000. pp. 9737-9764

[12] Malinowski ER. Factor Analysis in Chemistry, 3rd Ed. Technometrics. 2002;**36**(1):180-181

[13] Harshman RA. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-model factor analysis. Ucla Working Papers in Phonetics. 1970;**16**

[14] Carroll JD, Pruzansky S, Kruskal JB. CANDELINC: A general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. Psychometrika. 1980; **45**(1):3-24

[15] Wu H-L, Nie J-F, Yu Y-J, Yu R-Q. Multi-way chemometric methodologies and applications: A central summary of our research work. Analytica Chimica Acta. 2009;**650**(1):131-142

[16] Fleming CM, Kowalski BR. Second-order calibration and higher. Encyclopedia of Analytical Chemistry: Applications. Theory and Instrumentation. 2006

[17] Chen Z-P, Wu H-L, Jiang J-H, Li Y, Yu R-Q. A novel trilinear decomposition algorithm for second-order linear calibration. Chemometrics and Intelligent Laboratory Systems. 2000; **52**(1):75-86

[18] Xia AL, Wu HL, Fang DM, Ding YJ, Hu LQ, Yu RQ. Alternating penalty trilinear decomposition algorithm for second-order calibration with application to interference-free analysis of excitation–emission matrix fluorescence data. Journal of Chemometrics. 2005;**19**(2): 65-76

[19] Xia AL, Wu HL, Li SF, Zhu SH, Hu LQ, Yu RQ. Alternating penalty quadrilinear decomposition algorithm for an analysis of four-way data arrays. Journal of Chemometrics. 2007; **21**(3–4):133-144

[20] Law HG. Research methods for multimode data analysis. Praeger. 1984

[21] Wu HL, Shibukawa M, Oguma K. An alternating trilinear decomposition algorithm with application to calibration of HPLC–DAD for simultaneous determination of overlapped chlorinated aromatic hydrocarbons. Journal of Chemometrics. 1998;**12**(1):1-26

[22] Bro R, Kiers HA. A new efficient method for determining the number of components in PARAFAC models. Journal of Chemometrics. 2003;**17**(5):274-286

[23] Chen Z-P, Liu Z, Cao Y-Z, Yu R-Q. Efficient way to estimate the optimum number of factors for trilinear decomposition. Analytica Chimica Acta. 2001;**444**(2):295-307

[24] Li Y, Wu H-L, Qing X-D, Zuo Q, Chen Y, Yu R-Q. A novel method to estimate the chemical rank of three-way data for second-order calibration. Chemometrics and Intelligent Laboratory Systems. 2013;**127**:177-184

[25] Sanchez FC, Toft J, Van den Bogaert B, Massart D. Orthogonal projection approach applied to peak purity assessment. Analytical Chemistry. 1996;**68**(1):79-85

[26] Xie H-P, Jiang J-H, Long N, Shen G-L, Wu H-L, Yu R-Q. Estimation of chemical rank of a three-way array using a two-mode subspace comparison approach. Chemometrics and Intelligent Laboratory Systems. 2003;**66**(2):101-115

[27] Malinowski ER. Determination of the number of factors and the experimental error in a data matrix. Analytical Chemistry. 1977;**49**(4):612-617

[28] Xia A-L, Wu H-L, Zhang Y, Zhu S-H, Han Q-J, Yu R-Q. A novel efficient way to estimate the chemical rank of high-way data arrays. Analytica Chimica Acta. 2007;**598**(1):1-11

[29] Hu L-Q, Wu H-L, Jiang J-H, Han Q-J, Xia A-L, Yu R-Q. Estimating the chemical rank of three-way data arrays by a simple linear transform incorporating Monte Carlo simulation. Talanta. 2007;**71**(1):373-380

[30] Nie J-F, Wu H-L, Wang J-Y, Liu Y-J, Yu R-Q. The chemical rank estimation for excitation-emission matrix fluorescence data by region-based moving window subspace projection technique and Monte Carlo simulation. Chemometrics and Intelligent Laboratory Systems. 2010;**104**(2):271-280

[31] Amigo JM, Skov T, Bro R. ChroMATHography: Solving chromatographic issues with mathematical models and intuitive graphics. Chemical Reviews. 2010;**110**(8):4582-4605

[32] Zhang Y, Wu H-L, Xia A-L, Hu L-H, Zou H-F, Yu R-Q. Trilinear decomposition method applied to removal of three-dimensional background drift in comprehensive two-dimensional separation data. Journal of Chromatography A. 2007;**1167**(2):178-183

[33] Yu Y-J, Wu H-L, Fu H-Y, Zhao J, Li Y-N, Li S-F, et al. Chromatographic background drift correction coupled with parallel factor analysis to resolve coelution problems in three-dimensional chromatographic data: Quantification of eleven antibiotics in tap water samples by high-performance liquid chromatography coupled with a diode array detector. Journal of Chromatography A. 2013;**1302**:72-80

[34] Su Z-Y, Wu H-L, Liu Y-J, Xu H, Zhang J, Nie C-C, et al. Simultaneous determination of main effective constituents in traditional Chinese medicine Kudzuving root using HPLC-DAD coupled with second-order calibration based on alternating trilinear decomposition. Acat Chimica Sinica;**2011**(4):459-464

[35] Liu Y, Wu H, Zhu S, Kang C, Xu H, Su Z, et al. Rapid determination of costunolide and dehydrocostuslactone in human plasma sample and Chinese patent medicine Xiang Sha Yang Wei capsule using HPLC-DAD coupled with second-order calibration. Chinese Journal of Chemistry. 2012;**30**(5):1137-1143

[36] Ding Y, Wu H, Xia A, Cui H, Yu R. Simultaneous determination of two antituberculosis drugs using alternating fitting residue algorithm combined with HPLC-DAD. Journal of Analytical Science. 2008;**24**(1):1

[37] Zhao J, Wu H-L, Niu J-F, Yu Y-J, Yu L-L, Kang C, et al. Chemometric resolution of coeluting peaks of eleven antihypertensives from multiple classes in high performance

liquid chromatography: A comprehensive research in human serum, health product and Chinese patent medicine samples. Journal of Chromatography B. 2012;**902**:96-107

[38] Xiang SX, Kang C, Xie LX, Yin XL, Gu HW, Yu RQ. Fast quantitative analysis of four tyrosine kinase inhibitors in different human plasma samples using three-way calibration-assisted liquid chromatography with diode array detection. Journal of Separation Science. 2015;**38**(16):2781-2788

[39] Liu Z, Wu H-L, Li Y, Gu H-W, Yin X-L, Xie L-X, et al. Rapid and simultaneous determination of five vinca alkaloids in Catharanthus roseus and human serum using trilinear component modeling of liquid chromatography–diode array detection data. Journal of Chromatography B. 2016;**1026**:114-123

[40] Gu H-W, Wu H-L, Yin X-L, Li Y, Liu Y-J, Xia H, et al. Multi-targeted interference-free determination of ten β-blockers in human urine and plasma samples by alternating trilinear decomposition algorithm-assisted liquid chromatography–mass spectrometry in full scan mode: Comparison with multiple reaction monitoring. Analytica Chimica Acta. 2014;**848**:10-24

[41] Yin XL, Wu HL, Gu HW, Zhang XH, Sun YM, Hu Y, et al. Chemometrics-enhanced high performance liquid chromatography-diode array detection strategy for simultaneous determination of eight co-eluted compounds in ten kinds of Chinese teas using second-order calibration method based on alternating trilinear decomp. Journal of Chromatography A. 2014;**1364**:151

[42] Sun Y-M, Wu H-L, Wang J-Y, Liu Z, Zhai M, Yu R-Q. Simultaneous determination of eight flavonoids in propolis using chemometrics-assisted high performance liquid chromatography-diode array detection. Journal of Chromatography B. 2014;**962**:59-67

[43] Zhang X-H, Wu H-L, Wang J-Y, Tu D-Z, Kang C, Zhao J, et al. Fast HPLC-DAD quantification of nine polyphenols in honey by using second-order calibration method based on trilinear decomposition algorithm. Food Chemistry. 2013;**138**(1):62-69

[44] Yu Y-J, Wu H-L, Shao S-Z, Kang C, Zhao J, Wang Y, et al. Using second-order calibration method based on trilinear decomposition algorithms coupled with high performance liquid chromatography with diode array detector for determination of quinolones in honey samples. Talanta. 2011;**85**(3):1549-1559

[45] Liu Z, Wu HL, Xie LX, et al. Direct and interference-free determination of thirteen phenolic compounds in red wines using a chemometrics-assisted HPLC-DAD strategy for authentication of vintage year. Analytical Methods. 2017;**9**(22):3361-3374

[46] Wang T, Wu HL, Xie LX, Zhu L, Liu Z, Sun XD, et al. Fast and simultaneous determination of 12 polyphenols in apple peel and pulp by using chemometrics-assisted high-performance liquid chromatography with diode array detection. Journal of Separation Science. 2017;**40**(8):1651-1659

[47] Liu Z, Wu H-L, Xie L-X, Hu Y, Fang H, Sun X-D, et al. Chemometrics-enhanced liquid chromatography-full scan-mass spectrometry for interference-free analysis of multi-class

mycotoxins in complex cereal samples. Chemometrics and Intelligent Laboratory Systems. 2017;**160**:125-138

[48] Gu H-W, Wu H-L, Li S-S, Yin X-L, Hu Y, Xia H, et al. Chemometrics-enhanced full scan mode of liquid chromatography–mass spectrometry for the simultaneous determination of six co-eluted sulfonylurea-type oral antidiabetic agents in complex samples. Chemometrics and Intelligent Laboratory Systems. 2016;**155**:62-72

[49] Zhang Y, Wu H-L, Xia A-L, Han Q-J, Cui H, Yu R-Q. Interference-free determination of Sudan dyes in chilli foods using second-order calibration algorithms coupled with HPLC-DAD. Talanta. 2007;**72**(3):926-931

[50] Qing X-D, Wu H-L, Li Y-N, Nie C-C, Wang J-Y, Zhu S-H, et al. Simultaneous determination of pre-emergence herbicides in environmental samples using HPLC-DAD combined with second-order calibration based on self-weighted alternating trilinear decomposition algorithm. Analytical Methods. 2012;**4**(3):685-692

[51] Shen H-B, Yang J, Liu X-J, Chou K-C. Using supervised fuzzy clustering to predict protein structural classes. Biochemical and Biophysical Research Communications. 2005;**334**(2): 577-581

[52] Sun J, Wu H, Mo C, Lu J, Cui H, Yu R. Simultaneous determination of isomers of dihydroxybenzene using alternating trilinear decomposition algorithm combined with reversed-phase high performance liquid chromatography/diode array detection. Chinese Journal of Chromatography. 2002;**20**(5):385-389

[53] Lu J-Z, Wu H-L, Sun X-Y, Cui H, Sun J-Q, Yu R-Q. Simultaneous decomposition and determination of the complex dimethylphenol isomers by alternating trilinear decomposition algorithm combined with region selection. Chinese Journal of Analytical Chemistry. 2004;**32**(10):1278-1282