# Interpretation of comprehensive two-dimensional gas-chromatography data using advanced chemometrics

Zhongda Zeng [a, b], Jia Li [b], Helmut M. Hugel [c], Guowang Xu [b], Philip J. Marriott [1, *]

[a] *Australian Centre for Research on Separation Science, School of Chemistry, Monash University, Wellington Rd, Clayton 3800, Australia*

[b] *CAS Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China*

[c] *School of Applied Sciences, RMIT University, GPO Box 2476, Melbourne 3001, Australia*

## HIGHLIGHTS

- We present a comprehensive survey of GC×GC and GC×GC-MS data with chemometrics
- This review includes principles, theories and graphical tools for data processing
- We discuss deconvolution of 1D, coupled and GC×GC separations with FID/MS
- We consider retention, structure, $t_R$ shifts, orthogonality and image analysis

## ABSTRACT

The power of comprehensive two-dimensional gas chromatography (GC×GC) for the study of complex mixtures has been indisputably proved in the past several decades. This review encompasses the whole of GC×GC-related data processing and summarizes relevant applications. We include theoretical introduction to some specific methods and studies to aid readers' understanding of chemometrics strategies for advanced data interpretation.

*Keywords:*
Chemometrics
Comprehensive two-dimensional gas chromatography (GC×GC)
Data characteristics
Data processing
Deconvolution
Mass spectrometry
Multivariate curve resolution (MCR)
Multivariate data analysis
Orthogonality
Signal processing

* Corresponding author. Tel.: + 61 3 99059630; Fax: + 61 3 99058501.
*E-mail address:* philip.marriott@monash.edu (P. Marriott)

**Contents**

## 1. Introduction

Comprehensive two-dimensional gas chromatography (GC×GC) is a natural extension in the panoply of development of conventional separations, including single-dimensional chromatography (SDC) and heart-cut (H/C) techniques. H/C is also called multidimensional gas chromatography (MDGC), defined as "the process of selecting a (limited) region or zone of eluted compounds issuing from the end of one GC column, and subsequently subjecting the zone to a further GC displacement" [1].

In the 1990s, a typical example of MDGC separation comprised about 50 individual sampled regions each with around 20 peaks for pattern recognition of pharmaceuticals [2]. In the past decade, the number of "cross-samples" investigated in a specific study and the number of peaks contained in each sample significantly increased (e.g., to 1000 and 500, respectively), because of biology-driven studies, such as proteomics and metabolomics.

Our recent work reported that 2771 compounds were found in an investigation of a Chinese medicinal formulation (CMF) that included nine single herbs, using the platform of GC×GC with time-of-flight mass spectrometry (GC×GC-TOF-MS) [3]. The ability to resolve a sample of such complexity is an evident challenge or even an impossible task for SDC analysis.

However, successful applications exploited the power of GC×GC related techniques to compositions in mixtures of high complexity, such as herbal medicines and drugs, flavors, foods, petroleum and biological samples [4–11]. The outcomes and the performance of this technique have been introduced and reviewed, with frequent updates [5,12–20]. The basic experiment comprises the connection of two chromatographic columns with complementary polarity that together enhance the separation capacity of the arrangement; the columns are interfaced through a modulator device, which effectively decouples elution on each column [21,22]. The column set pairs two columns that are most often

defined as comprising a low polarity (LP) / polar (P) combination, a moderately polar (MP) / polar combination, a P / LP combination, or a P / MP combination. Note, however, that these are relative properties, since a very polar / P combination may perform similarly to a P / LP combination. Such smart configurations help to separate and to re-arrange further the peaks in the first dimension ($^1$D) compared to the second one ($^2$D), with a fixed modulation period ($P_M$) and same total analysis time [23,24]. Thus, peak capacity, a theoretical measure of the number of peaks that can be separated in the 2D space, can then ideally attain to the product $^1n_c \times {}^2n_c$, assuming the peak capacities in $^1$D and $^2$D separations are $^1n_c$ and $^2n_c$, respectively [25]. This is the essential advantage of GC×GC, enabling the investigation – and separation – of samples with hundreds or even thousands of chemical components in contrast to SDC and MDGC techniques.

Unlike the conventional data structures of SDC, MDGC and coupling of chromatographic and spectral instruments, GC×GC data have two special properties:

(1) 2D characteristics with specific retention properties and response in $^1$D and $^2$D dimensions; and,
(2) loss of raw chromatographic data in $^1$D, but continuous modulation of fractions in $^2$D for each $^1$D peak.

For GC×GC-MS data with different mass analyzers, such as quadrupole and TOF, ideally a single-component mass spectrum can be detected at each retention-time ($t_R$) measurement point throughout the 2D GC×GC separation plane. This effectively expands the original data to a three-dimensional (3D) data set, with $t_R$ in both $^1$D and $^2$D ($^1t_R$ and $^2t_R$, respectively), and spectral intensity at the scanned *m/z*, comprising the x-, y-, and z-axes, respectively. Further, time-dependent and sample-to-sample dynamic variations complicate data processing and information extraction (extended to a four-dimensional arrangement) (e.g., metabolite fingerprinting analysis in metabolomics analysis with evolution of treatment or environmental effects over time). One of the typical examples is correction of $t_R$ shifts among different but related samples of GC×GC-TOF-MS, or different types of detectors on the basis of GC×GC separation [26,27].

The complexity of GC×GC related data and high-throughput analysis for real mixtures make chemometrics widely applicable to this area [28–31], which has the power to expose buried information in "white, grey and black systems" with different degrees of prior knowledge of multi-components, and draws on multivariate statistics, mathematics and computer science [32], as shown in Fig. 1. Many of the reported reviews of GC×GC incorporate the relevance of chemometrics for the investigation of GC×GC data, and include theoretical development and novel applications [28,29,33]. This work further explores the nexus between GC×GC and chemometrics to mine out hidden information with mathematical interpretation, and aims to provide extra understanding to the researcher without a chemometrics background. Previously reported chemometrics tools for processing of coupled data are introduced to explain GC×GC data, such as multivariate curve resolution (MCR) for bilinear data decomposition based on the principles of the Beer-Lambert Law (BLL). In terms of the 2D, or even 3D, data characteristics introduced above, we review some specific research insights of GC×GC, such as orthogonality and image analysis.

First, chemometrics methods to deconvolute overlapping GC×GC peak clusters in $^1$D and $^2$D separations are introduced by using model or fitting techniques. Based on the 2D feature of GC×GC separation, conventional deconvolution methods for 2D or 3D data processing have been applied for GC×GC processing. This further helps to recover lost information of primary peaks. Second, MCR methods based on single or multiple runs are separately summarized, to extract chromatographic data and spectral profiles of pure components from GC×GC-TOF-MS data to support identification and quantification

[34,35]. Four important chemometrics methods for 2-way and 3-way data resolution are introduced in theory, with worked examples of processed GC×GC related data, including heuristic evolving latent projection (HELP) [36,37], parallel factor analysis (PARAFAC) [38,39], MCR-alternating least squares (MCR-ALS) [40,41], and alternative moving window factor analysis (AMWFA) [42,43].

Next, some new research topics applied to GC×GC data are expounded, exploiting the 2D separation characteristics and matrix data structure, such as $t_R$ alignment, orthogonality and image analysis. Last, but not least, some routine considerations of the GC×GC experiment related to data processing aided by chemometrics are reviewed, such as peak detection, experimental design and optimization, signal processing, and component-calibration models. This should familiarize the reader with an appreciation of various chemometrics tools for presentation and interpretation of GC×GC and GC×GC-MS data.

In addition, some commercial and freely downloadable programs or software for GC×GC data analysis are introduced and can be readily used following instructions [44–46]. This includes signal-to-noise filtering, baseline correction, retention-time alignment, normalization, peak picking, deconvolution, integration, and library searching and identification by using retention-index and MS libraries [47,48]. This allows chemometrics strategies to be readily employed by researchers with limited chemometrics experience, such as ChromaToF data-processing software (Leco, USA). However, discussion of this is not included in the present review. The nomenclature and the conventions used here follow our previous recommendations, recently updated [23,24].

## 2. GC×GC data deconvolution

Deconvolution of unresolved chromatographic peaks largely extends experimental capability by using chemometrics tools. It can save time, labor and money to acquire the information about target analytes for identification and quantification with existing data at hand [35]. Though some analytical scientists prefer to use classical formulae or may be reluctant to apply chemometrics, previously reported applications introduce the effectiveness of these methods. This approach can overcome limitations of analytical instruments and/or insufficiently optimized conditions, and meets the demand to process complicated mixtures, such as biological fluids. If the chromatogram of pure components can be independently extracted from contaminated (overlapping) clusters, peak area or height can be applied for relative quantification, and retention time to determine the retention index for identification [49,50], especially for SDC analysis without MS detection. After this, absolute quantification can be achieved with the help of standard calibration strategies, including:

(1) normalization method;
(2) internal standard method;
(3) external standard method; or,
(4) standard addition method.

Operational details can be followed *via* the summarized review work [35].

GC×GC separation amplifies the $^1$D chromatogram to form a 2D contour plot. Deconvolution can be divided into two parts, namely, full deconvolution of overlapping peaks in $^2$D, and recovery of chromatographic profile in $^1$D, in terms of Equation (1) for mass transfer from the $^1$D column to $^2$D. Deconvolution could also be achieved for $^1$D overlapping peak clusters after individual optimization of the chromatogram.

$$A_{i,total} = \sum_{i=1}^{i} A_i = \int_0^{t_0+(i-1)\times P_M} \mathbf{S}(t)dt \qquad i = 1, 2, \ldots, n_M \qquad (1)$$

where, notations $A_i$ and $A_{i,total}$, respectively, denote peak area of the i$^{th}$ fraction detected in $^2$D and total area of the fractions from the 1$^{st}$ to the i$^{th}$. Term **S**(t) is the signal profile of the primary peak with change of retention time. The three numbers $t_0$, $P_M$ and $n_M$ represent the heart-cut (H/C) position of the first modulated fraction, modulation period and number detected for a given $^1$D peak. Here, H/C refers to the repetitive sampling event of the first column peak.

For deconvolution of a chromatogram with multiple components in $^2$D, a range of conventional methods can be employed because of the similar characteristics of separation by SDC. This has been widely reported over several decades. The two techniques of non-linear least-squares (NLLS) analysis [51,52] and Fourier transform (FT) [53,54] generated many publications in this area. The former uses curve fitting and predefined functions of the chromatographic profile, and the latter is a reversible process of mathematical transformation from raw signal to complex frequencies, and then inverse deconvolution to pure peaks *via* FT operation. Another choice was to optimize the target chromatogram with an iterative technique and constrained conditions of peaks, such as unimodality and non-negativity [55,56]. Among the mathematical models to simulate the chromatogram, the polynomial modified Gaussian (PMG) function has been found most suitable to derive each eluting component [51,57]. Non-linear regression techniques for deconvolution were reviewed by Vivo-Truyols et al. [58], who proposed a new algorithm for alternative fitting of the original signal and the second derivatives. These deconvolution methods included Powell-1 and -2, multi-start local search (MSLS) and locally optimized genetic algorithm (LOGA) [59,60]. Globally optimal performance was achieved through automated selection of the most effective calculation procedure employing first-order multivariate selectivity. This enables users without strong mathematical background to process their data.

In addition, strategies to process data obtained from high-performance liquid chromatography (HPLC) and capillary electrophoresis (CE) [61,62], or a multi-batch approach, combined with multiple-wavelength chromatograms for 1D data analysis, may be extended to GC deconvolution [63].

However, characteristic information about each fraction in $^2$D can be found if a component is present in more than one modulated fraction of the primary peak. The sequential chromatograms may be chemometrically interrogated to develop new strategies for peak deconvolution, including simultaneous derivation of $^1$D and $^2$D chromatographic profiles of pure components. For instance, Zeng et al. [64] developed a method for simultaneous deconvolution and re-construction of primary and secondary overlapping peak clusters in GC×GC analysis. Non-linear least squares curve fitting (NLLSCF) was employed to optimize the $^2$D chromatogram of pure components with a selective elution window, and then each area obtained from the corresponding modulated fraction of the primary peak was further used to simulate the $^1$D peak and its peak-profile data. According to Equation (1), the principle for mass transfer from $^1$D to the $^2$D column allows recovery of individual overlapping peaks in the $^1$D separation. Kong et al. [65] investigated this issue through full parameter search using the same principle introduced above. The linear relationships of parameters of the exponentially modified Gaussian (EMG) model and the corresponding $t_R$ values may be used for deconvolution of primary peaks in GC×GC analysis. In general, these methods were established to deconvolute $^1$D overlapping peaks after obtaining the pure chromatographic profile in $^2$D. Thus, the conventional algorithms for SDC deconvolution are suitable for the study of modulated

fractions in $^2$D, and then simulation of $^1$D peaks generated by GC×GC.

The 2D characteristics of GC×GC data mean that conventional chemometrics tools for multi-run-based data resolution may be introduced for deconvolution. For example, PARAFAC and PARAFAC2 methods were simultaneously utilized for quantification of kerosene in gasoline with GC×GC analysis [66]. MCR-ALS was developed by Tauler et al. to resolve second-order data from more than one run. It has been utilized to process GC×GC data with flame-ionization detection (GC×GC-FID) [67] and LC×LC data [68]. The raw chromatograms were unfolded to 2D GC×GC structure similar to coupled data with chromatographic and spectral information. Correlation of the concentration of analytes with the response of pure components permitted quantitative analysis of essential oils in perfume. The results indicated the effectiveness of the MCR-ALS method to resolve target chromatograms from complicated mixtures and, further, to build multivariate models of GC×GC-FID separation. The generalized rank-annihilation method (GRAM), also based on more than one experimental run, was successfully applied to study GC×GC quantification [69]. After alignment of $t_R$s, it was utilized for analyte identification and accurate quantification of unresolved analyte peaks by Fraga et al [69]. These algorithms are introduced in detail in sub-section 3.2.2, including PARAFAC, MCR-ALS and GRAM methods.

## 3. MCR of overlapping peaks in GC×GC-MS

MCR was traditionally defined as "a model-free or a soft-modeling method that focuses on describing the evolution of the experimental multi-component measurements through their pure component contributions" [70]. The goal of MCR is to extract the pure chromatographic and spectral profiles **C** and **S** shown in Equation (2) from the raw data matrix **X**. The bold letter **E** denotes residual errors or experimental noise not explained by chemical components. If successful, qualitative and quantitative analysis of analytes can then be attained with simultaneous presence of co-eluting components. This is the mathematical definition of BLL with bilinear basis for MCR analysis, according to the principles given in Fig. 2.

$$\mathbf{X} = \mathbf{C}\mathbf{S}^T + \mathbf{E} \quad (2)$$

For multi-run-based data resolution, the goal is the same but attempts to employ new information provided by the additional data. The assumptions of MCR analysis include two parts:

(1) fundamental characteristics of signal (experimental data and peaks), such as bilinearity, non-negativity and unimodality; and,
(2) each component has an experimentally identical spectrum in different runs.

In **Fig. 3**, a flowchart to summarize the sampling process is given, to sketch the whole picture for MCR with chromatographic and spectral contributions.

In substance, GC×GC-MS data are an extension of conventional coupled data, such as GC-MS and HPLC-DAD (diode-array detector), though it has 3D or even higher attributes (refer to Introduction). Thus, all the reported resolution methods based on single-run data and multi-run data can then be applied for GC×GC-MS data processing. They are summarized below. However, modulation of a primary peak to fractions and further separation in $^2$D make data of these types complicated. For example, re-arrangement of the fractions with spectral information may help resolution of overlapping peak clusters with high complexity, since molecularly-specific information

among the sub-peaks in $^2$D may be helpful. This should be one of the most important study updates of GC×GC-MS in contrast to previous coupled chromatographic-spectral datasets.

### *3.1. Single-run-based data resolution*

Manne proposed resolution theorems to introduce underlying premises to obtain correctly target matrices **C** and **S** for the resolution of a single data set [71]. The two key points of the theorems were originally stated as follows:

(1) if all interfering compounds that appear inside the concentration window of a given analyte also appear outside this window, it is possible to calculate the concentration profile of the analyte; and,

(2) if, for every interferent the concentration window of the analyte has a sub-window where the interferent is absent, then it is possible to calculate the spectrum of the analyte.

This is the basis to develop new chemometrics methods for single coupled data resolution. In reported works, not so many applications can be found relevant to the GC×GC-MS process by using single-run-based data-resolution methods. It is suitable, of course, to mine out information about components buried in complicated datasets in terms of the modulation arrangement of GC×GC separation and principles for MCR analysis. The task includes stepwise processing of each modulation fraction following the steps shown in Fig. 3. However, the resolved matrices **C** and **S** are not completely the same within the sampled region of primary column eluate because of component overlap and H/C with a constant time window according to the $P_M$ value.

In summary, single-run-based methods can be divided into two categories with unique resolution and rational resolution, respectively.

#### *3.1.1. Unique resolution methods*

Employment of local rank information (LRI) is the crucial advantage of this type of method. This helps to extract $\mathbf{C}_i$ and $\mathbf{S}_i$ from each pure component with uniqueness property. LRI means the unique characteristics of an analyte hidden in a local evolutionary window where co-elution with other components is absent. Evolving factor analysis (EFA) was a considerable milestone in obtaining the elution window of each component and then LRI for chemometrics resolution [72]. It mined out the starting and end elution points of each peak though forward and backward derivation of eigenvalues with singular value decomposition (SVD) analysis of all gradually enlarged sub-matrices extracted from the whole data **X** [73]. Thereafter, several effective methods were developed on the basis of EFA to discover evolving elution information, but different strategies for full resolution, including fixed size moving window evolving factor analysis (FSMWEFA) [74], HELP [36, 37], window factor analysis (WFA) [75] and sub-window factor analysis (SFA) [76]. Among these methods, HELP, developed by Kvalheim and Liang et al., should have the widest application to resolve chromatographic data arising from studies of herbal medicines, biological samples, environmental pollutants, and others. The total literature citations to the two parts of this method number more than 600. It was recommended as having quite good performance [77,78]. Briefly, HELP can be decomposed to the following steps:

(1) finding of zero-component regions with noise only to define the detection limit and to simulate background data for subtraction;

(2) estimation of the number of components with eigenvalues from SVD operation, and further determination of selective regions via latent-projective graphs;

(3) local rank analysis to confirm the selectivity of regions found by visual detection; and,
(4) full rank resolution and component-stripping technique to determine **C** and **S** uniquely in Equation (2) by using the selective and zero-concentration regions. A conceptual interpretation of HELP introduced above is given in Fig. 4.

However, all the methods mentioned above were established on the evolutionary condition of components eluting out from the column, namely, "first-in-first-out". This means that the embedded overlap system exceeds the resolution abilities of these methods, where a minor peak is completely buried by a peak with large response along the whole elution window. Such a case challenges the conventional chemometrics methods, though some scientists attempted to solve this problem partially [79,80]. It may be resolved by the multi-run based methods introduced in sub-section 3.2.2.

It should be pointed out that full resolution of all overlapping peaks in GC×GC-MS data is difficult in most cases because of the modulation of $^{1}$D peaks to $^{2}$D fractions that still comprise many peaks and the complexity of real samples. Chemometrics techniques were introduced in our work for component correlation strategically to extract elution windows of analytes with MS channels from the whole GC×GC-TOF-MS dataset, which has a large size of more than 1.0 GB [3,42]. In addition, two chemometrics methods, that is, the similarity index (SI) and Fisher ratio (FR), were recently compared to locate peaks with significant concentration difference amongst samples [81]. An introduction to non-targeted cross-sample analysis has been reviewed elsewhere [82,83].

### *3.1.2. Rational resolution methods*

Most methods providing rational resolution results were built with iterative approaches, and generally include three steps:
(1) determination of initial set-point to start the iterative loop;
(2) proposal of a chemometrics method to generate new and qualified **C** or **S** on the basis of the characteristics of the chromatogram and/or spectral profile; and,
(3) termination of the iterative process with a predefined condition or attaining an acceptable precision.

Using iterative target transformation factor analysis (ITTFA) as an example [84,85], the abstract chromatogram obtained from principal component analysis (PCA) [86–88] can be utilized as the iterative initial point, and further calculation of new chromatographic profile after acquiring the rotation matrix. A changing threshold of **C** between the $n^{th}$ and $(n-1)^{th}$ iterative cycle is defined to determine continuance of the process or export of the final results. Other methods in this series include simple-to-use interactive self-modeling mixture analysis (SIMPLISMA) [89,90], orthogonal projection analysis (OPA) [91–93], and SIMPLEX [94].

Since these methods have ideal flexibility to cope with instrumental data with different structures, they were widely introduced to solve very complex chemical or biological problems. Especially, prior and external information can be applied to most of these methods for constrained searching and optimization, such as unimodality and non-negativity of chromatogram. This can speed up iterative convergence or improve the performance of the algorithm.

## *3.2. Multi-run-based data resolution*

The multi-run-based method means that more than one chromatographic run is involved to resolve a co-elution problem through additional information provided by other datasets. It can be the same sample analyzed with different experimental conditions, or

different, but related, samples with common chromatographic and/or spectral characteristics. For example, diseased biofluids and healthy controls containing both the same and different chemical components are instrumentally detected to find biomarkers for disease diagnosis in metabonomics, herbal medicines with different quality and therapy effects for disease treatment because of the different geographic origins of the herbs, prevailing climate/environmental effects, manufacture and/or storage conditions. For GC×GC-MS processing, primary peaks in $^{1}$D are modulated to fractions with spectral identification of each of the peaks or peak clusters in the fractions. Thus, it can mine out qualitative and quantitative information by using multi-run-based data resolution. Application examples are summarized below on the basis of trilinear decomposition (TLD) and non-trilinear methods [95,96]. It should be noted that the disadvantages and the limitations of single-run-based methods are largely overcome by this type of method (e.g., resolution of embedded overlapping multi-components).

Another example of multi-run-based data analysis is targeted comparison and analysis of components of particular interest in a sample. This can be achieved with chemometrics tools (not introduced in detail for brevity). For example, component correlation of complex mixtures combined with GC×GC-TOF-MS was investigated by Zeng et al. [3]. This allows the rapid discovery of the similarity of, or the difference between, samples with hundreds (if not thousands) of compounds. No full data resolution is required for such a study. The Dotmap was previously developed by Sinha et al. [97,98] to locate a specific compound in a predefined elution window of the whole GC×GC-TOF-MS data. These chemometrics tools provide effective choices for GC×GC data processing with MS detection. Non-targeted cross-sample analysis has been reviewed by Reichenbach et al. [82,83]. It includes the following five approaches to the study of sample classification and clustering, chemical fingerprinting, monitoring, and marker discovery:

(1) visual image comparisons;
(2) datapoint feature analysis;
(3) peak feature analysis;
(4) region feature analysis; and,
(5) peak-region feature analysis.

*3.2.1. Trilinear decomposition (TLD) methods*

Methods based on TLD techniques are natural generalizations of data resolution with bilinear characteristics, as shown in Equation (2). This is defined as the principle of linear additivity, suitable for each dimension of a cubic dataset including chromatographic and spectral detection, and dynamic change of samples (runs). The PARAFAC method is one of the most successful methods in this area with much reported work investigating GC×GC-TOF-MS data [66,99–103]. The detail and the decomposition flowchart of this algorithm is given in Fig. 5. For example, Snyder et al. [102] used this method for separation, identification and quantification of L-beta-methylamino-alanine, a neurotoxin possibly causing neurodegenerative disease, with GC×GC-TOF-MS analysis, and other biological studies with PARAFAC and GC×GC-TOF×MS analysis, including yeast metabolites exhibiting oscillatory behavior [103], and metabolite-profiling analysis of *Methylobacterium extorquens* $AM_1$ [104]. A chemical-weapons precursor as possible forensic signature and automated analysis of target and non-target analytes were reported with very similar steps by Hoggard et al. [38,39,105], and other applications [106]. The limitation of PARAFAC for resolution of isomers has been investigated by Sinha et al. [107,108] initiated by TLD.

Other important multi-way methods include GRAM [109–111] and direct TLD (DTD) developed by Kowalski et al. [112]. As introduced above, Fraga et al. [69] studied

GC×GC data deconvolution by using GRAM. Of course, it can be extended to identification and quantification of analytes with GC×GC-MS. The purpose of TLD methods is to find matrices **C** and **S** of pure components as introduced in Equation (2), and quantitative change of relative concentrations about composition in different runs (A). However, the trilinear model assumes that the chromatographic profile of a particular analyte has no shift of $t_R$ and variation of peak shape among runs. This is not easy to attain in real experiments, and may require data preprocessing to correct peak shifting and the shape of the chromatographic profile. Obviously, this is time consuming with manual intervention, and may lead to errors. But the methods introduced below are independent of the tri-linear model, and can be conveniently applied for GC×GC-MS data analysis.

*3.2.2. Non-trilinear decomposition (non-TLD) methods*

Like the single-run-based method for resolution with iterative searching, non-TLD methods provide a rational solution to acquire **C**, **S** and **A** introduced above. Such methods include MCR-ALS [40,41], PARAFAC2 [113,114] and Tucker3-ALS [115,116] with flexible features for resolution, without serious restriction to the constraint, as in the TLD methods.

Among these methods, MCR-ALS has been most successfully applied in the past two decades [96,117–120]; it is an iterative technique with identity as the algorithm introduced in Section 2. The two key features of this method can be summarized as follows:

(1) augment datasets collected from different runs to obtain **C** and **S** simultaneously in all mixtures; and,
(2) alternative optimization of **C** and **S** by using ALS.

After determination of the initial values of **C** and **S** for iteration operation (e.g., with the help of EFA methods), the constraints of chromatographic and spectral profiles can be applied to reduce rotational ambiguity for resolution. It has been effectively extended to GC×GC-MS data processing. The working procedure is illustrated in Fig. 6. For example, Parastar and Tauler et al. [100] utilized it for resolution and quantification of complex mixtures of polycyclic aromatic hydrocarbons (PAHs) in heavy fuel-oil samples combined with GC×GC-TOF-MS analysis. The chromatographic and spectral profiles of pure compositions of samples and standard mixtures were obtained with the help of the MCR-ALS method because of the convenience of this method for simultaneous analysis of several chromatographic runs, and effectiveness for third-order and even fourth-order data resolution, such as GC×GC-TOF-MS data [121]. The results were compared to the PARAFAC method with relatively better performance. However, there are some drawbacks of MCR-ALS method, as follows:

(1) iterative operation with potential failure to correct local minimum, unlike GRAM and DTD methods with unique outcomes;
(2) initial estimates needed before iterative analysis; and,
(3) advanced intervention required and unsuitability as a “black box” method.

PARAFAC2 is another method in this area proposed by Bro et al., which is a modified version of PARAFAC and is not limited by shifting $t_R$s and change of peak shape. It has been successfully applied [33]. The three reasons for $t_R$ shifts in GC×GC-TOF-MS analysis were first summarized by Skov et al. [122], and PARAFAC2 was then used to study severely shifted peak profiles and compared to PARAFAC. However, PARAFAC2 is a method with high computational cost and complexity. Further, the constraint conditions of chromatographic and spectral profiles, such as unimodality and non-negativity, can be used to optimize the results, as well as selective constraints to specific compounds.

The benefit of these methods is that they consider the resolution of more than one chromatographic run. The data dimensions (total running time) of chromatograms among the runs can be different, but the spectral range should be identical for all runs. For GC×GC-MS data analysis, such methods have evident advantages, in contrast to algorithms based on TLD, because of the frequent shift of $t_R$s in the $^2$D separation, and the change of peak shape. Experimentally, they are ideally the same for fractions of primary peaks and runs as expected.

AMWFA is an effective tool developed by Zeng and Liang et al. [42,43] to compare the similarity and the difference of different but related chromatograms by using evolutionary window information. It is another kind of multi-run-based method for data resolution, which can be utilized to resolve "contaminated" peak clusters through use of information buried in two matrices, say **X** and **Y**. The basic principle of AMWFA is illustrated in Fig. 7. The core idea is to transform the resolution problem of pure common spectra into a solution of the eigen equation of the data matrix. It is then readily possible to extract spectra for component identification, and simultaneous discovery of the number of common compounds. It has been a powerful algorithm to fill in the gap between 2D and 3D resolution methods. Two datasets are involved in this method, but no tri-linear model is required for **X** and **Y**, and unique results especially can be obtained with no iterative operation. Though AMWFA has not yet been applied to solve problems combined with GC×GC-MS analysis, it has potential advantages in terms of the theoretical basis and the wide applications of this method to process coupled datasets.

## 4. New insights from chemometrics for GC×GC structure interpretation with two-dimensional characteristics

### *4.1. Data preprocessing*

Data preprocessing is important to improve data quality and then to enhance the performance of chemometrics methods to obtain final results [123]. Here, two aspects are introduced, including peak detection and correction of time shift among runs. Further studies for signal processing of GC×GC are not included in this review [124,125]. The typical illustration to summarize GC×GC data processing is given in Fig. 1.

#### *4.1.1. Two-dimensional peak detection*

Peak detection in SDC and coupled column 2D datasets has been widely studied [126–129]. For example, Vivo-Truyols et al. [128] developed high-order derivatives and smoothing techniques for chromatograms to automate estimation of peak parameters, and the corresponding elution ranges. All conventional algorithms for SDC and coupled 2D studies can be employed to process GC×GC and GC×GC-MS directly, if the raw chromatograms are used without data conversion and dimensional transformation. However, peak detection in GC×GC can be quite different because of the 2D characteristics of separations in both $^1$D and $^2$D after data re-arrangement according to the principles for modulation, as mentioned in the Introduction (Section 1 above). In general, $t_R$s and intensity of the detected peaks are used to determine data points of analytes in both SDC and GC×GC analysis. The two main methods for GC×GC peak detection, including a two-step algorithm and the watershed algorithm, have been compared in previous studies [130]. The result shows that the two-step algorithm has higher accuracy than when no correction of $t_R$ shift was applied to the watershed algorithm. Latha et al.

[131] further investigated the two methods in 2011 after correcting shifts of $t_R$s with the two methods. It showed the watershed algorithm has better accuracy to detect resolved peaks.

In 2007, Peters et al. [132] proposed a method for detection of $^1$D peaks, and then determine the attribution and merge these peaks to the same component in $^1$D separation by using a decision-tree technique. The features of $t_R$ and separation regions of peaks in $^2$D are compared, and unimodality in $^1$D, difference of $t_R$s and regions of common peak were applied as criteria. It was demonstrated for a commercial air-freshener sample.

Peak detection is one of the important steps for data preprocessing in both SDC and GC×GC studies. If this can be attained correctly, it may help other aspects for information extraction (e.g., deconvolution) and then qualitative and quantitative analysis of analytes.

The limits of peak detection and quantification in comprehensive two-dimensional separation (C2DS) were theoretically investigated by de la Mata and Harynuk [133]. Because of the model consistency of $^2$D modulated fractions with the $^1$D chromatographic profile, a conventional Gaussian filter method was applied by Vivó-Truyols for peak detection.

Next, a new method was proposed to merge the corresponding sub-peaks of each primary peak effectively by using Bayesian inference [134].

In addition, Allen et al. and Filgueira et al. [135,136], respectively, proposed a singular value decomposition-based method to reduce the influence of background in LC×LC separation and to improve further the quality for peak detection and quantification. This should be potentially extendable to GC×GC peak detection.

#### *4.1.2. Correction of retention-time shift*

Alignment of shift of $t_R$s among runs is important for GC×GC (e.g., for peak comparison, chemometrics clustering, discriminant analysis, and other data processing steps) [137,138]. For high-throughput data analysis in metabolomics, automated alignment is required for rapid handling of hundreds or thousands of chromatograms. This challenges conventional methods, especially GC×GC-MS data with their spectral attributes.

Correlation optimized warping (COW) and dynamic time warping (DTW) are two of the most important methods to align shifts of SDC through correction of samples towards a reference chromatogram [139,140]. The detailed principles were introduced in the reported work. Through interpolation, the COW algorithm warps local regions to maximize the correlation between warped and reference chromatographic profiles. DTW was originally proposed for speech recognition, and is employed to align chromatograms through non-linearly warping two trajectories with an objective function of minimum distance between them. They have been extended for correction of $t_R$ shifts of GC×GC-TOF-MS and GC×GC by Vial et al [141]. and Zhang et al. [142], respectively.

On the basis of the 3D data structure of GC×GC-MS with chromatographic retention and MS information, an algorithm with simultaneous correlation optimization of distance and spectra (DISCO) was developed by Wang et al. [143]. After z-score transformation of raw retention data, distance minimization of 2D $t_R$s and maximization of Pearson's correlation coefficient of MS were applied as indices to determine alignment of the peaks. A progressive retention-time map-search method was utilized to correct the 2D shifts by using a local linear fitting technique. An improved version of this algorithm (DISCO2) has been reported to overcome the drawbacks of the previous method [144].

Recently, Jeong et al. [145] investigated alignment of GC×GC-MS data for the analysis of metabolomic profiling. With the help of an empirical Bayes model, matching confidence of peaks was calculated via posterior probability, and then the metabolite pairs

with high confidence were selected on the basis of the results. The representative landmark peaks were further generated for adjustment of $t_R$s for all the runs. Weusten et al. [26] first transformed GC×GC chromatograms with wrap-around separation in the 2D display as a surface of a 3D cylinder. Then, the cylindrical distance and mass spectral correlation were applied to define a combined similarity index and further clustering analysis of GC×GC-MS data. The former reflects similarity of chromatographic behavior, and the latter for chemical structure, respectively. An evident advantage of this method is the suitability for wrap-around treatment, which widely exists in GC×GC separation, arising when $^2t_R$ exceeds the $P_M$ setting. The study of GC×GC with high-resolution mass spectrometry (HRMS) was reported by Reichenbach et al. [146] on the basis of an informatics approach. The biggest challenge for alignment of this data type is the large size of the raw datasets. No comprehensive peak matching is required for this method by using a few reliable peaks and peak-based retention-plane windows.

Other techniques for GC×GC alignment include piecewise alignment [27], rank alignment [69], correlation optimized shifting [147], a bilinear peak alignment (BPA) a combined MCR technique [148], and GC×GC-MS alignment (e.g., Smith-Waterman local alignment-based algorithm) [149], novel mixture similarity algorithm with simultaneous peak distance and spectral similarity measures [150], and integrative normalization and comparative analysis [137].

In conclusion, shift correction is still a "hot", but difficult, research area in GC×GC separation, especially for GC×GC-MS data processing with high-throughput and high complexity. Very recently, an automated data-analysis strategy was introduced by Shellie and Harvey for GC×GC data reduction including processing of $t_R$ alignment. No user intervention was required from input of raw chromatogram, data transformation and preprocessing to output of results. Thus, it allows rapid analysis of batch samples [123].

### *4.2. Orthogonality study*

GC×GC distributes each peak in the 2D plane with $^1t_R$ and $^2t_R$ corresponding to the separation by two columns. Different experimental conditions will change the peak capacity (see Introduction, Section 1 above) and distribution of peaks in the 2D space [151]. Calculation of 2D peak capacity was re-addressed in recently-reported work [152,153]. Orthogonality is the quantitative evaluation of the separation performance (resolving power) estimated by defining an efficiency index according to a measure of effective use of ideal peak capacity ($n_{max}$). It is generally defined from 0 to 1, respectively corresponding to perfectly correlated separation (abbreviated as PCS) and orthogonal separation (abbreviated as OS). Maximizing orthogonality should correspond to best use of separation space and highest overall resolution of sample components, This helps to select the optimal column set and other conditions to separate complicated mixtures.

Most reported works were established based on assessment of rectangular bins, as introduced in Fig. 8. First, $^1t_R$ and $^2t_R$ are normalized in order to make the magnitudes of the $t_R$s comparable. The GC×GC 2D space is then divided into equal intervals along the two separation dimensions after determination of peak capacities in $^1$D and $^2$D, namely, $^1n_c$ and $^2n_c$. Next, occupation and distribution of peaks in the 2D virtual grid space (rectangular bins) is evaluated according to the orthogonality definition. For example, Gilar et al. [154] proposed an equation, as given in Equation (3) to calculate orthogonality with consideration of occupation percent of bins. The numerator and the denominator correspond to the practical and theoretical peak capacity of a GC×GC separation. Two years later, Watson et al. [155] modified the definition to Equation (4), since orthogonality is a function of $n_c$ and Equation (3) is only a special case with the $n_c$ limit of positive

infinity. However, the distribution correlation of peaks in the 2D plane is not taken into account in both these methods.

$$\sigma = \left(\sum \text{bins} - \sqrt{n_{s,\max}}\right) \Big/ \left(0.63 \times n_{s,\max}\right) \tag{3}$$

$$\sigma = \left(\sum \text{bins} - n_s\right) \Big/ \left(0.63 \times n_s^2 - n_c\right) \tag{4}$$

In 2011, an information-theory method was developed for orthogonality evaluation, still based on occupation of rectangular bins [156]. It is defined by using conditional entropy of 2D peak distribution.

Prior to this, another method on the basis of the principle of information theory was proposed by Slonecker et al. [157]. This is based on the mathematical consistency between chromatographic separation and the fundamental concept of entropy. In addition, geometric approaches were developed to study GC×GC separation power with factor analysis [158,159] and peak distribution in parallelograms or other patterns [152]. But the target analytes not diagonally distributed in the C2DS space cannot be ideally calculated.

However, there are some serious limitations with these methods as follows:

(1) the value of each orthogonality metric does not exclusively correspond to a single case of C2DS (i.e. the index is not sensitive enough to determine subtle differences in GC×GC separation);
(2) only derived data are used to define orthogonality, not the original data including retention information;
(3) only occupation of bins is considered, not how the bins are correlated in the 2D plane; and,
(4) orthogonality is restricted to 0–1 artificially. It is not achievable for real separations and may be unsuitable to some separation cases.

One or more shortcomings obtain for all these methods introduced above.

A comprehensive comparison has been provided by Gilar et al. [160], introducing the suitability and the pitfalls of these methods. Thus, new methods for orthogonality evaluation still need to be developed in the future [161–163]. The most recent contribution was made by ourselves [164], and divides the orthogonality metric into two parts (i.e. $C_{pert}$ and $C_{peaks}$) to introduce naturally` peak coverage percent, and distribution correlation of compounds spanning the 2D separation panel, respectively. They were further quantitatively estimated by “bin coverage” and a simple-linear regression model. This ideally overcomes the shortcomings of previous methods mentioned above.

### *4.3. Image processing*

As mentioned in the Introduction (Section 1, above), the high-dimensional characteristics of GC×GC data make it appropriate for image-structure interpretation and to be processed with image techniques. Most of the previous GC×GC studies in this area can be found in reported work [28,29,165]. Software for GC×GC data analysis and image processing include ChromaToF (Leco, USA), ImageJ 1.37v (Wayne Rasband, NIH, USA), GC Image developed by Reichenbach et al. [169], Image software (GC Image, Lincoln, NE), and Statistica (StatSoft, Tulsa, OK) [137,166–168]. In the literature, image-analysis techniques have been employed for comparative visualization, peak matching,

background removal, quantification and pattern recognition in GC×GC analysis. For example, ImageJ software was used for fingerprint recognition and comparison with PCA after obtaining areas of the chromatographic spots in contour plots [168,169].

On the basis of 2D gel-based image analysis in proteomics, Schmarr and Bernhardt developed a method for unbiased pattern comparison of GC×GC for profiling analysis of volatile compounds obtained from fruits [170,171]. It was summarized as follows:

(1) raw GC×GC chromatograms were converted into contour plots and further gray-scale images;
(2) variation among runs was compensated by image warping and merged to fusion image and spot patterns;
(3) quantification of analytes was achieved after determination of spot boundaries; and,
(4) it can be used for multivariate statistical analysis and pattern recognition.

In order to find the chemical difference from GC×GC chromatograms, visual comparison was attained after data preprocessing, such as registration and scaling to remove variations in $t_R$s and sample amounts [172]. Three methods for image comparison were introduced in detail, including the grayscale-difference method to obtain individual pixel differences between images, the colorized difference method to show pixel differences and values simultaneously, and the fuzzy difference method to remove possible variations in peak shape and alignment.

Another study of image processing in GC×GC was employed for discrimination analysis by Vial et al. [173]. The discriminant pixel approach (DPA) was introduced to find the most discriminating pixels with linkage to chromatographic peaks (i.e. the discriminatory power to each class is defined according to chromatogram pixels after data pre-processing and time alignment).

In addition, a method to remove GC×GC image background was proposed by Reichenbach et al. [174]. Structural and statistical properties of the data were used to estimate image-background levels, and the image background was further removed from the raw image to generate "pure" chromatogram data. The quantitative relationship between the peaks and chemicals was then obtained for rapid, accurate detection of GC×GC peaks.

## 5. Summary of other aspects of GC×GC studies with chemometrics

### *5.1. Experimental optimization*

Experimental optimization is important to analyze complicated mixtures and to maximize instrumental capability [175–178]. The complicated network and interplay of parameters for GC×GC analysis was reviewed by Harynuk and Gorecki [179]. The whole picture of experimental variables clearly shows the difficulty in deducing optimal conditions. In practice, design of experiment (DoE) [180–184], multi-objective and variable optimization [185–187] and employment of separation peak capacity and orthogonality [155,162,163,188–190] should be three pillars to achieve this purpose. For example, Omar et al. [181] optimized the conditions for GC×GC-MS analysis of essential oils of plants by using the DoE method and Multisimplex. Amongst the optimized parameters were $P_M$, discharge-time and first and second column flows.

In 2005, O'Hagan et al. [186] developed a closed-loop, multi-objective approach to optimize parameters for GC-TOF-MS analysis automatically, and further delivered the same strategy to GC×GC-TOF-MS in 2007 [187]. For serum separation in a metabolomics study, 18 experimental variables were optimized. Dorman et al. [185] fully

optimized and predicted separation in GC×GC by using computer simulation and modeling. Simultaneous optimization of all columns and run-time variables were studied on the basis of enthalpy and entropy.

Model approaches for prediction of two-dimensional $t_R$s and/or peak widths ($w_b$) were developed in recent studies [191,192]. This can help to simulate and to optimize further analytical conditions. For example, experimental data under isothermal separation were extended to obtain $t_R$ and $w_b$ of temperature programmed GC×GC analysis on the basis of chromatographic theory [193]. With the help of the solvation parameter model, Seeley successfully generated retention diagrams of $^1t_R$ and $^2t_R$ for GC×GC [191, 194]. This is a crucial step for the prediction of 2D separation before experimentation and further confirmation of retention properties. very recently, we developed a new method for determination of $t_R$s [50]. The concept of center of gravity (COG) was applied to estimate the $t_R$s of primary peaks ($^1t_R$) by using the peak area and $^2t_R$s of modulated fractions in the comprehensive separation. A moving window search strategy was further used to derive the $t_R$s of overlapping peaks simultaneously.

Separation quality of comprehensive two-dimensional LC (LC×LC) was estimated by defining a new chromatographic response function [151]. It was attained through extension of peak purity of SDC to LC×LC separation. This may be extended to GC×GC analysis for parameter optimization. It was reviewed recently by Bedani et al. [195].

*5.2. Pattern recognition*

Pattern-recognition techniques for both supervised and unsupervised analysis have been applied for GC×GC data treatment, as illustrated in Fig. 9. PCA and partial least squares discriminant analysis (PLS-DA) [196–199] are the two representative methods with most successful applications.

The principle of PCA has been widely introduced in the literature [86]. For processing GC×GC data, PCA can cluster samples with correlations. But PCA was mostly applied to analyze peak tables obtained from GC×GC or GC×GC-MS analysis. It is unable to process raw GC×GC related data with 2D or higher dimension data. For example, Vestner et al. [200] recently differentiated control wines and those fermented with different starter cultures of malolactic fermentation (MLF), and, using GC×GC-TOF-MS analysis, McGregor et al. [201] separated 12 dense non-aqueous phase liquids (DNAPLs)from former manufactured gas plants. Further applications include:

(1) discrimination of five different animal-derived fatty acids, including lard, chicken fat, beef tallow, mutton tallow and cod-liver oil [202];

(2) classification of radix ginseng with different ages [203] and *Notopterygium incisum* Ting ex H.T chang collected from different regions [204]; and,

(3) difference discovery of 54 chromatograms from three different species, namely, basil, peppermint, and sweet herb stevia [205], and other examples by the PCA method [27,206–209].

The techniques of GC×GC or GC×GC-MS were utilized as tools for sample analysis and component identification for all these studies. After alignment of GC×GC chromatograms with the 2D DTW method, Vial et al. [141] compared and classified three types of tobaccos by using independent component analysis (ICA) with comparison of PCA, and ICA extracts the original signals with a hypothesis of independence among the signals, but PCA found a sequence of uncorrelated principal components (PCs) including the variance. In addition, hierarchical clustering analysis (HCA) has some applications for pattern recognition of GC×GC data [170,210].

The PLS-DA method has been valuable for GC×GC data analysis with a supervised

strategy [211,212]. Prediction of total exposure time of petroleum mixtures to weathering effects allowed an environmental investigation to determine the reason for a fire. Zorzetti and Harynuk applied the PLS-DA method to predict the weathering levels (relatively fresh or highly weathered) in combination with GC×GC separation [213]. It was also utilized to detect gasoline samples with varying octane ratings and from several vendors.

For studies in metabolomics, PLS-DA with GC×GC-TOF-MS data successfully implemented:

(1) the discrimination analysis of HUC-1 (non-tumorigenic) and HUC T-2 (tumorigenic) cells with metabolic footprinting [196];
(2) two transgenic lines and the control line with terpenoid metabolic profiling analysis [197];
(3) chromatograms of control and exposed rats with aged and diluted side-stream cigarette smoke [198]; and,
(4) metabolite peak tables of diabetic patients and healthy controls [214].

Other techniques, including DPA (as introduced above), principal-component discriminant analysis, projection pursuit, back propagation-artificial neural networks (BP-ANN) and least squares-support vector machine (LS-SVM) were proposed to find differences of chemical profiles with GC×GC or GC×GC-MS data, and then classify samples [173,215,216].

### *5.3. Model calibration*

The principles and the methods for quantitative modeling have been comprehensively introduced elsewhere [217–219]. For the study of GC×GC analysis, PLS, including conventional PLS, trilinear PLS and N-way PLS, are the main chemometrics tools to predict chemical properties of components, such as retention time [117,220–224]. Compounds of environmental interest, such as polychlorinated biphenyls (PCBs), PAHs and gasoline were the main research objectives in past studies [66,225–228].

Interval multi-way PLS (iNPLS) was newly developed for target quantitative analysis of GC×GC [229]. The new idea is to split the 2D chromatogram into small sections as independent new (partial) chromatograms. The conventional NPLS method was used to establish models for each segment of the whole chromatogram with concentration of the target analyte. It was found effective for quantification of some allergens in perfume samples, even for poorly resolved peaks. Noorizadeh and Noorizadeh investigated 69 drugs with comparison of two linear methods and non-linear methods, including multiple linear regression (MLR) and PLS, and kernel PLS (KPLS) and Levenberg-Marquardt artificial neural network (LM-ANN) [221]. Genetic algorithm (GA) was used for variable selection before regression analysis. In the above-mentioned study by Zorzetti and Harynuk [213], several calibration methods were compared to predict the amount of exposure time of a petroleum mixture to weathering effects, including PLS, non-linear PLS (PolyPLS) and locally-weighted regression (LWR) [213]. The best multi-linear regression (BMLR) method was also used by Ren et al. [230] to predict $t_R$s of PCBs congeners for GC×GC-TOF-MS analysis.

The area of model calibration in GC×GC is essentially no different from conventional QSPR or QSAR studies. Prior chemometrics tools and research perspectives can be conveniently extended and combined with GC×GC separation.

## 6. Concluding remarks

Chemometrics comprises many methods to mine out the rich information existing in instrumental data, including GC×GC and GC×GC-MS. Cross-discipline support, such as mathematics, statistics and computer science, offer effective developments to process data with the help of updated tools. Though chemometrics has been successfully used in many research areas of GC×GC-related data studies, new developments are still ongoing. An urgent challenge is the automated, high-throughput data analysis of samples with hundreds or more chemical components, such as the investigation of metabolite profiling or fingerprinting of metabolomics. This is a continuing issue with chemometrics.

However, the special data structures of GC×GC and GC×GC-MS with 2D retention and spectral detection complicate data processing for chemometrics. Thus, future studies will probably exploit many research topics of chemometrics, such as data pretreatment to improve data quality, deconvolution methods to extend experiments for qualitative and quantitative analysis of analytes (with $^1t_R$, $^2t_R$ and MS data), pattern recognition to classify samples with certain similarities and differences, model calibration to establish quantitative relationships of explanatory variables and response variables, and special aspects of GC×GC and GC×GC-MS data processing, including orthogonality estimation and image analysis. To readers with an interest to enter this area, this should be a good initial and enlightening start, with introduction of fundamental principles and theories. It also provides an overview understanding to researchers with some experience in this field.

**Acknowledgments**

This work is financially supported by a Australian Research Council Discovery Grant (No. DP0988656), the National Basic Research Program of China (No. 2012CB518303) from the State Ministry of Science & Technology of China, the Creative Research Group Project (No. 21021004) from the National Natural Science Foundation of China, and the One Hundred Talent Program of the Dalian Institute of Chemical Physics, Chinese Academy of Sciences.

**References**

[1] P.J. Marriott, Gas chromatography multidimensional techniques. in: P. Worsfold, A. Townshend and C. Poole (Eds.), Encyclopedia of Analytical Science. Elsevier, Oxford, UK. 2005, p. 39.

[2] R.G. Brereton, Applied chemometrics for scientists, John Wiley & Sons, Chichester, West Sussex, UK, 2007.

[3] Z.D. Zeng, H.M. Hugel, P.J. Marriott, Component correlation between related samples by using comprehensive two-dimensional gas chromatography - time-of-flight mass spectrometry with chemometric tools, J. Chromatogr. A 1254 (2012) 98-106.

[4] P.Q. Tranchida, D. Sciarrone, P. Dugo, L. Mondello, Heart-cutting multidimensional gas chromatography: A review of recent evolution, applications, and future prospects, Anal. Chim. Acta 716 (2012) 66-75.

[5] P.J. Marriott, S.T. Chin, B. Maikhunthod, H.G. Schmarr, S. Bieri, Multidimensional gas chromatography, TrAC-Trend Anal. Chem. 34 (2012) 1-21.

[6] M.F. Almstetter, P.J. Oefner, K. Dettmer, Comprehensive two-dimensional gas chromatography in metabolomics, Anal. Bioanal. Chem. 402 (2012) 1993-2013.

[7] G.S. Frysinger, R.B. Gaines, Separation and identification of petroleum biomarkers by comprehensive two-dimensional gas chromatography, J. Sep. Sci. 24 (2001) 87-96.

[8] R.A. Shellie, W. Welthagen, J. Zrostlikova, J. Spranger, M. Ristow, O. Fiehn, R. Zimmermann, Statistical methods for comparing comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry results: Metabolomic analysis of mouse tissue extracts, J. Chromatogr. A 1086 (2005) 83-90.

[9] S.M. Song, P. Marriott, A. Kotsos, O.H. Drummer, P. Wynne, Comprehensive two-dimensional

gas chromatography with time-of-flight mass spectrometry (GC x GC-TOF-MS) for drug screening and confirmation, Forensic Sci. Int. 143 (2004) 87-101.

[10] J.E. Welke, C.A. Zini, Comprehensive Two-Dimensional Gas Chromatography for Analysis of Volatile Compounds in Foods and Beverages, J. Brazil. Chem. Soc. 22 (2011) 609-622.

[11] S. Zhu, X. Lu, K. Ji, K. Guo, Y. Li, C. Wu, G. Xu, Characterization of flavor compounds in Chinese liquor Moutai by comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry, Anal. Chim. Acta 597 (2007) 340-348.

[12] R. Shellie, P. Marriott, Opportunities for ultra-high resolution analysis of essential oils using comprehensive two-dimensional gas chromatography: a review, Flavour Frag. J. 18 (2003) 179-191.

[13] R.C.Y. Ong, P.J. Marriott, A review of basic concepts in comprehensive two-dimensional gas chromatography, J. Chromatogr. Sci. 40 (2002) 276-291.

[14] L. Mondello, P.Q. Tranchida, P. Dugo, G. Dugo, Comprehensive two-dimensional gas chromatography-mass spectrometry: A review, Mass Spectrom. Rev. 27 (2008) 101-124.

[15] H.J. Cortes, B. Winniford, J. Luong, M. Pursch, Comprehensive two dimensional gas chromatography review, J. Sep. Sci. 32 (2009) 883-904.

[16] M. Adahchour, J. Beens, U.A.T. Brinkman, Recent developments in the application of comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1186 (2008) 67-108.

[17] M. Adahchour, J. Beens, R.J.J. Vreuls, U.A.T. Brinkman, Recent developments in comprehensive two-dimensional gas chromatography (GC X GC) I. Introduction and instrumental set-up, TrAC-Trend Anal. Chem. 25 (2006) 438-454.

[18] M. Adahchour, J. Beens, R.J.J. Vreuls, U.A.T. Brinkman, Recent developments in comprehensive two-dimensional gas chromatography (GC x GC) II. Modulation and detection, TrAC-Trend Anal. Chem. 25 (2006) 540-553.

[19] M. Adahchour, J. Beens, R.J.J. Vreuls, U.A.T. Brinkman, Recent developments in comprehensive two-dimensional gas chromatography (GC x GC) III. Applications for petrochemicals and organohalogens, TrAC-Trend Anal. Chem. 25 (2006) 726-741.

[20] M. Adahchour, J. Beens, R.J.J. Vreuls, U.A.T. Brinkman, Recent developments in comprehensive two-dimensional gas chromatography (GC X GC) - IV. Further applications, conclusions and perspectives, TrAC-Trend Anal. Chem. 25 (2006) 821-840.

[21] P.Q. Tranchida, G. Purcaro, P. Dugo, L. Mondello, Modulators for comprehensive two-dimensional gas chromatography, TrAC-Trend Anal. Chem. 30 (2011) 1437-1461.

[22] M. Edwards, A. Mostafa, T. Gorecki, Modulation in comprehensive two-dimensional gas chromatography: 20 years of innovation, Anal. Bioanal. Chem. 401 (2011) 2335-2349.

[23] P.J. Marriott, P. Schoenmakers, Z.Y. Wu, Nomenclature and conventions in comprehensive multidimensional chromatography - an update, LC GC Europe 25 (2012) 266-+.

[24] P. Schoenmakers, P. Marriott, J. Beens, Nomenclature and conventions in comprehensive multidimensional chromatography, LC GC Europe 16 (2003) 335-339.

[25] R.B. Wilson, W.C. Siegler, J.C. Hoggard, B.D. Fitz, J.S. Nadeau, R.E. Synovec, Achieving high peak capacity production for gas chromatography and comprehensive two-dimensional gas chromatography by minimizing off-column peak broadening, J. Chromatogr. A 1218 (2011) 3130-3139.

[26] J.J.A.M. Weusten, E.P.P.A. Derks, J.H.M. Momrners, S. van der Wal, Alignment and clustering strategies for GC x GC-MS features using a cylindrical mapping, Anal. Chim. Acta 726 (2012) 9-21.

[27] K.M. Pierce, L.F. Wood, B.W. Wright, R.E. Synovec, A comprehensive two-dimensional retention time alignment algorithm to enhance chemometric analysis of comprehensive two-dimensional separation data, Anal. Chem. 77 (2005) 7735-7743.

[28] K.M. Pierce, R.E. Mohler, A Review of chemometrics applied to comprehensive two-dimensional separations from 2008-2010, Sep. Purif. Rev. 41 (2012) 143-168.

[29] Z.D. Zeng, H.M. Hugel, P.J. Marriott, Chemometrics in comprehensive multidimensional separations, Anal. Bioanal. Chem. 401 (2011) 2373-2386.

[30] K.M. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, Review of chemometric analysis techniques for comprehensive two dimensional separations data, J. Chromatogr. A 1255 (2012) 3-11.

[31] J.T.V. Matos, R. Duarte, A.C. Duarte, Trends in data processing of comprehensive two-dimensional chromatography: State of the art, J. Chromatogr. B 910 (2012) 31-45.

[32] Y.Z. Liang, O.M. Kvalheim, R. Manne, White, gray and black multicomponent systems - a classification of mixture problems and methods for their quantitative-analysis, Chemometr. Intell. Lab. 18 (1993) 235-250.

[33] K.M. Pierce, J.C. Hoggard, R.E. Mohler, R.E. Synovec, Recent advancements in comprehensive two-dimensional separations with chemometrics, J. Chromatogr. A 1184 (2008) 341-352.

[34] H.G.J. Mol, H.J. van der Kamp, G. van der Weg, M. van der Lee, A. Punt, T.C.d. Rijk, Validation of automated library-based qualitative screening of pesticides by comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry, J. AOAC Int. 94 (2011) 1722-1740.

[35] J.A. Murray, Qualitative and quantitative approaches in comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1251 (2012) 1-9.

[36] Y.Z. Liang, O.M. Kvalheim, H.R. Keller, D.L. Massart, P. Kiechle, F. Erni, Heuristic evolving latent projections - resolving 2-way multicomponent data .2. detection and resolution of minor constituents, Anal. Chem. 64 (1992) 946-953.

[37] O.M. Kvalheim, Y.Z. Liang, Heuristic evolving latent projections - resolving 2-way multicomponent data .1. selectivity, latent-projective graph, datascope, local rank, and unique resolution, Anal. Chem. 64 (1992) 936-946.

[38] J.C. Hoggard, R.E. Synovec, Parallel factor analysis (PARAFAC) of target analytes in GC x GC-TOF-MS data: Automated selection of a model with an appropriate number of factors, Anal. Chem. 79 (2007) 1611-1619.

[39] J.C. Hoggard, R.E. Synovec, Automated resolution of nontarget analyte signals in GC x GC-TOF-MS data using parallel factor analysis, Anal. Chem. 80 (2008) 6677-6688.

[40] R. Tauler, B. Kowalski, S. Fleming, Multivariate curve resolution applied to spectral data from multiple runs of an industrial-process, Anal. Chem. 65 (1993) 2040-2047.

[41] J. Jaumot, R. Gargallo, A. de Juan, R. Tauler, A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB, Chemometr. Intell. Lab. 76 (2005) 101-110.

[42] Z.D. Zeng, Y.Z. Liang, Y.L. Wang, X.R. Li, L.M. Liang, Q.S. Xu, C.X. Zhao, B.Y. Li, F.T. Chau, Alternative moving window factor analysis for comparison analysis between complex chromatographic data, J. Chromatogr. A 1107 (2006) 273-285.

[43] Z.D. Zeng, Y.Z. Liang, Z.H. Jiang, F.T. Chau, J.R. Wang, Quantification of target components in complex mixtures using alternative moving window factor analysis and two-step iterative constraint method, Talanta 74 (2008) 1568-1578.

[44] S. Castillo, I. Mattila, J. Miettinen, M. Oresic, T. Hyotylainen, Data analysis tool for comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry, Anal. Chem. 83 (2011) 3058-3067.

[45] M. Kallio, M. Kivilompolo, S. Varjo, M. Jussila, T. Hyotylainen, Data analysis programs for comprehensive two-dimensional chromatography, J. Chromatogr. A 1216 (2009) 2923-2927.

[46] A. Lommen, H.J. van der Kamp, H.J. Kools, M.K. van der Lee, G. van der Weg, H.G.J. Mol, metAlignID: A high-throughput software tool set for automated detection of trace level contaminants in comprehensive LECO two-dimensional gas chromatography time-of-flight mass spectrometry data, J. Chromatogr. A 1263 (2012) 169-178.

[47] P. Wojtowicz, J. Zrostlikova, T. Kovalczuk, J. Schurek, T. Adam, Evaluation of comprehensive

two-dimensional gas chromatography coupled to time-of-flight mass spectrometry for the diagnosis of inherited metabolic disorders using an automated data processing strategy, J. Chromatogr. A 1217 (2010) 8054-8061.

[48] S. Dasgupta, K. Banerjee, S.H. Patil, M. Ghaste, K.N. Dhumal, P.G. Adsule, Optimization of two-dimensional gas chromatography time-of-flight mass spectrometry for separation and estimation of the residues of 160 pesticides and 25 persistent organic pollutants in grape and wine, J. Chromatogr. A 1217 (2010) 3881-3889.

[49] C. von Muehlen, P.J. Marriott, Retention indices in comprehensive two-dimensional gas chromatography, Anal. Bioanal. Chem. 401 (2011) 2351-2360.

[50] Z.D. Zeng, H.M. Hugel, P.J. Marriott, Simultaneous estimation of retention times of overlapping primary peaks in comprehensive two-dimensional GC, J. Sep. Sci. (2013) in press. DOI: 10.1002/jssc.201300267.

[51] P. Nikitas, A. Pappa-Louisi, A. Papageorgiou, On the equations describing chromatographic peaks and the problem of the deconvolution of overlapped peaks, J. Chromatogr. A 912 (2001) 13-29.

[52] R.A. Vaidya, R.D. Hester, Deconvolution of overlapping chromatographic peaks using constrained non-linear optimization, J. Chromatogr. 287 (1984) 231-244.

[53] A. Economou, P.R. Fielden, A.J. Packham, Deconvolution of analytical peaks by means of the fast hartley transform, Analyst 121 (1996) 1015-1018.

[54] A. Economou, P.R. Fielden, A.J. Packham, Deconvolution of overlapping chromatographic peaks by means of fast Fourier and Hartley transforms, Analyst 121 (1996) 97-104.

[55] A. Felinger, Deconvolution of overlapping skewed peaks, Anal. Chem. 66 (1994) 3066-3072.

[56] P.B. Crilly, The use of a ratio method as an alternative to constrained deconvolution, J. Chemom. 5 (1991) 85-95.

[57] J.T.V. Matos, R. Duarte, A.C. Duarte, Chromatographic response functions in 1D and 2D chromatography as tools for assessing chemical complexity, TrAC-Trend Anal. Chem. 45 (2013) 14-23.

[58] G. Vivo-Truyols, J.R. Torres-Lapasio, A.M. van Nederkassel, Y. van der Heyden, D.L. Massart, Automatic program for peak detection and deconvolution of multi-overlapped chromatographic signals - Part II: Peak model and deconvolution algorithms, J. Chromatogr. A 1096 (2005) 146-155.

[59] G. Vivo-Truyols, J.R. Torres-Lapasio, A. Garrido-Frenich, M.C. Garcia-Alvarez-Coque, A hybrid genetic algorithm with local search - II. Continuous variables: multibatch peak deconvolution, Chemometr. Intell. Lab. 59 (2001) 107-120.

[60] G. Vivo-Truyols, J.R. Torres-Lapasio, M.C. Garcia-Alvarez-Coque, A hybrid genetic algorithm with local search: I. Discrete variables: optimisation of complementary mobile phases, Chemometr. Intell. Lab. 59 (2001) 89-106.

[61] V. Costa, L. Pasti, N. Marchetti, F. Dondi, A. Cavazzini, Automated instrumental method for on-line fraction analysis and peak deconvolution in gradient multicomponent overloaded high performance liquid chromatography, J. Chromatogr. A 1217 (2010) 4919-4924.

[62] V. Olazabal, L. Prasad, P. Stark, J.A. Olivares, Application of wavelet transforms and an approximate deconvolution method for the resolution of noisy overlapped peaks in DNA capillary electrophoresis, Analyst 129 (2004) 73-81.

[63] G. Vivo-Truyols, J.R. Torres-Lapasio, R.D. Caballero, M.C. Garcia-Alvarez-Coque, Peak deconvolution in one-dimensional chromatography using a two-way data approach, J. Chromatogr. A 958 (2002) 35-49.

[64] Z.D. Zeng, S.T. Chin, H.M. Hugel, P.J. Marriott, Simultaneous deconvolution and re-construction of primary and secondary overlapping peak clusters in comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1218 (2011) 2301-2310.

[65] H.W. Kong, F. Ye, X. Lu, L. Guo, J. Tian, G.W. Xu, Deconvolution of overlapped peaks based

on the exponentially modified Gaussian model in comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1086 (2005) 160-164.

[66] L.A.F. de Godoy, E.C. Ferreira, M.P. Pedroso, C.H.d.V. Fidelis, F. Augusto, R.J. Poppi, Quantification of kerosene in gasoline by Comprehensive Two-Dimensional Gas Chromatography and N-way multivariate analysis, Anal. Lett. 41 (2008) 1603-1614.

[67] L.A. Fonseca de Godoy, L.W. Hantao, M. Pozzobon Pedroso, R.J. Poppi, F. Augusto, Quantitative analysis of essential oils in perfume using multivariate curve resolution combined with comprehensive two-dimensional gas chromatography, Anal. Chim. Acta 699 (2011) 120-125.

[68] C. Tistaert, H.P. Bailey, R.C. Allen, Y. van der Heyden, S.C. Rutan, Resolution of spectrally rank-deficient multivariate curve resolution: alternating least squares components in comprehensive two-dimensional liquid chromatographic analysis, J. Chemom. 26 (2012) 474-486.

[69] C.G. Fraga, B.J. Prazen, R.E. Synovec, Objective data alignment and chemometric analysis of comprehensive two-dimensional separations with run-to-run peak shifting on both dimensions, Anal. Chem. 73 (2001) 5833-5840.

[70] A. de Juan, R. Tauler, Multivariate curve resolution (MCR) from 2000: Progress in concepts and applications, Crit. Rev. Anal. Chem. 36 (2006) 163-176.

[71] R. Manne, On the resolution problem in coupled chromatography, Chemometr. Intell. Lab. 27 (1995) 89-94.

[72] M. Maeder, Evolving factor-analysis for the resolution of overlapping chromatographic peaks, Anal. Chem. 59 (1987) 527-530.

[73] E.R. Malinowski, Factor Analysis in Chemistry, third ed., Wiley, New York, USA, 2002.

[74] H.R. Keller, D.L. Massart, Peak purity control in liquid-chromatography with photodiode-array detection by a fixed size moving window evolving factor-analysis, Anal. Chim. Acta 246 (1991) 379-390.

[75] E.R. Malinowski, Window factor-analysis - theoretical derivation and application to flow-injection analysis data, J. Chemom. 6 (1992) 29-40.

[76] R. Manne, H.L. Shen, Y.Z. Liang, Subwindow factor analysis, Chemometr. Intell. Lab. 45 (1999) 171-176.

[77] M. Jalali-Heravi, H. Parastar, Assessment of the co-elution problem in gas chromatography-mass spectrometry using non-linear optimization techniques, Chemometr. Intell. Lab. 101 (2010) 1-13.

[78] Z.Y. Wu, Z.D. Zeng, P.J. Marriott, Comparative qualitative analysis of nonylphenol isomers by gas chromatography-mass spectrometry combined with chemometric resolution, J. Chromatogr. A 1217 (2010) 7759-7766.

[79] C.J. Xu, Y.Z. Liang, J.H. Jiang, Resolution of the embedded chromatographic peaks by modified orthogonal projection resolution and entropy maximization method, Anal. Lett. 33 (2000) 2105-2128.

[80] F. Gong, Y.Z. Liang, Q.S. Xu, F.T. Chau, Gas chromatography-mass spectrometry and chemometric resolution applied to the determination of essential oils in Cortex Cinnamomi, J. Chromatogr. A 905 (2001) 193-205.

[81] H.P. Bailey, S.C. Rutan, Comparison of chemometric methods for the screening of comprehensive two-dimensional liquid chromatographic analysis of wine, Anal. Chim. Acta 770 (2013) 18-28.

[82] S.E. Reichenbach, X. Tian, C. Cordero, Q. Tao, Features for non-targeted cross-sample analysis with comprehensive two-dimensional chromatography, J. Chromatogr. A 1226 (2012) 140-148.

[83] J. Kiefl, C. Cordero, L. Nicolotti, P. Schieberle, S.E. Reichenbach, C. Bicchi, Performance evaluation of non-targeted peak-based cross-sample analysis for comprehensive two-dimensional gas chromatography-mass spectrometry data and application to processed

hazelnut profiling, J. Chromatogr. A 1243 (2012) 81-90.

[84] P.J. Gemperline, A priori estimates of the elution profiles of the pure components in overlapped liquid-chromatography peaks using target factor-analysis, Journal of Chemical Information and Computer Sciences 24 (1984) 206-212.

[85] B.G.M. Vandeginste, W. Derks, G. Kateman, Multicomponent self-modeling curve resolution in high-performance liquid-chromatography by iterative target transformation analysis, Anal. Chim. Acta 173 (1985) 253-264.

[86] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometr. Intell. Lab. 2 (1987) 37-52.

[87] M. Hubert, P.J. Rousseeuw, K. van den Branden, ROBPCA: A new approach to robust principal component analysis, Technometrics 47 (2005) 64-79.

[88] F.H. Ruymgaart, A robust principal component analysis, J. Multivariate Anal. 11 (1981) 485-497.

[89] W. Windig, J. Guilment, Interactive self-modeling mixture analysis, Anal. Chem. 63 (1991) 1425-1432.

[90] W. Windig, D.A. Stephenson, Self-modeling mixture analysis of 2nd-derivative near-infrared spectral data using the simplisma approach, Anal. Chem. 64 (1992) 2735-2742.

[91] F.C. Sanchez, V. van den Bogaert, S.C. Rutan, D.L. Massart, Multivariate peak purity approaches, Chemometr. Intell. Lab. 34 (1996) 139-171.

[92] F.C. Sanchez, S.C. Rutan, M.D.G. Garcia, D.L. Massart, Resolution of multicomponent overlapped peaks by the orthogonal projection approach, evolving factor analysis and window factor analysis, Chemometr. Intell. Lab. 36 (1997) 153-164.

[93] Q.S. Xu, Y.Z. Liang, On the equivalence of window factor analysis and orthogonal projection resolution, Chemometr. Intell. Lab. 45 (1999) 335-338.

[94] J.H. Jiang, Y.Z. Liang, Y. Ozaki, On simplex-based method for self-modeling curve resolution of two-way data, Chemometr. Intell. Lab. 65 (2003) 51-65.

[95] A. de Juan, S.C. Rutan, R. Tauler, D.L. Massart, Comparison between the direct trilinear decomposition and the multivariate curve resolution-alternating least squares methods for the resolution of three-way data sets, Chemometr. Intell. Lab. 40 (1998) 19-32.

[96] A. de Juan, R. Tauler, Comparison of three-way resolution methods for non-trilinear chemical data sets, J. Chemom. 15 (2001) 749-772.

[97] J.L. Hope, A.E. Sinha, B.J. Prazen, R.E. Synovec, Evaluation of the DotMap algorithm for locating analytes of interest based on mass spectral similarity in data collected using comprehensive. two-dimensional gas chromatography coupled with time-of-flight mass spectrometry, J. Chromatogr. A 1086 (2005) 185-192.

[98] A.E. Sinha, J.L. Hope, B.J. Prazen, E.J. Nilsson, R.M. Jack, R.E. Synovec, Algorithm for locating analytes of interest based on mass spectral similarity in GC x GC-TOF-MS data: analysis of metabolites in human infant urine, J. Chromatogr. A 1058 (2004) 209-215.

[99] R.C. Allen, S.C. Rutan, Semi-automated alignment and quantification of peaks using parallel factor analysis for comprehensive two-dimensional liquid chromatography-diode array detector data sets, Anal. Chim. Acta 723 (2012) 7-17.

[100] H. Parastar, J.R. Radovic, M. Jalali-Heravi, S. Diez, J.M. Bayona, R. Tauler, Resolution and quantification of complex mixtures of polycyclic aromatic hydrocarbons in heavy fuel oil sample by means of GC x GC-TOF-MS combined to multivariate curve resolution, Anal. Chem. 83 (2011) 9289-9297.

[101] J.M. Davis, S.C. Rutan, P.W. Carr, Relationship between selectivity and average resolution in comprehensive two-dimensional separations with spectroscopic detection, J. Chromatogr. A 1218 (2011) 5819-5828.

[102] L.R. Snyder, J.C. Hoggard, T.J. Montine, R.E. Synovec, Development and application of a comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry

method for the analysis of L-beta-methylamino-alanine in human tissue, J. Chromatogr. A 1217 (2010) 4639-4647.

[103] R.E. Mohler, B.P. Tu, K.M. Dombek, J.C. Hoggard, E.T. Young, R.E. Synovec, Identification and evaluation of cycling yeast metabolites in two-dimensional comprehensive gas chromatography-time-of-flight-mass spectrometry data, J. Chromatogr. A 1186 (2008) 401-411.

[104] X. Guo, M.E. Lidstrom, Metabolite profiling analysis of Methylobacterium extorquens AM1 by comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry, Biotechnol. Bioeng. 99 (2008) 929-940.

[105] J.C. Hoggard, J.H. Wahl, R.E. Synovec, G.M. Mong, C.G. Fraga, Impurity profiling of a chemical weapon precursor for possible forensic signatures by comprehensive two-dimensional gas chromatography/mass spectrometry and chemometrics, Anal. Chem. 82 (2010) 689-698.

[106] M. Chen, S.E. Reichenbach, J. Shi, Automated unmixing of comprehensive two-dimensional chemical separations with mass spectrometry. in, 2005 IEEE International Conference on Electro/Information Technology. 2005, pp. 518-523.

[107] A.E. Sinha, C.G. Fraga, B.J. Prazen, R.E. Synovec, Trilinear chemometric analysis of two-dimensional comprehensive gas chromatography-time-of-flight mass spectrometry data, J. Chromatogr. A 1027 (2004) 269-277.

[108] A.E. Sinha, J.L. Hope, B.J. Prazen, C.G. Fraga, E.J. Nilsson, R.E. Synovec, Multivariate selectivity as a metric for evaluating comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry subjected to chemometric peak deconvolution, J. Chromatogr. A 1056 (2004) 145-154.

[109] E. Sanchez, B.R. Kowalski, Generalized rank annihilation factor-analysis, Anal. Chem. 58 (1986) 496-499.

[110] E. Sanchez, L.S. Ramos, B.R. Kowalski, Generalized rank annihilation method .1. application to liquid-chromatography diode-array ultraviolet detection data, J. Chromatogr. 385 (1987) 151-164.

[111] L.S. Ramos, E. Sanchez, B.R. Kowalski, Generalized rank annihilation method .2. analysis of bimodal chromatographic data, J. Chromatogr. 385 (1987) 165-180.

[112] E. Sanchez, B.R. Kowalski, Tensorial resolution: A direct trilinear decomposition, J. Chemom. 4 (1990) 29-45.

[113] R. Bro, C.A. Andersson, H.A.L. Kiers, PARAFAC2 - Part II. Modeling chromatographic data with retention time shifts, J. Chemom. 13 (1999) 295-309.

[114] H.A.L. Kiers, J.M.F. Ten Berge, R. Bro, PARAFAC2 - Part I. A direct fitting algorithm for the PARAFAC2 model, J. Chemom. 13 (1999) 275-294.

[115] C.A. Andersson, R. Bro, Improving the speed of multi-way algorithms: Part I. Tucker3, Chemometr. Intell. Lab. 42 (1998) 93-103.

[116] P. Paatero, C.A. Andersson, Further improvements of the speed of the Tucker3 three-way algorithm, Chemometr. Intell. Lab. 47 (1999) 17-20.

[117] T. Azzouz, R. Tauler, Application of multivariate curve resolution alternating least squares (MCR-ALS) to the quantitative analysis of pharmaceutical and agricultural samples, Talanta 74 (2008) 1201-1210.

[118] M. Esteban, C. Arino, J.M. Diaz-Cruz, M.S. Diaz-Cruz, R. Tauler, Multivariate curve resolution with alternating least squares optimisation: a soft-modelling approach to metal complexation studies by voltammetric techniques, TrAC-Trend Anal. Chem. 19 (2000) 49-61.

[119] M. Garrido, F.X. Rius, M.S. Larrechi, Multivariate curve resolution-alternating least squares (MCR-ALS) applied to spectroscopic data from monitoring chemical reactions processes, Anal. Bioanal. Chem. 390 (2008) 2059-2066.

[120] C.B. Zachariassen, J. Larsen, F. van den Berg, R. Bro, A. de Juan, R. Tauler, Comparison of PARAFAC2 and MCR-ALS for resolution of an analytical liquid dilution system, Chemometr. Intell. Lab. 83 (2006) 13-25.

[121] H.P. Bailey, S.C. Rutan, Chemometric resolution and quantification of four-way data arising from comprehensive 2D-LC-DAD analysis of human urine, Chemometr. Intell. Lab. 106 (2011) 131-141.

[122] T. Skov, J.C. Hoggard, R. Bro, R.E. Synovec, Handling within run retention time shifts in two-dimensional chromatography data using shift correction and modeling, J. Chromatogr. A 1216 (2009) 4020-4029.

[123] P.M. Harvey, R.A. Shellie, Data reduction in comprehensive two-dimensional gas chromatography for rapid and repeatable automated data analysis, Anal. Chem. 84 (2012) 6501-6507.

[124] L.L.P. van Stee, U.A.T. Brinkman, Peak clustering in two-dimensional gas chromatography with mass spectrometric detection based on theoretical calculation of two-dimensional peak shapes: The 2DAid approach, J. Chromatogr. A 1218 (2011) 7878-7885.

[125] M.J.E. Trudgett, G. Guiochon, R.A. Shalliker, Theoretical description of a new analytical technique: Comprehensive online multidimensional fast Fourier transform separations, J. Chromatogr. A 1218 (2011) 3545-3554.

[126] W. Windig, J.M. Phalp, A.W. Payne, A noise and background reduction method for component detection in liquid chromatography mass spectrometry, Anal. Chem. 68 (1996) 3602-3606.

[127] W. Windig, W.F. Smith, Chemometric analysis of complex coupled data - Improvements of the component detection algorithm, J. Chromatogr. A 1158 (2007) 251-257.

[128] G. Vivo-Truyols, J.R. Torres-Lapasio, A.M. van Nederkassel, Y. van der Heyden, D.L. Massart, Automatic program for peak detection and deconvolution of multi-overlapped chromatographic signals - Part I: Peak detection, J. Chromatogr. A 1096 (2005) 133-145.

[129] G. Cao, H. Cai, X.D. Cong, X. Liu, X.Q. Ma, Y.J. Lou, K.M. Qin, B.C. Cai, Global detection and analysis of volatile components from sun-dried and sulfur-fumigated herbal medicine by comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry, Analyst 137 (2012) 3828-3835.

[130] G. Vivo-Truyols, H.G. Janssen, Probability of failure of the watershed algorithm for peak detection in comprehensive two-dimensional chromatography, J. Chromatogr. A 1217 (2010) 1375-1385.

[131] I. Latha, S.E. Reichenbach, Q. Tao, Comparative analysis of peak-detection techniques for comprehensive two-dimensional chromatography, J. Chromatogr. A 1218 (2011) 6792-6798.

[132] S. Peters, G. Vivo-Truyols, P.J. Marriott, P.J. Schoenmakers, Development of an algorithm for peak detection in comprehensive two-dimensional chromatography, J. Chromatogr. A 1156 (2007) 14-24.

[133] A.P. de la Mata, J.J. Harynuk, Limits of detection and quantification in comprehensive multidimensional separations. 1. a theoretical look, Anal. Chem. 84 (2012) 6646-6653.

[134] G. Vivo-Truyols, Bayesian approach for peak detection in two-dimensional chromatography, Anal. Chem. 84 (2012) 2622-2630.

[135] M.R. Filgueira, C.B. Castells, P.W. Carr, A simple, robust orthogonal background correction method for two-dimensional liquid chromatography, Anal. Chem. 84 (2012) 6747-6752.

[136] R.C. Allen, M.G. John, S.C. Rutan, M.R. Filgueira, P.W. Carr, Effect of background correction on peak detection and quantification in online comprehensive two-dimensional liquid chromatography using diode array detection, J. Chromatogr. A 1254 (2012) 51-61.

[137] M.F. Almstetter, I.J. Appel, K. Dettmer, M.A. Gruber, P.J. Oefner, Comparison of two algorithmic data processing strategies for metabolic fingerprinting by comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry, J. Chromatogr. A 1218 (2011) 7031-7038.

[138] S. Kempa, J. Hummel, T. Schwemmer, M. Pietzke, N. Strehmel, S. Wienkoop, J. Kopka, W. Weckwerth, An automated GCxGC-TOF-MS protocol for batch-wise extraction and alignment of mass isotopomer matrixes from differential C-13-labelling experiments: a case study for

photoautotrophic-mixotrophic grown Chlamydomonas reinhardtii cells, J. Basic Microb. 49 (2009) 82-91.

[139] G. Tomasi, F. van den Berg, C. Andersson, Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data, J. Chemom. 18 (2004) 231-241.

[140] N.P.V. Nielsen, J.M. Carstensen, J. Smedsgaard, Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimised warping, J. Chromatogr. A 805 (1998) 17-35.

[141] J. Vial, H. Nocairi, P. Sassiat, S. Mallipatu, G. Cognon, D. Thiebaut, A. Teillet, D.N. Rutledge, Combination of dynamic time warping and multivariate analysis for the comparison of comprehensive two-dimensional gas chromatograms Application to plant extracts, J. Chromatogr. A 1216 (2009) 2866-2872.

[142] D. Zhang, X. Nuang, F.E. Regnier, M. Zhang, Two-dimensional correlation optimized warping algorithm for aligning GCxGC-MS data, Anal. Chem. 80 (2008) 2664-2671.

[143] B. Wang, A. Fang, J. Heim, B. Bogdanov, S. Pugh, M. Libardoni, X. Zhang, DISCO: Distance and spectrum correlation optimization alignment for two-dimensional gas chromatography time-of-flight mass spectrometry-based metabolomics, Anal. Chem. 82 (2010) 5069-5081.

[144] B. Wang, A.Q. Fang, X. Shi, S.H. Kim, X. Zhang, DISCO2: A comprehensive peak alignment algorithm for two-dimensional gas chromatography time-of-flight mass spectrometry. in: D. Huang, Y. Gan, P. Premaratne and K. Han (Eds.), Lecture Notes in Bioinformatics. Springer-Verlag Berlin, Heidelberger Plaz 3, D-14197 Berlin, Germany. 2012, pp. 486-491.

[145] J. Jeong, X. Shi, X. Zhang, S. Kim, C. Shen, Model-based peak alignment of metabolomic profiling from comprehensive two-dimensional gas chromatography mass spectrometry, BMC Bioinformatics 13 (2012).

[146] S.E. Reichenbach, X. Tian, Q. Tao, E.B. Ledford, Jr., Z. Wu, O. Fiehn, Informatics for cross-sample analysis with comprehensive two-dimensional gas chromatography and high-resolution mass spectrometry (GCxGC-HRMS), Talanta 83 (2011) 1279-1288.

[147] V.G. van Mispelaar, A.C. Tas, A.K. Smilde, P.J. Schoenmakers, A.C. van Asten, Quantitative analysis of target components by comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1019 (2003) 15-29.

[148] H. Parastar, M. Jalali-Heravi, R. Tauler, Comprehensive two-dimensional gas chromatography (GC x GC) retention time shift correction and modeling using bilinear peak alignment, correlation optimized shifting and multivariate curve resolution, Chemometr. Intell. Lab. 117 (2012) 80-91.

[149] S. Kim, I. Koo, A. Fang, X. Zhang, Smith-Waterman peak alignment for comprehensive two-dimensional gas chromatography-mass spectrometry, BMC Bioinformatics 12 (2011).

[150] S. Kim, A. Fang, B. Wang, J. Jeong, X. Zhang, An optimal peak alignment for comprehensive two-dimensional gas chromatography mass spectrometry using mixture similarity measure, Bioinformatics 27 (2011) 1660-1666.

[151] R. Duarte, J.T.V. Matos, A.C. Duarte, A new chromatographic response function for assessing the separation quality in comprehensive two-dimensional liquid chromatography, J. Chromatogr. A 1225 (2012) 121-131.

[152] R. Duck, H. Sonderfeld, O.J. Schmitz, A simple method for the determination of peak distribution in comprehensive two-dimensional liquid chromatography, J. Chromatogr. A 1246 (2012) 69-75.

[153] I. Koo, Y.P. Zhao, J. Zhang, S. Kim, X. Zhang, A method of calculating the second dimension hold-up time for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1260 (2012) 193-199.

[154] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography, Anal. Chem. 77 (2005) 6426-6434.

[155] N.E. Watson, J.M. Davis, R.E. Synovec, Observations on "orthogonality" in comprehensive

two-dimensional separations, Anal. Chem. 79 (2007) 7924-7927.

[156] M.R. Pourhaghighi, M. Karzand, H.H. Girault, Orthogonality of two-dimensional separations based on conditional entropy, Anal. Chem. 83 (2011) 7676-7681.

[157] P.J. Slonecker, X.D. Li, T.H. Ridgway, J.G. Dorsey, Informational orthogonality of two dimensional chromatographic separations, Anal. Chem. 68 (1996) 682-689.

[158] S. Zhu, Effect of column combinations on two-dimensional separation in comprehensive two-dimensional gas chromatography: Estimation of orthogonality and exploring of mechanism, J. Chromatogr. A 1216 (2009) 3312-3317.

[159] Z.Y. Liu, D.G. Patterson, M.L. Lee, Geometric approach to factor-analysis for the estimation of orthogonality and practical peak-capacity in comprehensive 2-dimensional separations, Anal. Chem. 67 (1995) 3840-3845.

[160] M. Gilar, J. Fridrich, M.R. Schure, A. Jaworski, Comparison of orthogonality estimation methods for the two-dimensional separations of peptides, Anal. Chem. 84 (2012) 8722-8732.

[161] C.J. Venkatramani, J.Z. Xu, J.B. Phillips, Separation orthogonality in temperature-programmed comprehensive two-dimensional gas chromatography, Anal. Chem. 68 (1996) 1486-1492.

[162] D. Ryan, P. Morrison, P. Marriott, Orthogonality considerations in comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1071 (2005) 47-53.

[163] B. Ornais, M. Courtiade, N. Charon, J. Ponthus, D. Thiebaut, Considerations on orthogonality duality in comprehensive two-dimensional gas chromatography, Anal. Chem. 83 (2011) 7550-7554.

[164] Z.D. Zeng, H.M. Hugel, P.J. Marriott, A modeling approach for orthogonality of comprehensive two-dimensional separations, Anal. Chem. 85 (2013) 6356-6363.

[165] S.E. Reichenbach, M.T. Ni, V. Kottapalli, A. Visvanathan, Information technologies for comprehensive two-dimensional gas chromatography, Chemometr. Intell. Lab. 71 (2004) 107-120.

[166] T. Sfetsas, C. Michailof, A. Lappas, Q. Li, B. Kneale, Qualitative and quantitative analysis of pyrolysis oil by gas chromatography with flame ionization detection and comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry, J. Chromatogr. A 1218 (2011) 3317-3325.

[167] C. Oh, X. Huang, F.E. Regnier, C. Buck, X. Zhang, Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry peak sorting algorithm, J. Chromatogr. A 1179 (2008) 205-215.

[168] L.T. Vaz-Freire, M.D.R. Gomes da Silva, A.M. Costa Freitas, Comprehensive two-dimensional gas chromatography for fingerprint pattern recognition in olive oils produced by two different techniques in Portuguese olive varieties Galega Vulgar, Cobrancosa e Carrasquenha, Anal. Chim. Acta 633 (2009) 263-270.

[169] S.E. Reichenbach, M. Ni, V. Kottapalli, A. Visvanathan, E.B. Ledford, J. Oostdijk, H.C. Trap, Chemical warfare agent detection in complex environments with comprehensive two-dimensional gas chromatography. in: P.J. Gardner (Ed.), Porceedings of the Society of Photo-optical Instrumentation Engineers (Spie). Spie-Int Soc Optical Engineering, Bellingham, WA 98227-0010 USA. 2003, pp. 28-36.

[170] H.G. Schmarr, J. Bernhardt, Profiling analysis of volatile compounds from fruits using comprehensive two-dimensional gas chromatography and image processing techniques, J. Chromatogr. A 1217 (2010) 565-574.

[171] M. Daszykowski, I. Stanimirova, A. Bodzon-Kulakowska, J. Silberring, G. Lubec, B. Walczak, Start-to-end processing of two-dimensional gel electrophoretic images, J. Chromatogr. A 1158 (2007) 306-317.

[172] B.V. Hollingsworth, S.E. Reichenbach, Q.P. Tao, A. Visvanathan, Comparative visualization for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1105 (2006) 51-58.

[173] J. Vial, B. Pezous, D. Thiebaut, P. Sassiat, B. Teillet, X. Cahours, I. Rivals, The discriminant

pixel approach: A new tool for the rational interpretation of GCxGC-MS chromatograms, Talanta 83 (2011) 1295-1301.

[174] S.E. Reichenbach, M.T. Ni, D.M. Zhang, E.B. Ledford, Image background removal in comprehensive two-dimensional gas chromatography, J. Chromatogr. A 985 (2003) 47-56.

[175] K. Banerjee, S.H. Patil, S. Dasgupta, D.P. Oulkar, S.B. Patil, R. Savant, P.G. Adsule, Optimization of separation and detection conditions for the multiresidue analysis of pesticides in grapes by comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry, J. Chromatogr. A 1190 (2008) 350-357.

[176] P. Jandera, Comprehensive two-dimensional liquid chromatography - practical impacts of theoretical considerations. A review, Cent. Eur. J. Chem. 10 (2012) 844-875.

[177] A. Mostafa, M. Edwards, T. Gorecki, Optimization aspects of comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1255 (2012) 38-55.

[178] K.M. Kalili, A. de Villiers, Systematic optimisation and evaluation of on-line, off-line and stop-flow comprehensive hydrophilic interaction chromatography x reversed phase liquid chromatographic analysis of procyanidins, Part I: Theoretical considerations, J. Chromatogr. A 1289 (2013) 58-68.

[179] J. Harynuk, T. Gorecki, Experimental variables in GCxGC: A complex interplay, Am. Lab. 39 (2007) 36-39.

[180] S. Morales-Munoz, R.J.J. Vreuls, M.D.L. de Castro, Dynamic ultrasound-assisted extraction of environmental pollutants from marine sediments for comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection, J. Chromatogr. A 1086 (2005) 122-127.

[181] J. Omar, B. Alonso, M. Olivares, A. Vallejo, N. Etxebarria, Optimization of comprehensive two-dimensional gas-chromatography (GC x GC) mass spectrometry for the determination of essential oils, Talanta 88 (2012) 145-151.

[182] R.A. Shellie, L.L. Xie, P.J. Marriott, Retention time reproducibility in comprehensive two-dimensional gas chromatography using cryogenic modulation - An intralaboratory study, J. Chromatogr. A 968 (2002) 161-170.

[183] A.G. Shepherd, V. van Mispelaar, J. Nowlin, W. Genuit, M. Grutters, Analysis of naphthenic acids and derivatization agents using two-dimensional gas chromatography and mass spectrometry: impact on flow assurance predictions, Energ. Fuel. 24 (2010) 2300-2311.

[184] A. Vallejo, M. Olivares, L.A. Fernandez, N. Etxebarria, S. Arrasate, E. Anakabe, A. Usobiaga, O. Zuloaga, Optimization of comprehensive two dimensional gas chromatography-flame ionization detection-quadrupole mass spectrometry for the separation of octyl- and nonylphenol isomers, J. Chromatogr. A 1218 (2011) 3064-3069.

[185] F.L. Dorman, P.D. Schettler, L.A. Vogt, J.W. Cochran, Using computer modeling to predict and optimize separations for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1186 (2008) 196-201.

[186] S. O'Hagan, W.B. Dunn, M. Brown, J.D. Knowles, D.B. Kell, Closed-loop, multi-objective optimization of analytical instrumentation: Gas chromatography/time-of-flight mass spectrometry of the metabolomes of human serum and of yeast fermentations, Anal. Chem. 77 (2005) 290-303.

[187] S. O'Hagan, W.B. Dunn, J.D. Knowles, D. Broadhurst, R. Williams, J.J. Ashworth, M. Cameron, D.B. Kell, Closed-loop, multi-objective optimization of two-dimensional gas chromatography/mass spectrometry for serum metabolomics, Anal. Chem. 79 (2007) 464-476.

[188] J.M. Davis, D.R. Stoll, P.W. Carr, Dependence of effective peak capacity in comprehensive two-dimensional separations on the distribution of peak capacity between the two dimensions, Anal. Chem. 80 (2008) 8122-8134.

[189] G. Semard, V. Peulon-Agasse, A. Bruchet, J.P. Bouillon, P. Cardinael, Convex hull: A new method to determine the separation space used and to optimize operating conditions for

comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1217 (2010) 5449-5454.

[190] J.M. Davis, D.R. Stoll, P.W. Carr, Effect of first-dimension undersampling on effective peak capacity in comprehensive two-dimensional separations, Anal. Chem. 80 (2008) 461-473.

[191] J.V. Seeley, E.M. Libby, K.A.H. Edwards, S.K. Seeley, Solvation parameter model of comprehensive two-dimensional gas chromatography separations, J. Chromatogr. A 1216 (2009) 1650-1657.

[192] J.L. Adcock, M. Adams, B.S. Mitrevski, P.J. Marriott, Peak modeling approach to accurate assignment of first-dimension retention times in comprehensive two-dimensional chromatography, Anal. Chem. 81 (2009) 6797-6804.

[193] S. Zhu, S. He, D.R. Worton, A.H. Goldstein, Predictions of comprehensive two-dimensional gas chromatography separations from isothermal data, J. Chromatogr. A 1233 (2012) 147-151.

[194] J.V. Seeley, S.K. Seeley, Model for predicting comprehensive two-dimensional gas chromatography retention times, J. Chromatogr. A 1172 (2007) 72-83.

[195] F. Bedani, P.J. Schoenmakers, H.G. Janssen, Theories to support method development in comprehensive two-dimensional liquid chromatography - A review, J. Sep. Sci. 35 (2012) 1697-1711.

[196] K.K. Pasikanti, J. Norasmara, S. Cai, R. Mahendran, K. Esuvaranathan, P.C. Ho, E.C.Y. Chan, Metabolic footprinting of tumorigenic and nontumorigenic uroepithelial cells using two-dimensional gas chromatography time-of-flight mass spectrometry, Anal. Bioanal. Chem. 398 (2010) 1285-1293.

[197] C. Ma, H. Wang, X. Lu, H. Wang, G. Xu, B. Liu, Terpenoid metabolic profiling analysis of transgenic Artemisia annua L. by comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry, Metabolomics 5 (2009) 497-506.

[198] T. Groeger, R. Zimmermann, Application of parallel computing to speed up chemometrics for GC x GC-TOF-MS based metabolic fingerprinting, Talanta 83 (2011) 1289-1294.

[199] L. Zhang, Z.D. Zeng, C.X. Zhao, H.W. Kong, X. Lu, G.W. Xu, A comparative study of volatile components in green, oolong and black teas by using comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry and multivariate data analysis, J. Chromatogr. A (2013) in press. http://dx.doi.org/10.1016/j.chroma.2013.1006.1022.

[200] J. Vestner, S. Malherbe, M. Du Toit, H.H. Nieuwoudt, A. Mostafa, T. Gorecki, A.G.J. Tredoux, A. de Villiers, Investigation of the volatile composition of pinotage wines fermented with different malolactic starter cultures using comprehensive two-dimensional gas chromatography coupled to time-of-flight mass spectrometry (GCxGC-TOF-MS), J. Agric. Food. Chem. 59 (2011) 12732-12744.

[201] L.A. McGregor, C. Gauchotte-Lindsay, N.N. Daeid, R. Thomas, P. Daly, R.M. Kalin, Ultra resolution chemical fingerprinting of dense non-aqueous phase liquids from manufactured gas plants by reversed phase comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1218 (2011) 4755-4763.

[202] S.T. Chin, Y.B.C. Man, C.P. Tan, D.M. Hashim, Rapid profiling of animal-derived fatty acids using fast gc x gc coupled to time-of-flight mass spectrometry, J. Am. Oil Chem. Soc. 86 (2009) 949-958.

[203] Y. Qiu, X. Lu, T. Pang, C. Ma, X. Li, G. Xu, Determination of radix ginseng volatile oils at different ages by comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry, J. Sep. Sci. 31 (2008) 3451-3457.

[204] Y. Qiu, X. Lu, T. Pang, S. Zhu, H. Kong, G. Xu, Study of traditional Chinese medicine volatile oils from different geographical origins by comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry (GC x GC-TOF-MS) in combination with multivariate analysis, J. Pharm. Biomed. Anal. 43 (2007) 1721-1727.

[205] K.M. Pierce, J.L. Hope, J.C. Hoggard, R.E. Synovec, Principal component analysis based method to discover chemical differences in comprehensive two-dimensional gas

chromatography with time-of-flight mass spectrometry (GC x GC-TOF-MS) separations of metabolites in plant samples, Talanta 70 (2006) 797-804.

[206] P. Korytar, J. Parera, P.E.G. Leonards, J. de Boer, U.A.T. Brinkman, Quadrupole mass spectrometer operating in the electron-capture negative ion mode as detector for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1067 (2005) 255-264.

[207] P. Korytar, J. Parera, P.E.G. Leonards, F.J. Santos, J. de Boer, U.A.T. Brinkman, Characterization of polychlorinated n-alkanes using comprehensive two-dimensional gas chromatography-electron-capture negative. ionisation time-of-flight mass spectrometry, J. Chromatogr. A 1086 (2005) 71-82.

[208] R.E. Mohler, K.M. Dombek, J.C. Hoggard, E.T. Young, R.E. Synovec, Comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry analysis of metabolites in fermenting and respiring yeast cells, Anal. Chem. 78 (2006) 2700-2709.

[209] J.Y. Gardner, D.E. Brillhart, M.M. Benjamin, L.G. Dixon, L.M. Mitchell, J.M.D. Dimandja, The use of GC x GC/TOF MS with multivariate analysis for the characterization of foodborne pathogen bacteria profiles, J. Sep. Sci. 34 (2011) 176-185.

[210] K.J. Johnson, R.E. Synovec, Pattern recognition of jet fuels: comprehensive GC x GC with ANOVA-based feature selection and principal component analysis, Chemometr. Intell. Lab. 60 (2002) 225-237.

[211] T. Groeger, W. Welthagen, S. Mitschke, M. Schaeffer, R. Zimmermann, Application of comprehensive two-dimensional gas chromatography mass spectrometry and different types of data analysis for the investigation of cigarette particulate matter, J. Sep. Sci. 31 (2008) 3366-3374.

[212] L.W. Hantao, H.G. Aleme, M.M. Passador, E.L. Furtado, F.A.D. Ribeiro, R.J. Poppi, F. Augusto, Determination of disease biomarkers in Eucalyptus by comprehensive two-dimensional gas chromatography and multivariate data analysis, J. Chromatogr. A 1279 (2013) 86-91.

[213] B.M. Zorzetti, J.J. Harynuk, Using GC x GC-FID profiles to estimate the age of weathered gasoline samples, Anal. Bioanal. Chem. 401 (2011) 2423-2431.

[214] X. Li, Z. Xu, X. Lu, X. Yang, P. Yin, H. Kong, Y. Yu, G. Xu, Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabonomics: Biomarker discovery for diabetes mellitus, Anal. Chim. Acta 633 (2009) 257-262.

[215] V.G. van Mispelaar, A.K. Smilde, O.E. de Noord, J. Blomberg, P.J. Schoenmakers, Classification of highly similar crude oils using data sets from comprehensive two-dimensional gas chromatography and multivariate techniques, J. Chromatogr. A 1096 (2005) 156-164.

[216] Y.N. Ni, M.H. Mei, S. Kokot, One- and two-dimensional gas chromatography-mass spectrometry and high performance liquid chromatography-diode-array detector fingerprints of complex substances: A comparison of classification performance of similar, complex Rhizoma Curcumae samples with the aid of chemometrics, Anal. Chim. Acta 712 (2012) 37-44.

[217] R. Kaliszan, QSRR: quantitative structure-(chromatographic) retention relationships, Chem. Rev. 107 (2007) 3212-3246.

[218] P. Liu, W. Long, Current mathematical methods used in QSAR/QSPR studies, International Journal of Molecular Sciences 10 (2009) 1978-1998.

[219] D.T. Stanton, QSAR and QSPR model interpretation using partial least squares (PLS) analysis, Curr. Comput. Aided Drug. Des. 8 (2012) 107-127.

[220] M. Liu, X. Sheng, M. Mu, J. Zhao, P. Kong, Q. Wang, H. Xu, X. Hu, Z. Zeng, Determining octane number of gasoline by improved PLS-multidimensional gas chromatography. in: V. EspositoVinzi, M. Tenenhaus and R. Guan (Eds.), PLS '09: Proceedings of the 6th International Conference on Partial Least Squares and Related Methods. Publishing House Electronics Industry, Beijing, Peoples R China. 2009, pp. 237-240.

[221] H. Noorizadeh, M. Noorizadeh, QSRR-based estimation of the retention time of opiate and sedative drugs by comprehensive two-dimensional gas chromatography, Med. Chem. Res. 21

(2012) 1997-2005.

[222] K.M. Pierce, S.P. Schale, Predicting percent composition of blends of biodiesel and conventional diesel using gas chromatography-mass spectrometry, comprehensive two-dimensional gas chromatography-mass spectrometry, and partial least squares analysis, Talanta 83 (2011) 1254-1259.

[223] B.J. Prazen, K.J. Johnson, A. Weber, R.E. Synovec, Two-dimensional gas chromatography and trilinear partial least squares for the quantitative analysis of aromatic and naphthene content in naphtha, Anal. Chem. 73 (2001) 5677-5682.

[224] G.T. Ventura, B.R.T. Simoneit, R.K. Nelson, C.M. Reddy, The composition, origin and fate of complex mixtures in the maltene fractions of hydrothermal petroleum assessed by comprehensive two-dimensional gas chromatography, Org. Geochem. 45 (2012) 48-65.

[225] M.P. Pedroso, L.A. Fonseca de Godoy, E.C. Ferreira, R.J. Poppi, F. Augusto, Identification of gasoline adulteration using comprehensive two-dimensional gas chromatography combined to multivariate data processing, J. Chromatogr. A 1201 (2008) 176-182.

[226] A.A. D'Archivio, A. Incani, F. Ruggieri, Retention modelling of polychlorinated biphenyls in comprehensive two-dimensional gas chromatography, Anal. Bioanal. Chem. 399 (2011) 903-913.

[227] L.A. Fonseca de Godoy, M.P. Pedroso, E.C. Ferreira, F. Augusto, R.J. Poppi, Prediction of the physicochemical properties of gasoline by comprehensive two-dimensional gas chromatography and multivariate data processing, J. Chromatogr. A 1218 (2011) 1663-1667.

[228] P.M. Bastos, P. Haglund, The use of comprehensive two-dimensional gas chromatography and structure-activity modeling for screening and preliminary risk assessment of organic contaminants in soil, sediment, and surface water, J. Soil. Sediment. 12 (2012) 1079-1088.

[229] L.A. Fonseca de Godoy, M.P. Pedroso, L.W. Hantao, R.J. Poppi, F. Augusto, Quantitative analysis by comprehensive two-dimensional gas chromatography using interval Multi-way Partial Least Squares calibration, Talanta 83 (2011) 1302-1307.

[230] Y. Ren, H. Liu, X. Yao, M. Liu, An accurate QSRR model for the prediction of the GCxGC-TOF-MS retention time of polychlorinated biphenyl (PCB) congeners, Anal. Bioanal. Chem. 388 (2007) 165-172.

# Captions

**Fig. 1.** Illustration of the importance of chemometrics to extract information from GC×GC related data for real scientific problems: results interpretation, including experimental optimization, data pre-treatment, multivariate curve resolution (MCR), pattern recognition (PR), quantitative modeling and some new insights of GC×GC study with 2D data characteristics.

**Fig. 2.** Bilinear decomposition of coupled chromatographic data with spectral information on the basis of the Beer-Lambert Law (BLL). Matrix **E** is explained in Equation (1). (A) Raw data **X** can be decomposed for the situation of N pure chemical components. (B) Each component shown in (A) can be decomposed by multiplication of **C** and **S** (pure compositions). (C) Matrices **C** and **S** include chromatographic and spectral profiles of all the N compounds (i.e. from $\mathbf{C}_1$ to $\mathbf{C}_N$ and $\mathbf{S}_1$ to $\mathbf{S}_N$, respectively). This is the basis and the goal of MCR analysis.

**Fig. 3.** Flowchart of the procedure for MCR analysis.

**Fig. 4.** Illustrative introduction of the HELP method. (A) A two-component system (peaks a and b) within a co-elution window around retention time points 43 to 62. (B) Latent-projective graphs (LPGs) proposed by Kvalheim and Liang [36,37], which can be applied to find the selective elution regions (SERs) of the two compositions, respectively. For example, the projection curve of PC1-PC2 from retention points 32 to 42 corresponds to the SER of peak a, and from 63 to 81 to peak b, since the LPG is almost a linear relationship in these two regions.

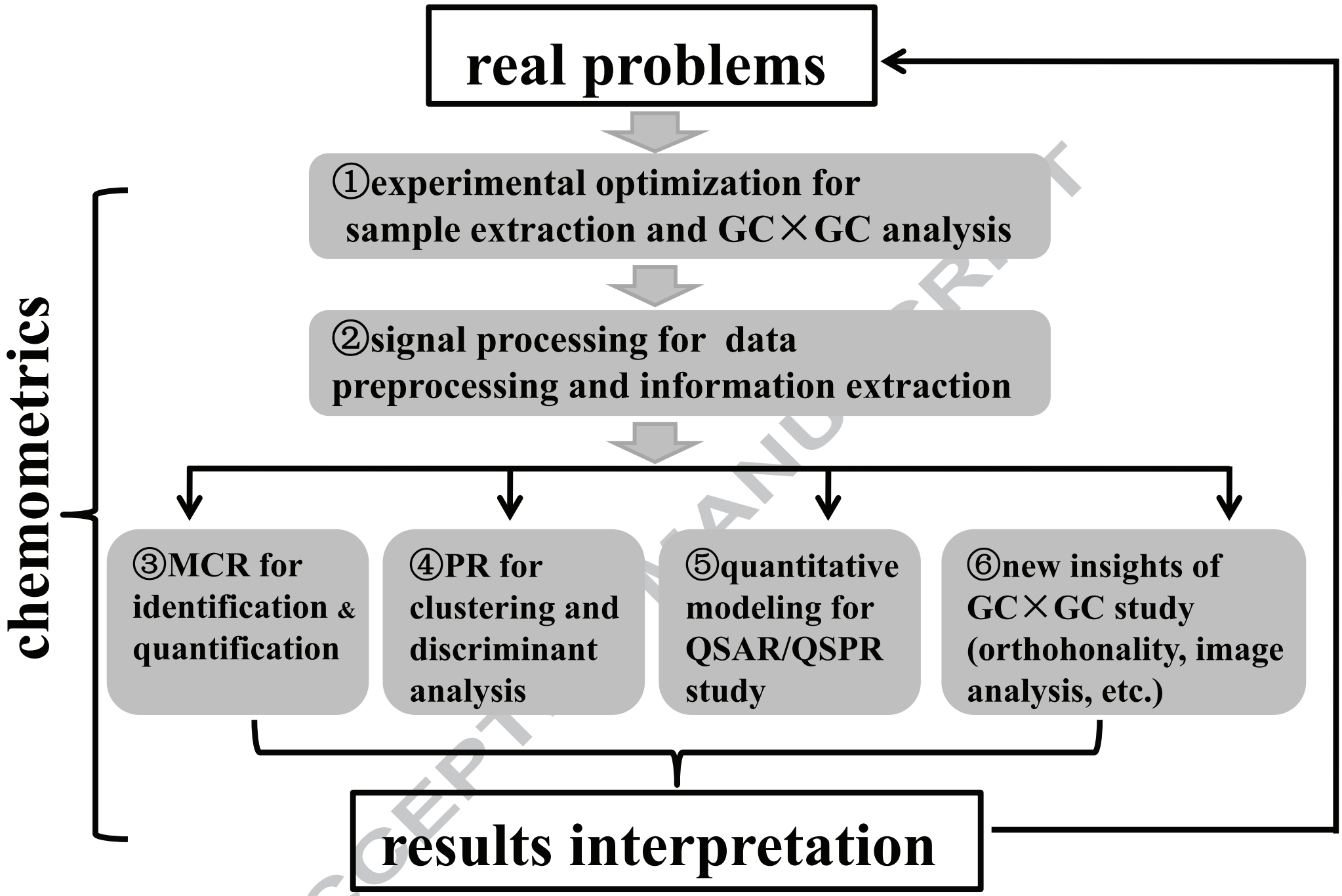
**Fig. 5.** A tri-linear decomposition of cubic data **X** with the PARAFAC model. The three parts (A), (B) and (C) have similar meanings as given in Fig. 2, but with a combination of more than one set of 2D bilinear data in **X**. For tri-linear decomposition, linear additivity is satisfied for each dimension, including **A**, **C** and **S** given in Fig. 5. Here, **C** and **S** have the same meaning as given in Equation (1), and **A** denotes the concentration of each component in different samples/runs.

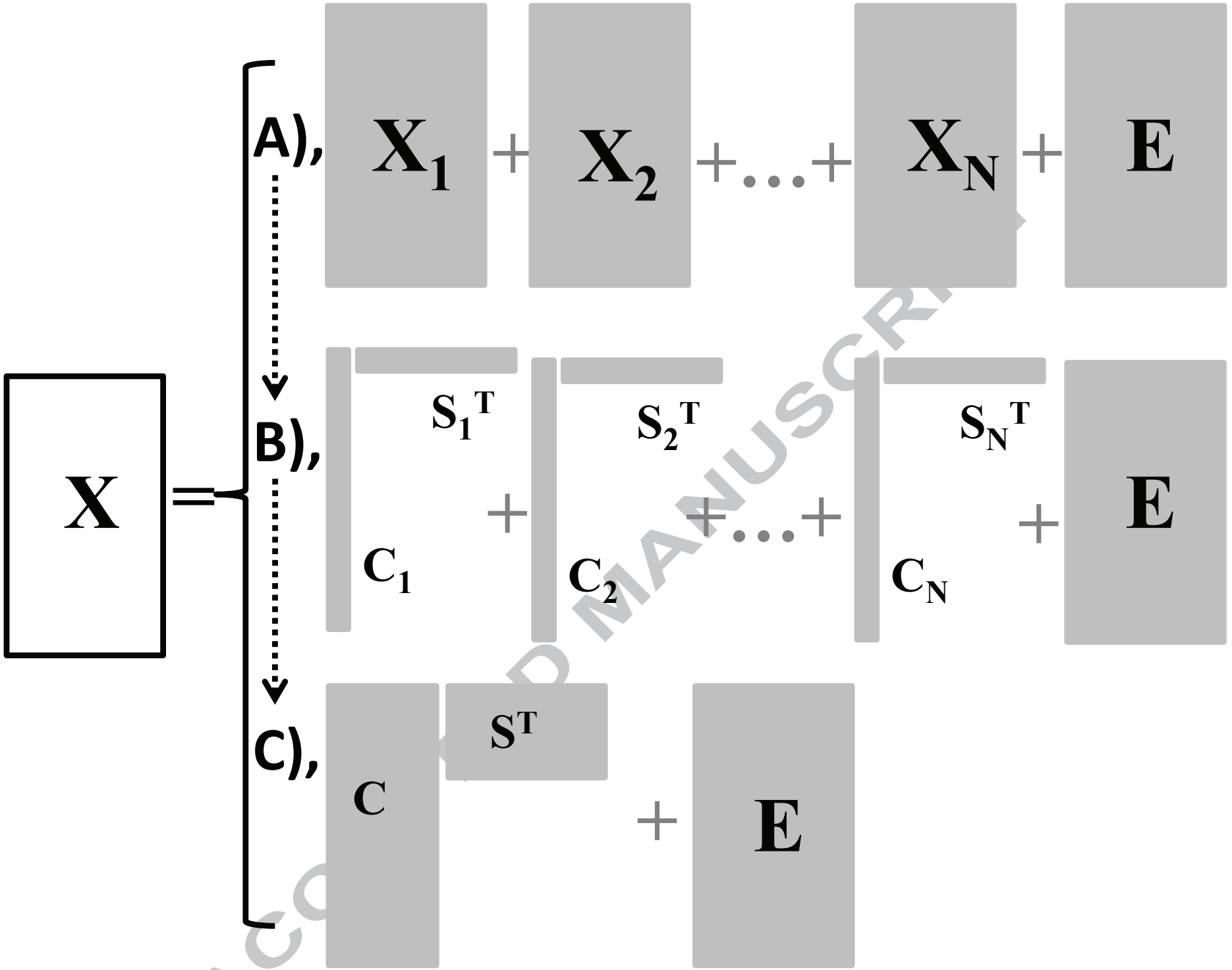
**Fig. 6.** Flowchart of the MCR-ALS method for GC×GC and GC×GC-MS data. Details are provided in the text.

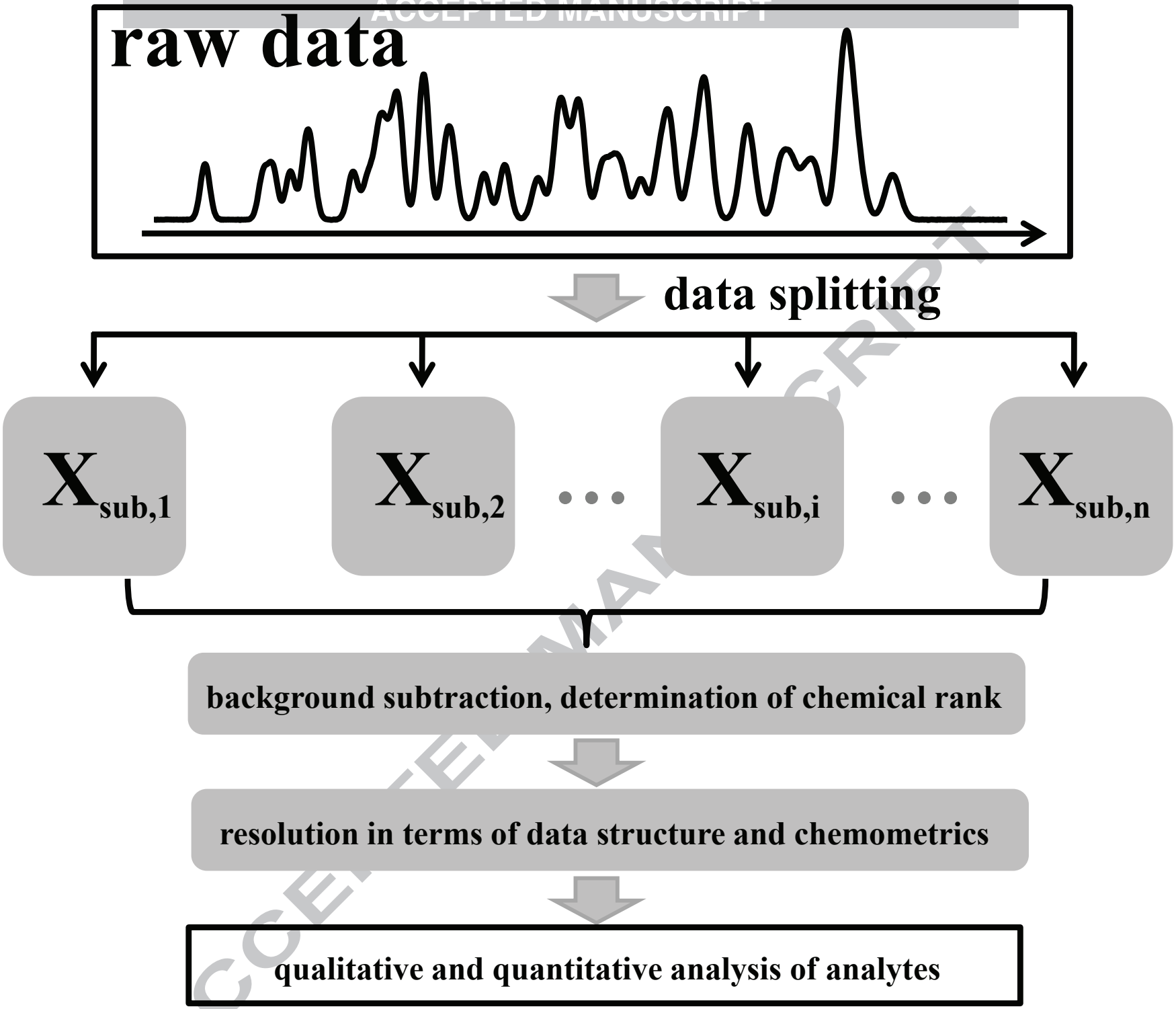
**Fig. 7.** The principle of AMWFA for alternative search to find spectra of pure components. Matrices **X** and **Y** respectively represent the two datasets to provide a moving search window and a base window. If a common compound exists in both data **X** and **Y**, it can be denoted by the abstract spectra of these two matrices after SVD analysis, respectively. Then, the spectrum of common compounds can be extracted through solving an eigen equation problem.

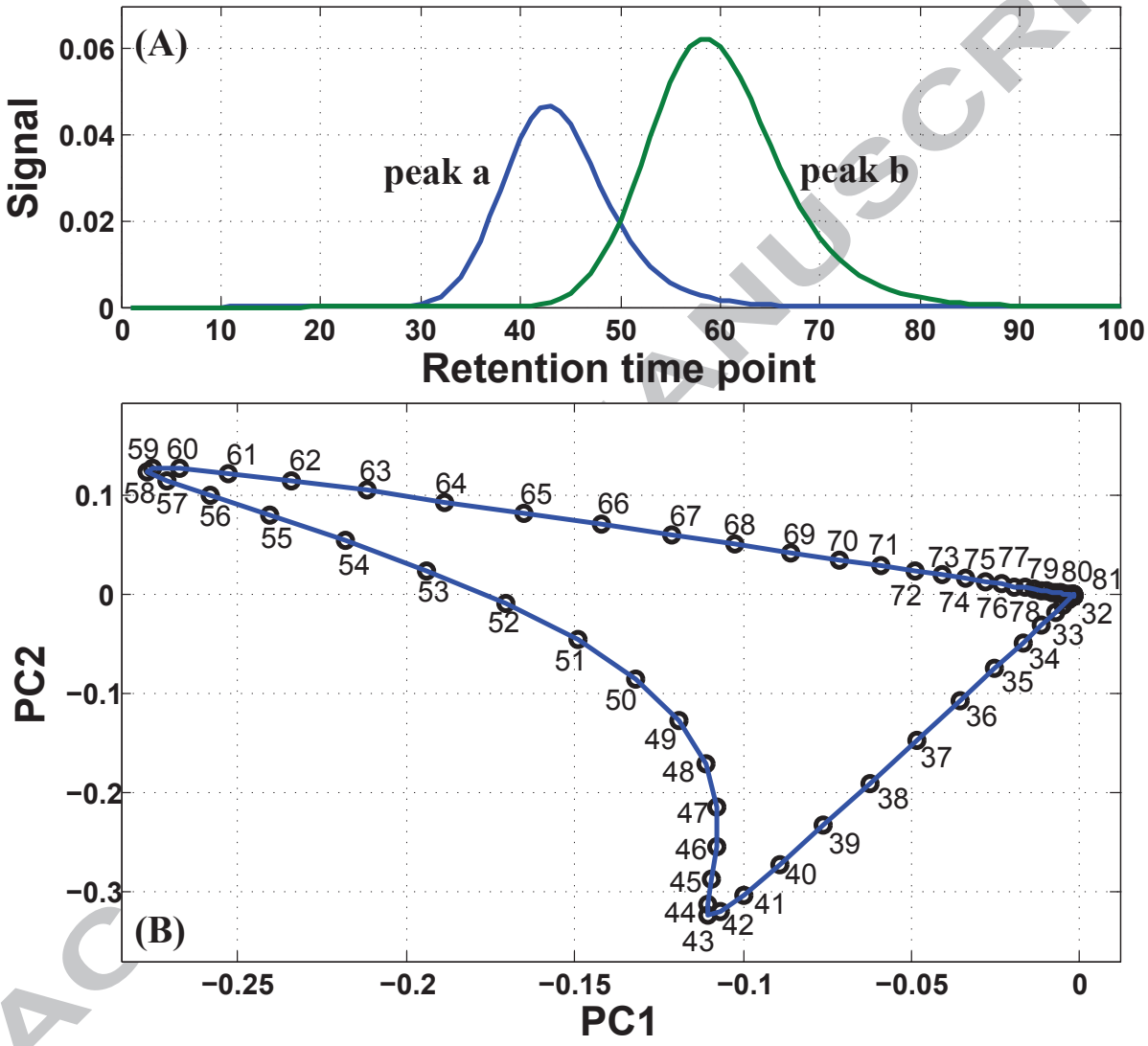
**Fig. 8.** Graphical representation of a 2D GC×GC separation. Denotations ${}^1n_c$ and ${}^2n_c$ correspond to peak capacities in ${}^1$D and ${}^2$D dimensions, respectively. The rectangular bins can be obtained to divide the 2D space, and are then employed for orthogonality evaluation.

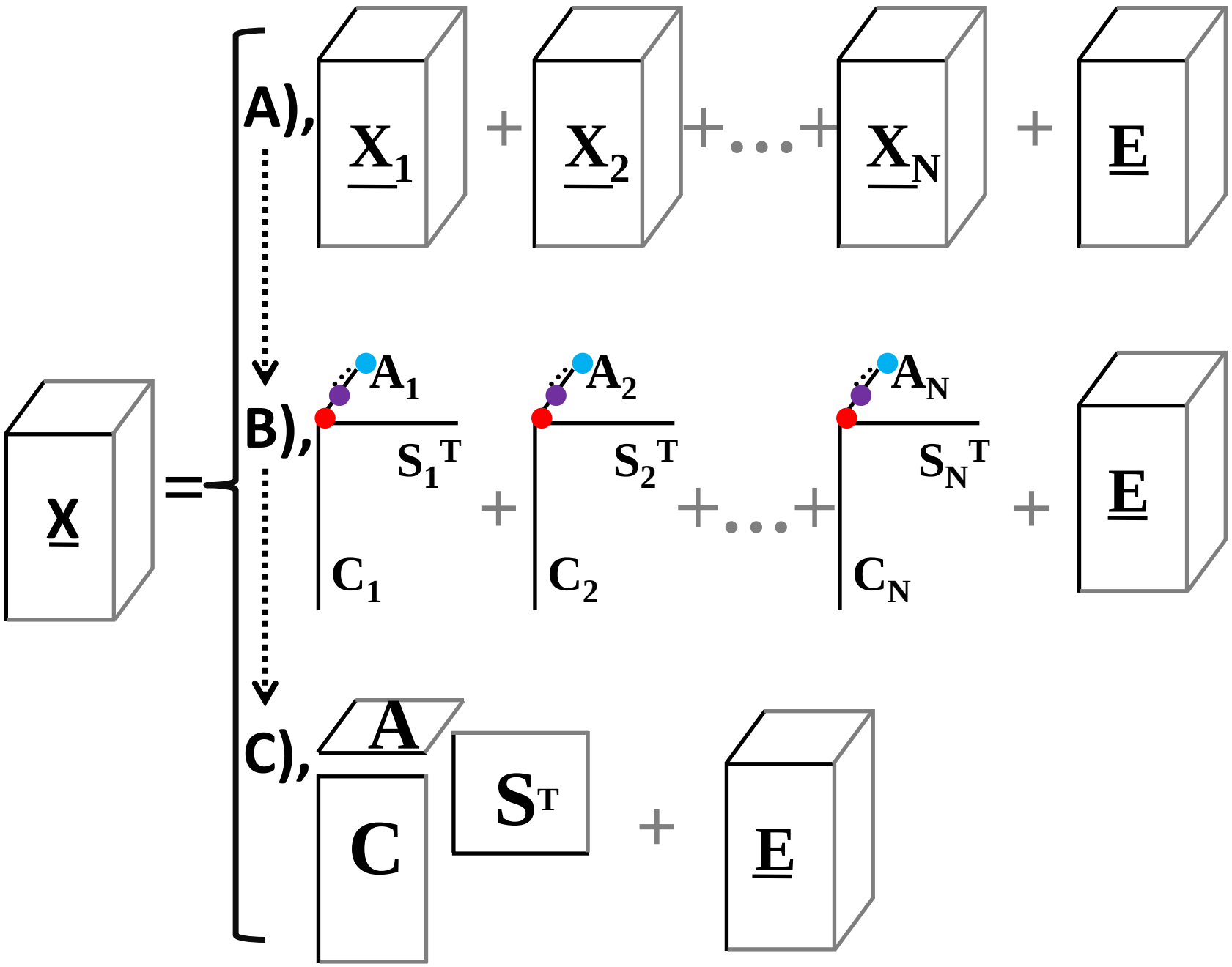
**Fig. 9.** Procedure to introduce unsupervised and supervised pattern recognitions (PRs) of GC×GC data for sample classification. The main methods for PR are included in this Figure.

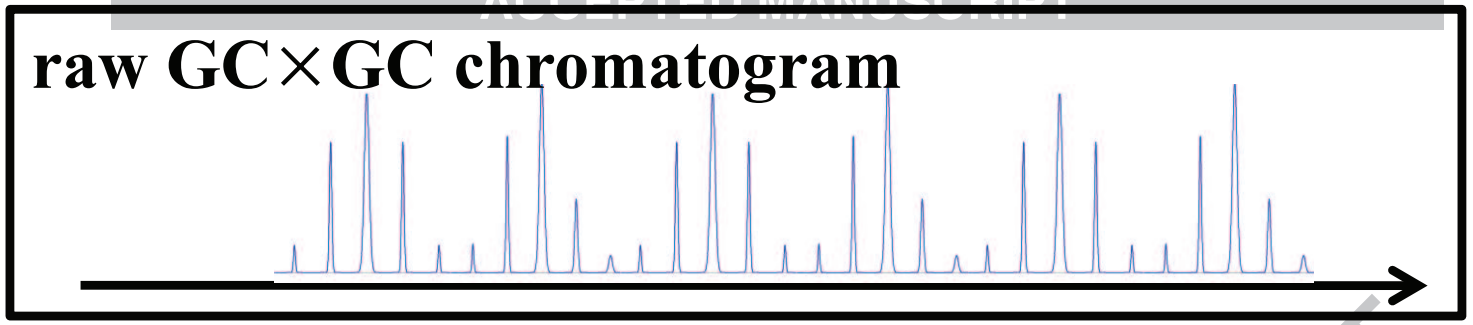









X
=
A),
X1
+
X2
+…+
XN
+
E
B),
A1
S1T
C1
+
A2
S2T
C2
+…+
AN
SNT
CN
+
E
C),
A
C
ST
+
E




data conversion

GC×GC 2D plane

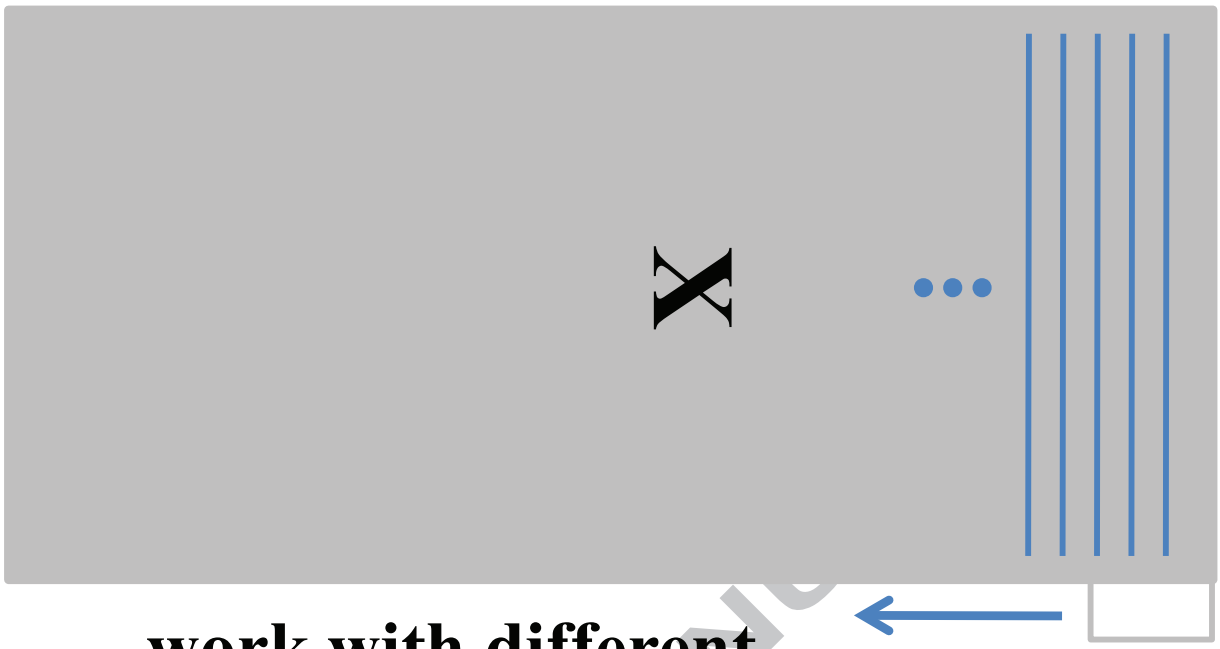
FA: fractions arrangement

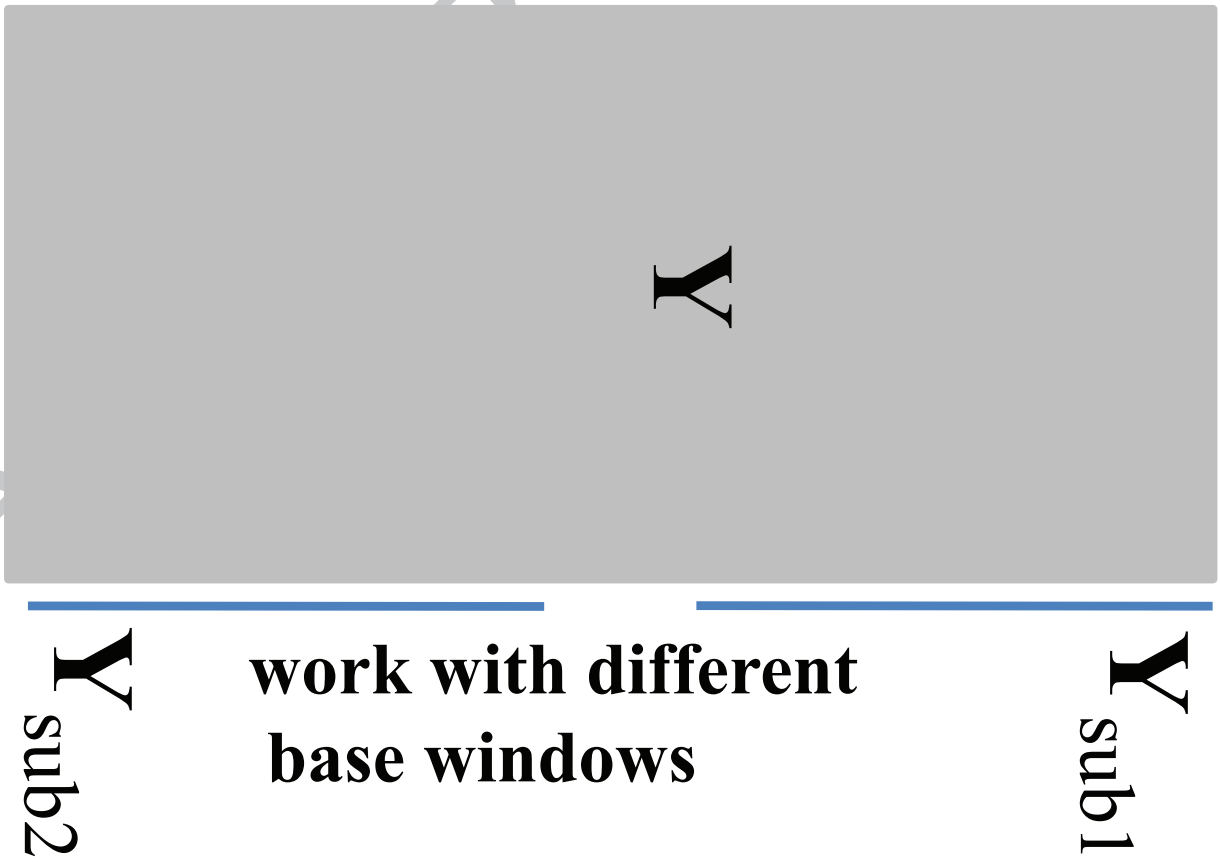
**data augment**
- ① FA of GC×GC data, 2-order data
- ② FA of GC×GC-MS data, 3-order data
- ③ FA of GC×GC-MS data with >1 run, 4-order data

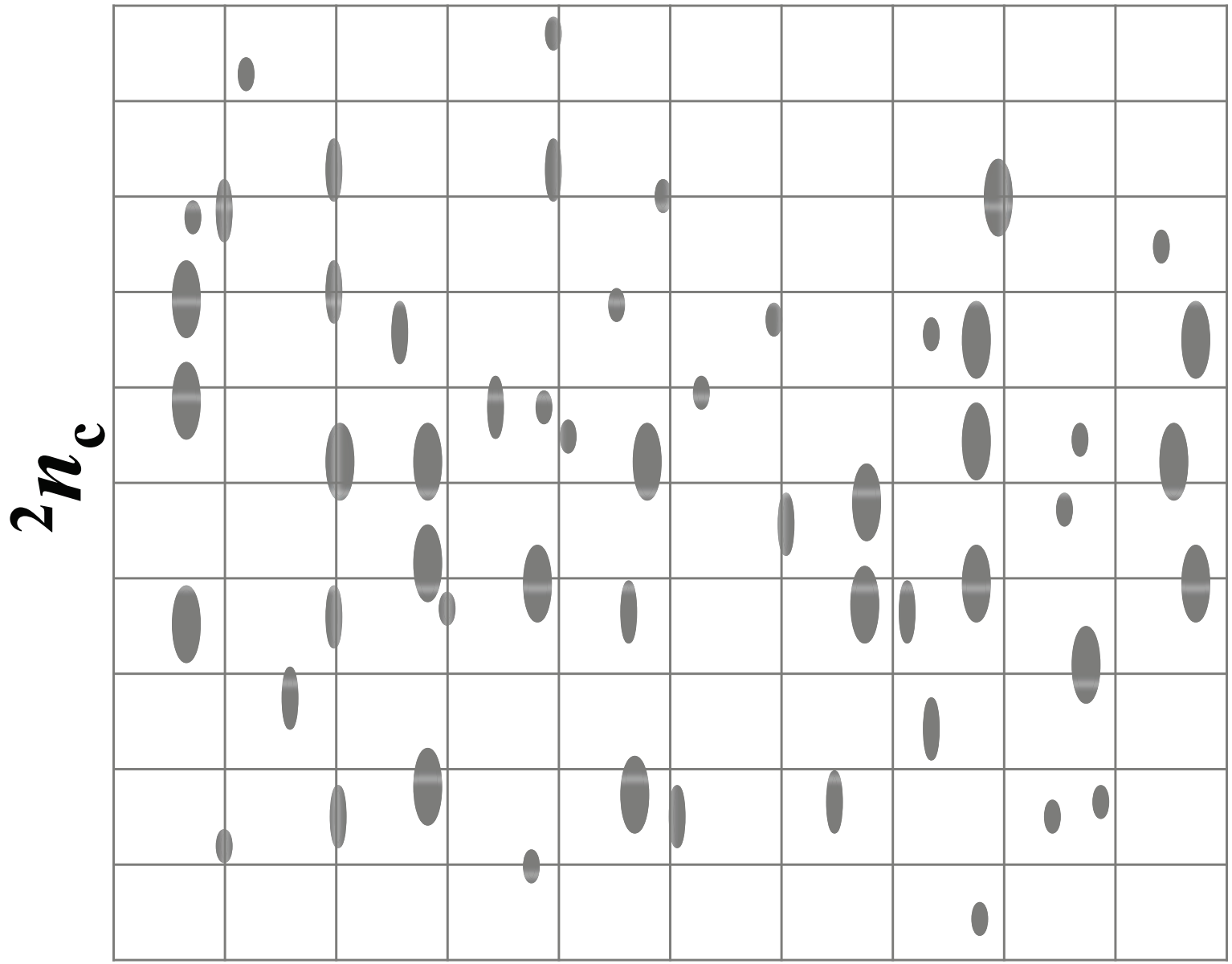
**estimation of chemical rank, and estimation of initial values for iteration**

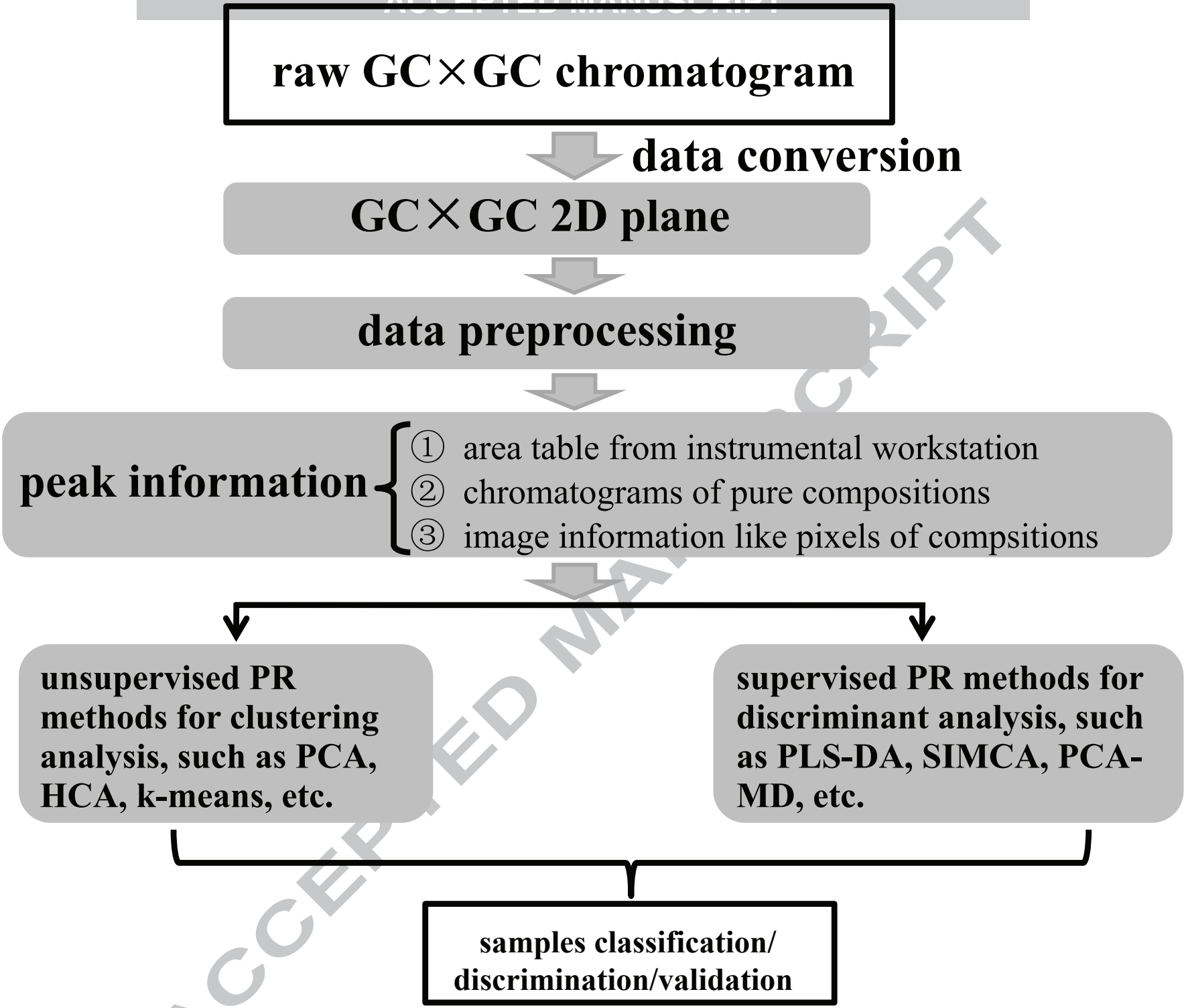
**ALS optimization** — **constraints**
- ① nonnegativity
- ② unimodality
- ③ zero-concentration regions

**qualitative and quantitative analysis**







- We present a comprehensive survey of GC×GC and GC×GC-MS data with chemometrics
- This review includes principles, theories and graphical tools for data processing
- We discuss deconvolution of 1D, coupled, and GC×GC separations with FID/MS
- We consider retention structure, $t_R$ shifts, orthogonality and image analysis