



HAL
open science

Independent components analysis (ICA) at the "cocktail-party" in analytical chemistry

Yulia Monakhova, Douglas Rutledge

► **To cite this version:**

Yulia Monakhova, Douglas Rutledge. Independent components analysis (ICA) at the "cocktail-party" in analytical chemistry. *Talanta*, 2019, pp.120451. 10.1016/j.talanta.2019.120451 . hal-02328547

HAL Id: hal-02328547


<https://agroparistech.hal.science/hal-02328547v1>

Submitted on 23 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Independent components analysis (ICA) at the “cocktail-party” in analytical chemistry

 The corrections made in this section will be reviewed and approved by journal production editor.

Yulia B. Monakhova^{a,b,c,*} yul-monakhova@mail.ru, Douglas N. Rutledge^{d,e}

^aSpectral Service AG, Emil-Hoffmann-Straße 33, 50996, Cologne, Germany

^bInstitute of Chemistry, Saratov State University, Astrakhanskaya Street 83, 410012, Saratov, Russia

^cInstitute of Chemistry, Saint Petersburg State University, 13B Universitetskaya Emb., St Petersburg, 199034, Russia

^dUMR Ingénierie Procédés Aliments, AgroParisTech, INRA, Université Paris-Saclay, Massy, France

^eNational Wine and Grape Industry Centre, Charles Sturt University, Wagga Wagga, Australia

*Corresponding author. Spectral Service AG, Emil-Hoffmann-Straße 33, 50996, Köln, Germany.

Abstract

Independent components analysis (ICA) is a probabilistic method, whose goal is to extract underlying component signals, that are maximally independent and non-Gaussian, from mixed observed signals. Since the data acquired in many applications in analytical chemistry are mixtures of component signals, such a method is of great interest. In this article recent ICA applications for quantitative and qualitative analysis in analytical chemistry are reviewed. The following experimental techniques are covered: fluorescence, UV-VIS, NMR, vibrational spectroscopies as well as chromatographic profiles. Furthermore, we reviewed ICA as a preprocessing tool as well as existing hybrid ICA-based multivariate approaches. Finally, further research directions are proposed. Our review shows that ICA is starting to play an important role in analytical chemistry, and this will definitely increase in the future.

Keywords: Independent components analysis; Chemometrics; Spectroscopy; Chromatography

1 Introduction

Experimental measurements acquired by an analytical device usually contain contributions from several interesting phenomena mixed together. A fundamental goal in the analytical community is to find the underlying, original signals that provide information about the system, but often cannot be directly or clearly seen in the observed data. Overcoming this challenge requires mathematical methods that can be efficiently applied to large data sets of chemical data and that are able to extract the underlying chemical features of interest.

Besides well-known and routinely applied chemometric methods, such as principal components analysis (PCA), cluster analysis or partial least squares regression (PLS), other interesting computational approaches are efficient but, unfortunately, not routinely used in analytical chemistry. Independent components analysis (ICA) is one such methods that deals with problems which can be symbolized by the “cocktail-party-problem” [1–3]. In this simple model, a number of microphones are situated at different locations in a room where a cocktail party is taking place. Different sound sources exist in this room: music being played, people talking, noises from the streets, etc. These sounds represent the “Source signals”. The microphones acquire different mixtures of these sounds. Providing that there are at least as many microphones as original sources signals, ICA is able to extract the original sound sources from the set of mixtures, along with the corresponding mixing proportions for each observed signal mixture. It is obvious that situations quite similar to the cocktail-party problem are to be found in analytical chemistry. For example, one would like to find the chemically interpretable spectral profiles characteristic of components or of phenomena, but one can only experimentally observe mixtures of the components.

Principal components analysis (PCA) is a classical technique in statistical data analysis, feature extraction and data reduction that aims to calculate orthogonal directions of greatest dispersion of the individuals in the multidimensional space defined by the original variables. The dispersion of the individuals being given by the variance-covariance matrix of the data, PCA is based on these second-order statistics. PCA gives vectors defining these directions of greatest dispersion (Loadings) and the coordinates of the individuals on these directions (Scores).

ICA on the other hand aims to extract “pure”, unmixed source signals from mixtures. ICA decomposes the $M \times N$ matrix of observed signal mixtures, \mathbf{X} (where M is the number of measured mixture signals, N denotes the number of variables), into a $K \times N$ matrix of K unknown source signals \mathbf{S} , and an $M \times K$ mixing matrix \mathbf{A} , containing the proportions of the sources in each mixture. All vectors are understood as column vectors.

In matrix notation,

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S}$$

The application of ICA to mixture decomposition is based on the following considerations. The pure component spectra, \mathbf{S} , show only weak dependencies, whereas mixing make the observed signals in \mathbf{X} more dependent than the sources. ICA seeks a transformation or decomposition matrix \mathbf{W} (an estimate of \mathbf{A}^{-1}) that “eliminates” the dependencies caused by mixing, $\mathbf{S} = \mathbf{W} \mathbf{X}$. The two main quantitative criteria of statistical

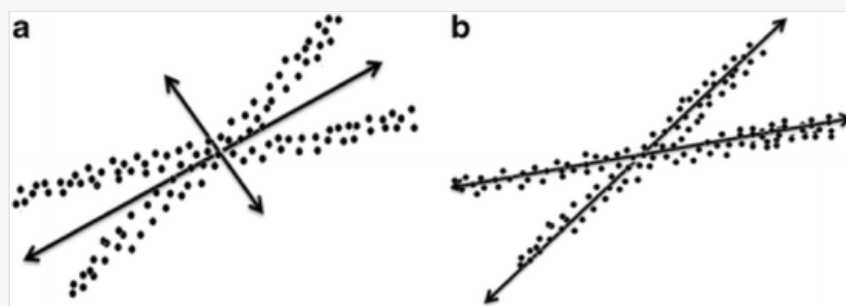
dependence are minimization of mutual information and maximization of non-Gaussianity [1–3]. The minimization of mutual information (MMI) family of ICA algorithms uses measures like Kullback-Leibler Divergence and maximum entropy. ICA algorithms based on the non-Gaussianity are motivated by the Central Limit Theorem and use kurtosis and negentropy [1–3].

ICA does not have the same statistical basis as multivariate methods based on the second-order moment (e.g., principal components analysis (PCA), factor analysis (FA) and partial least squares (PLS)) or curve resolution methods such as simple-to-use interactive self-modeling mixture analysis (SIMPLISMA) or other competing approaches such as multivariate curve resolution - alternating least squares (MCR-ALS) [4–6]. ICA looks for non-mixed, statistically independent signals. Since, as a result of the Central Limit Theorem, mixtures of signals have more Gaussian distributions than the individual signals mixed together, ICA extracts from the mixtures, vectors with maximal non-Gaussianity as measured by their fourth-order moment. ICA techniques usually do not use any training data and do not assume *a priori* knowledge about the parameters of mixing systems.

The main differences between PCA and ICA are schematically shown in Fig. 1. The first step in PCA is to perform eigenvalue decomposition of a covariance matrix. In PCA, the first eigenvector v_1 is a least-squares fit of the data cloud and best represents the variance of the data. PC1 is the vector of Scores or projections of each data point onto v_1 . The contributions of the original variables to the definition of v_1 are called the Loadings. The second eigenvector v_2 is orthogonal to v_1 , and extracts most of the remaining data variance. In contrast, the independent components IC1 and IC2 are the first two source signals and, like the Loadings of a PCA, they have contributions from the original variables. IC1 and IC2 are not only orthogonal, but also independent. There is no basic criterion for ordering the ICs, as they are all equally important, whereas in PCA the first extracted components are considered more important than the later ones, since it is assumed that more variance implies more information. Associated with these ICs are latent variables (comparable to PCA Scores) containing the Proportions of the ICs in each observed signal. There is no reason for these latent variables to be mutually orthogonal [2]. To compute the ICs, the whitened eigenvectors v_1 and v_2 are rotated using the unmixing matrix, W .

alt-text: Fig. 1

Fig. 1



Differences between PCA (a) and ICA (b). Reprinted with the permission of SpringerNature [80].



The basic ICA theory was mainly developed in the early 1990s and later summarized in several reviews. Its first applications were developed in statistics, neural networks, pattern recognition, and information theory (e.g., medical signal analysis, speech recognition, image processing, fault detection and others) [2,3,7–10]. Even now these areas remain the main application spheres of ICA approaches [11,12]. ICA has also received much attention as an effective data-mining tool for microarray gene expression data, including gene clustering, classification and identification [13]. On the other hand, the number of studies using ICA-based methods in analytical chemistry, compared to the total number of works employing this method, is still extremely small.

In this review, ICA is highlighted regarding its applications in analytical chemistry using spectroscopic and chromatographic signals as experimental data. To be concise, the fundamentals of ICA are not reviewed in this article, because sufficient literature is available on this topic [2,3,7–10]. On the contrary, an update of the main application areas of ICA in analytical chemistry since the year 2012 is not available. Other topics such as ICA-pretreatment of multivariate data and hybrid approaches will also be covered.

2 ICA applications in analytical chemistry

ICA has attracted great interest from chemists and has been applied to different types of experimental data including fluorescence, UV-VIS, NMR, vibrational spectroscopies as well as chromatographic approaches [7, 14–41]. In this regard ICA has been beneficial in solving practical problems in the fields of environment, agriculture, food products, petroleum, and medicine.

2.1 Spectroscopy

Fluorescence. Fluorescence spectroscopy can be a fast, nondestructive and sensitive technique, able to provide information about chemical composition of complex systems containing fluorophores. The increasing use of this technique has been facilitated by improved instruments and new data analysis techniques such as multivariate and multiway chemometric tools. One of these tools is ICA, which is particularly useful to improve the interpretation of fluorescence data in cases of spectral overlap and changes caused by environmental and matrix effects [15–25]. The main application areas of ICA to handle fluorescence profiles are environmental monitoring and food control.

Regarding environmental studies, synchronous fluorescence spectroscopy in combination with ICA was used to retrieve major spectral contributions diagnosing various dissolved organic carbon (DOC) constituents (tyrosine, tryptophane, humic and fulvic acids) as well as their proportions in environmental samples of natural organic matter. The DOC concentration estimation was based on a polynomial regression model between the ICA mixing matrix (the matrix of the Proportions of the Source Signals) and the reference values. The bias was estimated to be less than 6% and 13% for fulvic acid and humic acid, respectively. The proposed approach is interesting, because it can be performed using a portable spectrometer and thereby makes a stabilization step unnecessary [15]. Likewise, an ICA-based method was developed to demonstrate the feasibility of the automatic fluorimetric determination of three organic pollutants in model mixtures [16].

ICA was also applied to complete 3D fluorescence spectra of environmental samples. For example, spectral features of different phenolic homologs (phenol, thymol and m-cresol) in three-component mixtures were

extracted by ICA [17]. The proposed method increased the correct identification rate and allowed efficient water pollution monitoring. In another study, 3D solid-phase front-face fluorescence spectroscopic data of different simulated natural soil organic matter (SOM) samples were treated by ICA in order to identify major components of the organic matter as well as to reveal interactions in natural soils (e.g., self-quenching in case of cellulose) [18].

More applications of ICA are available in the area of fluorescence for food control and food monitoring. For example, Cantal-type cheese with added NaCl and KCl were analyzed for their structure at the molecular level by synchronous fluorescence spectroscopy (SFS) during gentle heating and cooling [19]. ICA treatment revealed three significant components (coenzyme/Maillard reaction products, tryptophan and vitamin A) and their proportions were correlated with rheological measurements to explain cheese texture modifications, which is important for the manufacturing of new cheese products [19].

The adulteration of orange juice by addition of grapefruit juice at the level of 1% was uncovered by 3D-front-face fluorescence spectroscopy followed by ICA [20]. Multiple linear regression (MLR) was performed to estimate the added percentage of grapefruit juice from the IC proportions based on the dataset of juice mixtures and two commercial samples labeled with known composition of both juices. Another interesting example was provided by Garcia et al. who applied 3D front-face fluorescence spectroscopy in combination with ICA to follow the development of dough [21].

ICA was also applied to monitor the antioxidant effect during the thermal evolution of corn oil samples with or without addition of natural or synthetic antioxidants [22,23]. 17 extracted ICs representing either individual fluorophores, which are degraded or formed during heating, or interpretable artifacts (baseline variations or Raleigh diffusion). The results were confirmed by NMR spectroscopy and classical methods for the determination of anisidine and viscosity values [23]. ICA applied to the 3D front-face fluorescence spectroscopic data of olive oil samples submitted to accelerated aging conditions in presence of polypropylene and polylactide plastic packaging materials revealed five significant IC profiles corresponding to natural olive oil fluorophores and newly formed oxidation products [24]. The results were used to monitor migration of antioxidants, which slow oil degradation and stabilize oils in contact with plastic.

ICA was also of interest in fluorescence beverage control. For example, 3D front-face fluorescence spectroscopic data of several wine samples for a range of pH values were modeled by ICA [25]. The results have shown that the introduction of pH as an additional parameter facilitated the discrimination of red wines according to grape variety. The reason for clearer discrimination is the pH-induced alterations in the structure of polyphenols, which makes the differences in concentrations of marker compounds more pronounced.

ICA was also applied to recover spatial structures and the corresponding spectra from multispectral fluorescence tomographic images in *in vivo* phantom and animal experiments. Moreover, fluorophore concentrations can be estimated. One such examples where ICA was applied to spectroscopic images is given by Pu et al. [26].

To conclude, ICA has been shown to be an effective tool for the analysis of 1D and 3D fluorescence spectra as well as fluorescence images by extracting pure underlying signals from a set of mixed experimental signals

with unknown proportions [15–25]. It allows the decomposition of the experimental data resulting in the extraction of individual fluorophores or interpretable artifacts, which facilitates the interpretation of the results and can be used in environmental monitoring and food control. This is possible with ICA, but not with PCA. When the 3D array, with a matrix for each sample, is unfolded into a 2D matrix, with a row vector for each sample, ICA recovers independent non-Gaussian source signals which can be folded back to give interpretable source signal matrices. PCA on the other hand nearly always extracts Loadings that are mixtures of source signals. Their interpretation in the case of normal spectral data can be delicate; in the case of unfolded 3D arrays, the results almost never resemble the spectra of fluorophores.

UV-VIS. Ultraviolet–visible spectroscopy (UV–VIS) is a complementary method to fluorescence spectroscopy, where energy absorbed during electronic transitions from the ground state to the excited state is measured. Despite being routinely used in analytical chemistry for the quantitative determination of different analytes, such as transition metal ions, highly conjugated organic compounds, and biological macromolecules, UV/Vis spectroscopy has not received much attention for ICA applications.

In the study of Rohart et al. the interactions of milk proteins and different polysaccharides under varying conditions (pH, ionic strength and temperature) were examined based on methylene blue spectrophotometric analysis combined with ICA modeling [27]. ICA identified associative interactions and characterized the strength of them under different experimental conditions. The results can be used to control interactions between different ingredients involved in formulated foods and, therefore, information can be had regarding the overall texture and stability of the products.

A fast and reliable UV-VIS spectroscopic method for multicomponent quantitative analysis of targeted compounds with overlapping signals in complex mixtures has been established [14]. The proposed approach is based on the preliminary extraction of qualitative and quantitative information from experimental profile of a calibration system by ICA. Using this quantitative model and the results obtained by ICA for unknown mixtures, the absolute analyte concentrations in multicomponent mixtures and authentic samples could then be calculated without reference solutions. The proposed method was tested on the analysis of vitamins and caffeine in energy drinks and aromatic hydrocarbons in motor fuel, with 10% error. The results demonstrated that the proposed method is a promising tool for rapid simultaneous multicomponent analysis in the case of spectral overlap and the absence/inaccessibility of reference materials [14].

Kernel independent component analysis (KICA) has been used for feature selection using hyperspectral imaging (HSI) technique, where the spectra were recorded within the spectral region of 400–720 nm. ICA results supplemented by support vector machine classifier were utilized for the identification of four varieties of qualified and adulterated oil products [28].

NMR. An historical overview of the applications of blind source separation (BSS) methodologies in NMR spectroscopy, including methods specifically designed for this type of spectroscopy is given in Ref. [29]. ICA applications were mainly restricted to food and biological objects. For example, the utility of ICA to analyze *in vivo* brain tissue samples was tested through the analysis of an artificial data [30]. Another promising application is the ICA separation of targeted resonances from confounding artifacts in *in vivo* NMR experiments [31].

ICA was applied as a statistical approach in the development of an innovative analytical methodology to uncover the most common adulterations and quality deviations of honey [32]. According to the ICA modeling, two marker compounds of honey adulteration with exogenous sugars were identified. In another study, various ICA algorithms were applied for simultaneous ¹H NMR spectroscopic determination of up to 8 organic substances in complex mixtures. Among others, constituents of honey, soft drinks, and liquids used in electronic cigarettes were studied [33]. Later, a quantitative NMR method for analysis of targeted compounds with overlapping signals in complex mixtures was established. The method is based on the combination of ICA for spectral deconvolution and the PULCON principle (pulse length based concentration determination) for quantification [34]. In the context of food surveillance, NMR spectroscopy and ICA were also used to discriminate 46 authentic rice samples according to type (Basmati, non-Basmati long-grain rice, and round-grain rice). ICA was found to be superior to classical PCA regarding the verification of rice authenticity [35].

ICA deconvolution considerably improves the application range of direct NMR spectroscopy for the analysis of complex mixtures. The results demonstrated that ICA is a promising tool for rapid simultaneous quantification of up to six components in case of spectral overlap and the absence of reference materials. NMR spectroscopy combined with ICA could be used as a screening method in the control of food and biological samples. ICA is a promising technique for decomposing NMR spectral data into components resembling metabolite resonances, and therefore has the potential to provide a data-driven alternative for making group comparisons to the use of metabolite concentrations derived from the curve-fitting of individual spectra.

Vibrational spectroscopy. As a nondestructive and convenient tool, vibrational spectroscopy has been widely shown to be a useful technique for inspection of complex objects due to its advantages of speed, noninvasive measurement, ease of use, and minimal sample preparation requirements.

Regarding mid infrared spectroscopy (MIR), one effective application was to retrieve qualitative and quantitative information about formulations of plastic materials, which can be important for automated plastic waste separation and to control the content of additives in polymer materials. For example, the combination of ICA and MIR showed great potential in sorting plastic wastes from five kinds of conventional polymers as well as differentiating low- and high-density polyethylene materials [36]. Among 28 significant ICs, only five ICs contained reliable information for the differentiation of the plastic materials. 100% discrimination rates based on the selected independent component proportions (ICA “scores”) were obtained regardless the samples properties or the spectrometer used. Attenuated total reflectance-mid infrared spectroscopy (ATR-MIR) was used to identify and quantify four different plasticizers as potential migrants in polylactide polymer material [37] (Fig. 1). A clear classification model was obtained using an ICA model with 11 significant ICs. To establish a quantitative model, the ICA proportions corresponding to specific ICs were correlated with the reference concentrations determined by gravimetric analysis of plasticizers. Sufficiently high values of validation parameter proved the good predictability of the proposed qualitative and quantitative models.

MIR spectroscopy was also used in environmental studies for the investigation of the mucilage aggregate formation process from several algae. The interpretation of the ICA decomposition results revealed that the aggregation process consisted of two phases: production of mono- and oligosaccharides with oligopeptides and their subsequent polymerization into supramolecular structures [38]. Later the MIR spectra of several samples

of normal and anomalous size aggregates of organic matter were collected and then submitted to ICA, which allowed a good spectral decomposition together with the extraction of a higher number of significant and chemically interpretable components [39].

ICA was also combined with near infrared spectroscopy (NIR) for the simultaneous quantitative analysis of two bioactive compounds, gentiopicroside and swertiamarin, in different parts of the medicinal plant *Gentiana scabra* Bunge [40]. The predictive ability of the model was evaluated based on the coefficients of variation (0.24–0.42%) and the standard deviations (0.07–1.29%). The approach can contribute to the quality control of medicinal plants during and after cultivation. Another interesting application was to use ICA for processing NIR hyperspectral images to determine the spatial distribution of peanut traces as an allergenic food contaminant in wheat flour at different stages of food manufacturing [41].

Functional near infrared spectroscopy (fNIRS) is a noninvasive method to capture brain activities according to the measurements of changes in both oxyhemoglobin and deoxyhemoglobin concentrations. Considering this experimental technique ICA was able to separate distinct spatial patterns of brain activity, which can be correlated with the specific neural events, such as finger tapping tasks [42]. ICA was also used in the time-spectral domain to isolate physiological sources of fNIRS acquired on the human forehead [43]. The IC profiles suggested that one of the two significant components represented vasodilation of cerebral arterioles while the other one represented the washout of deoxyhemoglobin either in cerebral capillaries and venules or in extra cerebral tissue.

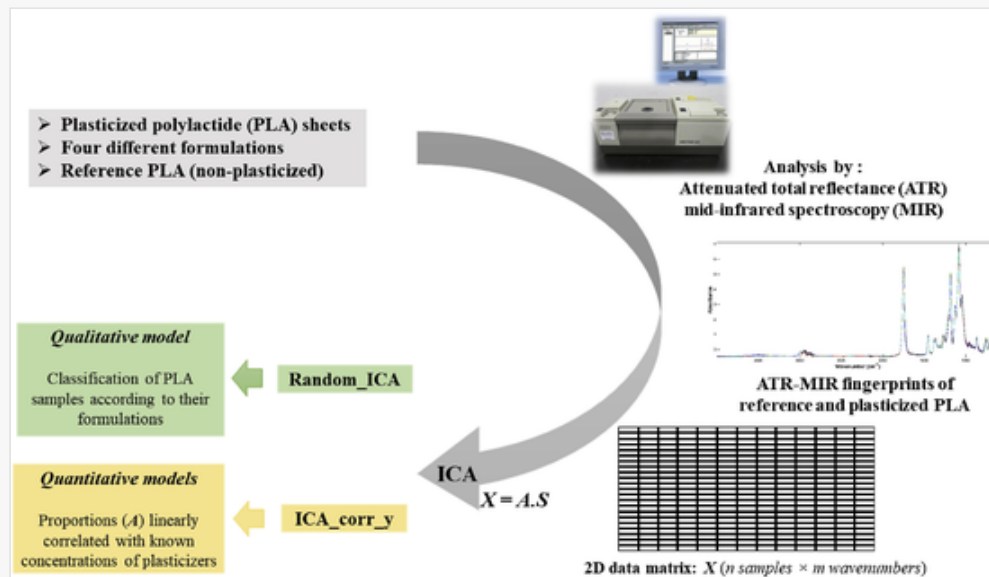
Raman spectroscopy and imaging. Raman spectroscopy and imaging is another area of application for ICA. For example, an ICA-based method was able to differentiate between single- and multi-component mixture spectra without prior knowledge of the analyzed sample in an automated manner. The method has been applied to artistic pigment identification [44].

Raman imaging. Chemical imaging spectroscopy is a technique where molecular and spatial information of a sample can be obtained simultaneously. The spatial information can be obtained by using imaging techniques and is becoming important for the control quality of different products. However, coupling spectroscopy and imaging techniques generates a huge amount of data, which can only be interpreted by using multivariate techniques.

The combination of Raman imaging and ICA has shown great potential for investigating biological and forensic samples, and pharmaceuticals. The combination of Raman imaging and ICA was used to detect ammonium nitrate/fuel oil explosives on banknotes with a limit of detection of $70 \mu\text{g}/\text{cm}^2$ after an ATM explosion experiment [45] and to detect differences in pen inks [46]. The capability of Raman spectroscopy together with ICA as a diagnostic tool for oral cancer was demonstrated in Ref. [47]. ICA was also used to examine the distribution of active ingredients and major excipients (lactose and avicel) in a pharmaceutical tablet Bipreterax® (Fig. 2) [48] as well as for the study of the homogeneity of constituents in semi-solid pharmaceutical formulations [49] (see Fig. 3 Fig. 3).

alt-text: Fig. 2

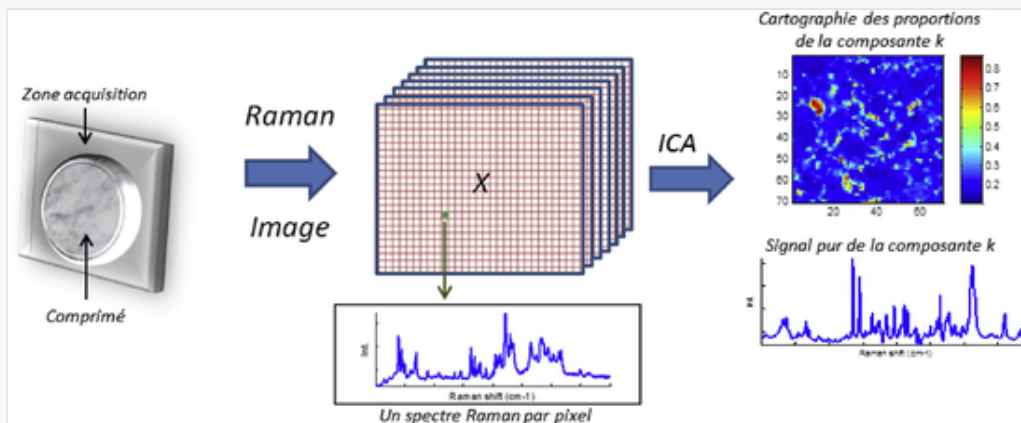
Fig. 2



ICA application on ATR-MIR data for polyactide analysis. Reprinted with the permission of Elsevier [34].

alt-text: Fig. 3

Fig. 3



ICA application on Raman images of the Bipreterax® drug product. Reprinted with the permission of Elsevier [42].


2.2 Chromatographic methods

Chromatographic profiles of complex samples involve high complexity data that must be effectively resolved to produce chemically meaningful results. However, in recent years there have not been many practical applications of ICA published regarding chromatographic techniques. In the case of standard chromatographic data (e.g. GC-FID), this is in part because it is only possible to apply multivariate analysis methods if the peaks are aligned so as to be in the same column for each sample. In the case of coupled instruments (e.g. GC-MS, LC-UV), another possibility is to have the spectral data associated with chromatogram in the columns.

ICA may be coupled with chromatographic data to extract the chromatographic profile of a targeted component from matrix interferences and background noise. As an example, the quantitative methodology to analyze the triazine pesticide prometryn in hair samples by total ion chromatography was developed [50]. The standard additions method was employed for the quantification. The efficiency of the method was validated with spiked samples, and the recoveries were in the range of 99–105%.

ICA was also used to address incomplete separation problems occurring during chromatographic analysis of complex matrices. ICA resolved the elution and spectral profiles in GC-MS and GC×GC-MS data with different numbers of components, degrees of overlap and noise, with acceptable statistical and calibration results. The approach was successfully applied to the quantification of phenanthrene and anthracene in the aromatic fraction of heavy fuel oil (HFO) analyzed by GC×GC-MS [51]. ICA was also successfully used to numerically separate coeluting LC peaks of 19 antimalarial drugs and estimate their unbiased retention times [52]. Independent components regression (ICR) was found to be an efficient tool for GC-MS compound identification, avoiding the influence of outlier data [53].

A new ICA approach called MetICA was applied to experimental mass spectrometry (MS)-based, non-targeted metabolomics data, which allowed to understand how non-targeted metabolomics data reflect biological nature and technical phenomena. The authors concluded that an optimal ICA model should be selected by optimizing the number of reliable components instead of just trying to fit the data [54].

It is clear that the range of applications of ICA to different types of chromatographic profiles should increase in the near future similarly to the alternative technique, MCR-ALS, which is widely used in chromatographic data analysis [45–57]. 

2.3 Other instrumental techniques

Apart from spectroscopic and chromatographic measurements, a small number of studies used less common types of instrumental measurements. For example, ICA enabled estimation of the Faradaic and double layer charging current components in differential pulse and anodic stripping voltammetry. Individual components were recovered with R² of more than 0.98 in comparison with the ground truth data. Voltammograms of Pb(II) and Cd(II) were constructed using only Faradaic component. Validation studies demonstrated the possibility of significant improvement of sensitivity of these voltamperic determinations [58].

The aim of another work was to propose an alternative way for wine classification and prediction based on electronic nose (e-nose) measurements. In this regard, ICA was used as a dimensionality reduction technique. Successful results have been obtained in most cases for prediction and classification [59].

3 Data pretreatment

Besides direct application to spectroscopic analysis of overlapping profiles, ICA is also an effective method for pretreatment of experimental data. ICA was used to eliminate Rayleigh and Raman scattering from the excitation-emission fluorescence data cube so that this preprocessed data could then be analyzed by parallel factor analysis (PARAFAC) resulting in more informative and noise-free models [60]. In subsequent studies, it

was shown that ICA could be used instead of PARAFAC for the complete analysis of such multi-way data [18–25]. 

ICA was used as a tool for identifying the original hemodynamic functional brain near-infrared response as the primary IC in the presence of different types of noise (instrumental noise, motion artifacts, interferences) [61]. Moreover, physiological noises (low-frequency oscillation and respiratory signals) were separated from unknown noise identified as motion artifacts. The approach was successfully applied to determine hemoglobin values of eight volunteers. A *t*-test was used to show the superiority of the ICA approach in comparison with the classical low-pass filter method [61].


It should be stressed that ICA-based preprocessing is not limited to multiway methods such as PARAFAC or to fluorescence data. ICA preprocessing could be used to selectively eliminate certain sources of variation from data sets in order to facilitate PLS regression, discriminant analysis, clustering or any other chemometric technique.

4 Determination of the number of significant components

When performing ICA, it is very useful to assess the statistical significance of the resolved components. The determination of the optimum number of ICs is a crucial step before ICA modeling. Indeed, extracting too few ICs may result in non-pure signals, still consisting of mixtures, whereas calculating too many may excessively decompose source signals and introduce noise.

In order to address this problem several methods have been proposed in the literature during recent years. The number of ICs imposed can be iteratively increased from one to 5–7 to observe the consistency of certain compounds and to detect any substantial increase of noise [15]. In the case where the ICA results are used to develop a quantitative predictive model, the behavior of the standard error of validation (SEV) as consecutive ICs are extracted can be used to restrict the number of significant ICs in the model [40]. For transcriptomics datasets a possibility of prioritizing components with respect to their reproducibility in multiple ICA runs has been shown [62].

Principal object analysis (POA), which represents simply a PCA on the transposed dataset combined with FastICA algorithm, was proposed as a robust and accurate unsupervised method to retrieve maximal spectral information regarding the number of components, the respective signal sources and their contributions [63]. Three parameters based on *a priori* information (relative bias, relative success index and signal source distribution) were used to estimate the performance of the method. The approach was applied in the context of the unsupervised data treatment of synchronous fluorescence signals of standard mixtures of tyrosine, tryptophan and phenanthroline.

The consequences of having too few or too many components in the ICA model on the resolved spectral profiles during NIR examination of distribution of active ingredient and major excipients in a pharmaceutical tablet Bipreterax® were discussed. For example, with too many the signal contribution of some substances can be spread over several ICs. On the contrary, the use of too few components led to the non-detection of a low concentration compound [47]. 

Other methods based on initial data matrix manipulations have been proposed [64,65]. The first one, called Random_ICA, which can be regarded as a generalization of the ICA_by_blocks method [65], performs several random splittings of the initial data matrix into two blocks and then compares the outputs from ICA runs with increasing numbers of ICs. The Kaiser-Meyer-Olkin (KMO) index of the transposed residual matrices obtained after progressive extraction of ICs can be also regarded as a means of determining whether the observed signals in the residual matrix are still mixtures. In another method, the correlations between calculated ICA outputs (proportions or source signals) with known physico-chemical information (known concentrations or spectra) were found to be helpful to select the optimal number of ICs. The methods were successfully tested using various simulated and experimental data sets based on IR, GC-MS, and 3D fluorescence measurements [64].

5 Hybrid approaches based on ICA

It is known that in chemometrics it is common to apply a cascade of multivariate approaches to the dataset under study. For example, PCA and cluster analysis are often applied to detect outliers and provide exploratory analysis of the data before the application of more advanced chemometric approaches. Another promising but as yet largely unexplored research direction is to increase the efficiency of classical chemometric methods by synergistic combination of their advantages within a new approach.

In the study of Shu, the ability of ICA to extract hidden information was used for feature extraction before constructing a support vector machine (SVM) model for on-line multivariate process monitoring [66]. A case study of the Tennessee Eastman benchmark process was used to confirm the efficiency of the proposed methodology by comparing it to several traditional methods, including ICA, PCA and ICA-PCA [66]. Similarly, a Genetic ICA algorithm was introduced as the feature extraction method to improve the prediction accuracy of the Back-propagation neural network method for the simultaneous determination of four exchangeable cations (K^+ , Na^+ , Ca^{2+} and Mg^{2+}) to monitor the irrigation water quality [67]. The ICA mixing matrix has also been used as an input to obtain a Bayesian posterior for a spatial source distribution [68].

ICA based on the minimization of mutual information can be used as an alternative to PCA as a preprocessing tool for LDA and FDA classification. To illustrate the performance of this ICA/DA methodology, four representative nuclear magnetic resonance (NMR) data sets of wine samples were used. The average increase in the percentage of correct classification for ICA/DA in comparison with PCA/DA varied between $6 \pm 1\%$ and $8 \pm 2\%$. The use of independent components (ICs) instead of principle components (PCs) resulted in improved classification performance of the DA methods [69].

The use of hybrid chemometric approaches also resulted in improved quality of multivariate quantitative analysis. The idea is to use the calculated mixing matrix (proportions) of chemically interpretable ICs as the input for a regression procedure for multicomponent determination. This method is called IC-based regression (ICR) [70,71] and is advantageous in those cases where the estimated ICs are very similar to the real sources since the corresponding mixing matrix gives the relative-concentration profiles of each component in the mixtures. The suitability of ICR was demonstrated using spectroscopic signals of complex biological samples [53].

In the study by Gao et Ren, the idea to combine ICA denoising properties with Elman recurrent neural network (ERNN) was implemented for the simultaneous spectrofluorimetric determination of biphenyl, naphthalene and benzotriazol in model mixtures, which are well-known carcinogenic organic pollutants in biological and environmental samples. The mixing matrix \mathbf{A} , derived from ICA, was used as input data to the ERNN model, which was trained by the descent back-propagation approach [72].

Another research group has proposed to combine ICA with PCA. A new approach accommodated one or more Gaussian components in the ICA model and uses PCA to characterize contributions from this inseparable Gaussian subspace [73]. Successfully tested on the well-known Fisher's iris and Howells' craniometric data sets, mixed ICA/PCA was proven to be of potential interest in any chemical investigation, where the authenticity of blindly separated non-Gaussian sources might otherwise be questionable.

Microarray data have an important role in identification and classification of cancer tissues. Two novel classification approaches, namely, multi-resolution independent component analysis based support vector machines (MICA-SVM) and linear discriminant analysis (MICA-LDA) were successfully tested on microarray data for cancer diagnostics [74]. Later, a similar approach was successfully applied to three cancer datasets (leukemia, breast cancer and lung cancer datasets) [75].

6 Conclusions

It is probably fair to say that in the last 10 years, ICA has shown its potential to become one of the standard multivariate tools in analytical chemistry. The generality and potential usefulness of the methodology is no longer in question, and ICA for some datasets can now compete with traditional multivariate methods, such as PCA or MCR-ALS [76]. In this review it is shown that different ICA algorithms have been successfully applied in many application areas. For the application of ICA in analytical chemistry, further developments will need to address two central challenges, namely, using *a priori* information during ICA-modeling and the choice of an ICA algorithm in each particular case.

According to our experience and reported research results, all or most ICA algorithms can successfully recover interesting sources signals in many situations. However, the accuracy of ICA algorithms used for different kinds of spectral data may not be the same [7]. Unfortunately, so far there is no general criterion for the selection of an ICA algorithm in signal processing for analytical chemistry, therefore comparison of several ICA algorithms for a specific task is recommended. The ICA algorithm needs to be more carefully selected in particular situations, for example, when the sources are continuously distributed or broad bands. Mutual Information Least Dependent Component Analysis (MILCA) and kernel ICA (KICA) are robust ICA algorithms that can be used when distribution and correlation of the sources are unknown [7,78]. Few efforts are needed to compare various ICA algorithms and to compare ICA with other multivariate approaches using different analytical datasets.

Currently, the problem of incorporating *a priori* information into ICA modeling is rather acute. Including such information not only improves convergence, but also extends the application of ICA techniques to experimental data that do not exactly satisfy ICA assumptions. As a proof of principle, an adaptive weighted

summation algorithm, which takes into account available reference signals, has been introduced in the framework of multi-objectives optimization [79].

Notes

The authors declare no competing financial interest.


Uncited references

[77]. 

Acknowledgements

Y. Monakhova acknowledges support of the Russian Science Foundation (project 18-73-10009).

References

 The corrections made in this section will be reviewed and approved by journal production editor.

- [1] M.S. Pedersen, D. Wang, J. Larsen, U. Kjems, Two-microphone separation of speech mixtures, *IEEE Trans. Neural Netw.* 19 (2008) 475–492.
- [2] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, *Neural Netw.* 13 (2000) 411–430.
- [3] A. Hyvärinen, Independent component analysis: recent advances, *Philos. Trans A Math. Phys. Eng. Sci.* 371 (2012) 20110534.
- [4] B.R., Multivariate analysis in metabolomics, *Curr. Metabolomics* 1 (2013) 92–107.
- [5] H. Abdi, L.J. Williams, Partial least squares methods: partial least squares correlation and partial least square regression, *Methods Mol. Biol.* 930 (2013) 549–579.
- [6] C.J. Gaskin, B. Happell, *Int. J. Nurs. Stud.* 51 (2014) 511–521.
- [7] G. Wang, Q. Ding, Z. Hou, Independent component analysis and its applications in signal processing for analytical chemistry, *Trends Anal. Chem.* 27 (2008) 368–376.
- [8] D.N. Rutledge, D. Jouan-Rimbaud Bouveresse, Independent components analysis with the JADE algorithm, *Trends Anal. Chem.* (2013) 22–32.
- [9] C. Jutten, J. Herault, Blind source separation of sources, part I: an adaptive algorithm based on neuromimetic architecture, *Signal Process.* 24 (1991) 1–10.
- [10] P. Comon, Independent component analysis, a new concept?, *Signal Process.* 36 (1994) 287–314.

- [11] J.R. Wessel, Testing multiple psychological processes for common neural mechanisms using EEG and independent component analysis, *Brain Topogr.* 31 (2018) 90–100.
- [12] F. Artoni, A. Delorme, S. Makeig, A visual working memory dataset collection with bootstrap Independent Component Analysis for comparison of electroencephalographic preprocessing pipelines, *Data Brief.* 22 (2018) 787–793.
- [13] W, C.R. Vanderburg, H. Gunshin, J.T., X, A review of independent component analysis application to microarray gene expression data, *Biotechniques* 45 (2008) 501–520.
- [14] Y.B. Monakhova, S.P. Mushtakova, Multicomponent quantitative spectroscopic analysis without reference substances based on ICA modelling, *Anal. Bioanal. Chem.* 409 (2017) 3319–3327.
- [15] F, J.C. da Costa Pereira J., H.D. Burrows, Direct estimation of dissolved organic carbon using synchronous fluorescence and independent component analysis (ICA): advantages of a multivariate calibration, *Environ. Monit. Assess.* 187 (2015) 703.
- [16] L, S. Ren, Integrating independent component analysis with artificial neural network to analyze overlapping fluorescence spectra of organic pollutants, *J. Fluoresc.* 22 (2012) 1595–1602.
- [17] X. Yu, Y. Zhang, G. Yin, N. Zhao, X. Xiao, C. Lu, Y. Gao, W. Zhang, Discrimination of three dimensional fluorescence spectra based on wavelet analysis and independent component analysis, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 124 (2014) 52–58.
- [18] F. Ammari, R. Bendoula, D. Jouan-Rimbaud Bouveresse, D.N. Rutledge, J.M. Roger, 3D front face solid-phase fluorescence spectroscopy combined with Independent Components Analysis to characterize organic matter in model soils, *Talanta* 125 (2014) 146–152.
- [19] M. Loudiyi, R. Karoui, D.N. Rutledge, M.C. Montel, E. Rifa, A. Aït-Kaddour, Fluorescence spectroscopy coupled with independent components analysis to monitor molecular changes during heating and cooling of Cantal-type cheeses with different NaCl and KCl contents, *J. Sci. Food Agric.* 98 (2018) 963–975.
- [20] F. Ammari, L. Redjal, D.N. Rutledge, Detection of orange juice frauds using front-face fluorescence spectroscopy and Independent Components Analysis, *Food Chem.* 168 (2015) 211–217.
- [21] R. Garcia, A. Boussard, L. Rakotozafy, J. Nicolas, J. Potus, D.N. Rutledge, C.B. Cordella, 3D-front-face fluorescence spectroscopy and independent components analysis: a new way to monitor bread dough development, *Talanta* 147 (2016) 307–314.
- [22] F. Ammari, C.B.Y. Cordella, N. Boughanmi, D.N. Rutledge, Independent components analysis applied to 3D-front-face fluorescence spectra of edible oils to study the antioxidant effect of *Nigella sativa* L. extract on the thermal stability of heated oils, *Chemometr. Intell. Lab. Syst.* 113 (2012) 32–42.

- [23] F. Ammari, D. Jouan-Rimbaud-Bouveresse, N. Boughanmi, D.N. Rutledge, Study of the heat stability of sunflower oil enriched in natural antioxidants by different analytical techniques and front-face fluorescence spectroscopy combined with Independent Components Analysis, *Talanta* 99 (2012) 323–329.
- [24] A. Kassouf, M. El Rakwe, H. Chebib, V. Ducruet, D.N. Rutledge, J. Maalouly, Independent components analysis coupled with 3D-front-face fluorescence spectroscopy to study the interaction between plastic food packaging and olive oil, *Anal. Chim. Acta* 839 (2014) 14–25.
- [25] R. Saad, D.J. Bouveresse, N. Locquet, D.N. Rutledge, Using pH variations to improve the discrimination of wines by 3D front face fluorescence spectroscopy associated to Independent Components Analysis, *Talanta* 153 (2016) 278–284.
- [26] H. Pu, G. Zhang, W. He, F. Liu, H. Guang, Y. Zhang, J. Bai, J. Luo, Resolving fluorophores by unmixing multispectral fluorescence tomography with independent component analysis, *Phys. Med. Biol.* 59 (2014) 5025–5042.
- [27] A. Rohart, D. Jouan-Rimbaud Bouveresse, D.N. Rutledge, C. Michon, Spectrophotometric analysis of polysaccharide/milk protein interactions with methylene blue using Independent Components Analysis, *Food Hydrocolloids* 43 (2015) 769–776.
- [28] Z. Han, J. Wan, L. Deng, K. Liu, Oil Adulteration identification by hyperspectral imaging using QHM and ICA, *PLoS One* 11 (2016) e0146547.
- [29] I. Toumi, S. Caldarelli, B. Torrèsani, A review of blind source separation in NMR spectroscopy, *Prog. Nucl. Magn. Reson. Spectrosc.* 81 (2014) 37–64.
- [30] R. Kalyanam, D. Boutte, K. E Hutchison, V. D Calhoun, Application of ICA to realistically simulated ^1H -MRS data, *Brain Behav.* 5 (2015) e00345.
- [31] R. Kalyanam, D. Boutte, C. Gasparovic, K.E. Hutchison, V.D. Calhoun, Group independent component analysis of MR spectra, *Brain Behav.* 3 (2013) 229–242.
- [32] M. Spiteri, E. Jamin, F. Thomas, A. Rebours, M. Lees, K.M. Rogers, D.N. Rutledge, Fast and global authenticity screening of honey using ^1H -NMR profiling, *Food Chem.* 189 (2015) 60–66.
- [33] Y.B. Monakhova, A.M. Tsikin, T. Kuballa, D.W. Lachenmeier, S.P. Mushtakova, Independent component analysis (ICA) algorithms for improved spectral deconvolution of overlapped signals in ^1H NMR analysis: application to foods and related products, *Magn. Reson. Chem.* 52 (2014) 231–240.
- [34] Y.B. Monakhova, D.W. Lachenmeier, T. Kuballa, S.P. Mushtakova, Standardless multicomponent qNMR analysis of compounds with overlapped resonances based on the combination of ICA and PULCON, *Magn. Reson. Chem.* 53 (2015) 821–828.

Y.B. Monakhova, D.N. Rutledge, A. Roßmann, H. Waiblinger, M. Mahler, M. Ilse, T. Kuballa, D.W. Lachenmeier, Determination of rice type by ^1H NMR spectroscopy in combination with different chemometric tools, *J. Chemom.* 28 (2014) 83–92.


- [36] A. Kassouf, J. Maalouly, D.N. Rutledge, H. Chebib, V. Ducruet, Rapid discrimination of plastic packaging materials using MIR spectroscopy coupled with independent components analysis (ICA), *Waste Manag.* 34 (2014) 2131–2138.
- [37] A. Kassouf, A. Ruellan, D. Jouan-Rimbaud Bouveresse, D.N. Rutledge, S. Domenek, J. Maalouly, H. Chebib, V. Ducruet, Attenuated total reflectance-mid infrared spectroscopy (ATR-MIR) coupled with independent components analysis (ICA): a fast method to determine plasticizers in polylactide (PLA), *Talanta* 147 (2016) 569–580.
- [38] M. Mecozzi, M. Pietroletti, M. Scarpiniti, R. Acquistucci, M.E. Conti, Monitoring of marine mucilage formation in Italian seas investigated by infrared spectroscopy and independent component analysis, *Environ. Monit. Assess.* 184 (2012) 6025–6036.
- [39] Y.B. Monakhova, A.M. Tsikin, S.P. Mushtakova, M. Mecozzi, Independent component analysis and multivariate curve resolution to improve spectral interpretation of complex spectroscopic data sets: application to infrared spectra of marine organic matter aggregates, *Microchem. J.* 118 (2015) 211–222.
- [40] Y.-K. Chuang, I.-C. Yang, Y.M. Lo, C.-Y. Tsai, S. Chen, Integration of independent component analysis with near-infrared spectroscopy for analysis of bioactive components in the medicinal plant *Gentiana scabra* Bunge, *J. Food Drug Anal.* 22 (2014) 336–344.
- [41] P. Mishra, C. Cordella, D.N. Rutledge, P. Barreiro, J.-M. Roger, D. Belen, Application of Independent Components Analysis with the JADE algorithm and NIR hyperspectral imaging for revealing food adulteration, *J. Food Eng.* 168 (2016) 7–15.
- [42] Z. Yuan, Spatiotemporal and time-frequency analysis of functional near infrared spectroscopy brain signals using independent component analysis, *J. Biomed. Opt.* 18 (2013) 106011.
- [43] I. Schelkanova, V. Toronov, Independent component analysis of broadband near-infrared spectroscopy data acquired on adult human head, *Biomed. Opt. Express* 3 (2012) 64–74.
- [44] J.J. González-Vidal, R. Pérez-Pueyo, M.J. Soneira, S. Ruiz-Moreno, Independent component analysis-based algorithm for automatic identification of Raman spectra applied to artistic pigments and pigment mixtures, *Appl. Spectrosc.* 69 (2015) 314–322.
- [45] M.R. Almeida, L.P.L. Logrado, J.J. Zacca, D.N. Correa, R.J. Poppi, Raman hyperspectral imaging in conjunction with independent component analysis as a forensic tool for explosive analysis: the case of an ATM explosion, *Talanta* 174 (2017) 628–632.

C.A. Teixeira, R.J. Poppi, Discriminating blue ballpoint pens inks in questioned documents by Raman imaging and mean-field approach independent component analysis (MF-ICA), *Microchem. J.* 144 (2019) 411–418.

- [47] P. Meksiarun, M. Ishigaki, V.A.C. Huck-Pezzei, C.W. Huck, K. Wongravee, H. Sato, Y. Ozaki, Comparison of multivariate analysis methods for extracting the paraffin component from the paraffin-embedded cancer tissue spectra for Raman imaging, *Sci. Rep.* 7 (2017) 44890.
- [48] M. Boiret, D.N. Rutledge, N. Gorretta, Y.M. Ginot, J.M. Roger, Application of independent component analysis on Raman images of a pharmaceutical drug product: pure spectra determination and spatial distribution of constituents, *J. Pharm. Biomed. Anal.* 90 (2014) 78–84.
- [49] H. Mitsutake, S.R. Castro, E. de Paula, R.J. Poppi, D.N. Rutledge, M.C. Breitkreitz, Comparison of different chemometric methods to extract chemical and physical information from Raman images of homogeneous and heterogeneous semi-solid pharmaceutical formulations, *Int. J. Pharm.* 552 (2018) 119–129.
- [50] W. Yu, W. Cai, X. Shao, Chemometric approach for fast analysis of prometryn in human hair by GC-MS, *J. Sep. Sci.* 36 (2013) 2277–2282.
- [51] M. Zarghani, H. Parastar, Joint approximate diagonalization of eigenmatrices as a high-throughput approach for analysis of hyphenated and comprehensive two-dimensional gas chromatographic data, *J. Chromatogr., A* 1524 (2017) 188–201.
- [52] B. Debrus, P. Lebrun, J.M. Kindenge, F. Lecomte, A. Ceccato, G. Caliaro, J.M. Mbay, B. Boulanger, R.D. Marini, E. Rozet, P. Hubert, Innovative high-performance liquid chromatography method development for the screening of 19 antimalarial drugs based on a generic approach, using design of experiments, independent component analysis and design space, *J. Chromatogr., A* 1218 (2011) 5205–5215.
- [53] X. Domingo-Almenara, A. Perera, N. Ramírez, N. Cañellas, X. Correig, J. Brezmes, Compound identification in gas chromatography/mass spectrometry-based metabolomics by blind source separation, *J. Chromatogr., A* 1409 (2015) 226–233.
- [54] Y. Liu, K. Smirnov, M. Lucio, R.D. Gougeon, H. Alexandre, P. Schmitt-Kopplin, MetICA: independent component analysis for high-resolution mass-spectrometry based non-targeted metabolomics, *BMC Bioinf.* 17 (2016) 114.
- [55] Y. Izadmanesh, E. Garreta-Lara, J.B. Ghasemi, S. Lacorte, V. Matamoros, R. Tauler, Chemometric analysis of comprehensive two dimensional gas chromatography-mass spectrometry metabolomics data, *J. Chromatogr., A* 1488 (2017) 113–125.
- [56] M. Navarro-Reig, J. Jaumot, A. Baglai, G. Vivó-Truyols, P.J. Schoenmakers, R. Tauler, Untargeted comprehensive two-dimensional liquid chromatography coupled with high-resolution

mass spectrometry analysis of rice metabolome using multivariate curve resolution, *Anal. Chem.* 89 (2017) 7675–7683.

- [57] O. Monago-Maraña, R.L. Pérez, G.M. Escandar, A. Muñoz de la Peña, T. Galeano-Díaz, Combination of liquid chromatography with multivariate curve resolution-alternating least-squares (MCR-ALS) in the quantitation of polycyclic aromatic hydrocarbons present in paprika samples, *J. Agric. Food Chem.* 64 (2016) 8254–8262.
- [58] F. Ciepiela, M. Jakubowska, Faradaic and capacitive current estimation by means of Independent Components Analysis and 1kHz sampling, *Talanta* 170 (2017) 158–164.
- [59] T. Aguilera, J. Lozano, J.A. Paredes, F.J. Alvarez, J.I. Suárez, Electronic nose based on independent component analysis combined with partial least squares and artificial neural networks for wine prediction, *Sensors* 12 (2012) 8055–8072.
- [60] D. Jouan-Rimbaud Bouveresse, H. Benabid, D.N. Rutledge, Independent component analysis as a pretreatment method for parallel factor analysis to eliminate artefacts from multiway data, *Anal. Chim. Acta* 589 (2007) 216–224.
- [61] H. Santosa, M.J. Hong, S.P. Kim, K.S. Hong, Noise reduction in functional near-infrared spectroscopy signals by independent component analysis, *Rev. Sci. Instrum.* 84 (2013) 073106.
- [62] U. Kairov, L. Cantini, A. Greco, A. Molkenov, U. Czerwinska, E. Barillot, A. Zinovyev, Determining the optimal number of independent components for reproducible transcriptomic data analysis, *BMC Genomics* 18 (1) (2017) 712.
- [63] J.C. Pereira, J.C. Azevedo, H.G. Knapik, H.D. Burrows, Unsupervised component analysis: PCA, POA and ICA data exploring - connecting the dots, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 165 (2016) 69–84.
- [64] A. Kassouf, D. Jouan-Rimbaud Bouveresse, D.N. Rutledge, Determination of the optimal number of components in independent components analysis, *Talanta* 179 (2018) 538–545.
- [65] D. Jouan-Rimbaud Bouveresse, A. Moya-González, F. Ammari, D.N. Rutledge, Two novel methods for the determination of the number of components in independent components analysis models, *Chemometr. Intell. Lab. Syst.* 112 (2012) 24–32.
- [66] C.-C. Hsu, M.-C. Chen, L.-S. Chen, Integrating independent component analysis and support vector machine for multivariate process monitoring, *Comput. Ind. Eng.* 59 (2010) 145–156.
- [67] L. Wang, D. Yang, C. Fang, Z. Chen, P.J. Lesniewski, M. Mallavarapu, R. Naidu, Application of neural networks with novel independent component analysis methodologies to a Prussian blue modified glassy carbon electrode array, *Talanta* 131 (2015) 395–403.
- [68] J. Puuronen, A. Hyvärinen, A Bayesian inverse solution using independent component analysis, *Neural Netw.* 50 (2014) 47–59.

- [69] Y.B. Monakhova, R. Godelmann, T. Kuballa, S.P. Mushtakova, D.N. Rutledge, Independent components analysis to increase efficiency of discriminant analysis methods (FDA and LDA): application to NMR fingerprinting of wine, *Talanta* 141 (2015) 60–65.
- [70] H. Kaneko, M. Arakawa, K. Funatsu, Development of a new regression analysis method using independent component analysis, *J. Chem. Inf. Model.* 48 (2008) 534–541.
- [71] M.G. Gustafsson, Independent component analysis yields chemically interpretable latent variables in multivariate regression, *J. Chem. Inf. Model.* 45 (2005) 1244–1255.
- [72] L. Gao, S. Ren, Integrating independent component analysis with artificial neural network to analyze overlapping fluorescence spectra of organic pollutants, *J. Fluoresc.* 22 (2012) 1595–15602.
- [73] R.P. Woods, L.K. Hansen, S. Strother, How many separable sources? Model selection in independent components analysis, *PLoS One* 10 (2015) e0118877.
- [74] H. Han, X.L. Li, Multi-resolution independent component analysis for high-performance tumor classification and biomarker discovery, *BMC Bioinf.* 12 (2011) S7.
- [75] H. Saberhari, M. Shamsi, M. Joroughi, F. Golabi, M.H. Sedaaghi, Cancer classification in microarray data using a hybrid selective independent component analysis and ν -support vector machine algorithm, *J. Med. Signals Sens.* 4 (2014) 2 91-298.
- [76] H. Parastar, M. Jalali-Heravi, R. Tauler, Is independent component analysis appropriate for multivariate resolution in analytical chemistry?, *Trends Anal. Chem.* 31 (2012) 134–143.
- [77] Y.B. Monakhova, S.A. Astakhov, A. Kraskov, S.P. Mushtakova, Independent components in spectroscopic analysis of complex mixtures, *Chemometr. Intell. Lab. Syst.* 103 (2010) 108–115.
- [79] Y. Shi, W. Zeng, N. Wang, L. Zhao, A new method for independent component analysis with priori information based on multi-objective optimization, *J. Neurosci. Methods* 283 (2017) 72–82.
- [78]  L. Liu, C. Li, Y. Lei, J. Yin, J. Zhao, Feature extraction for hyperspectral remote sensing image using weighted PCA-ICA, *Arab. J. Geosci.* 10 (2017) 307.

Graphical abstract

alt-text: Image 1



Highlights

- ICA applications for spectral and chromatographic data modeling are reviewed.
 - Applications for food, drug and environmental control are covered.
 - ICA plays important role in data preprocessing.
 - Hybrid ICA-based approaches have a potential in analytical chemistry.
-

Queries and Answers

Query: Your article is registered as a regular item and is being processed for inclusion in a regular issue of the journal. If this is NOT correct and your article belongs to a Special Issue/Collection please contact r.saravanakumar@elsevier.com immediately prior to returning your corrections.

Answer:



Query: Please confirm that given names and surnames have been identified correctly and are presented in the desired order and please carefully verify the spelling of all authors' names.

Answer:



Query: Please confirm that the provided email “yul-monakhova@mail.ru” is the correct address for official communication, else provide an alternate e-mail address to replace the existing one, because private e-mail addresses should not be used in articles as the address for communication.

Answer:

Query: Please note that author's telephone/fax numbers are not published in Journal articles due to the fact that articles are available online and in print for many years, whereas telephone/fax numbers are changeable and therefore not reliable in the long term.

Answer: 

Query: Please note that 'Fig.3' was not cited in the text. Please check that the citation(s) suggested by the copyeditor are in the appropriate place, and correct if necessary.

Answer: 

Query: Uncited references: This section comprises references that occur in the reference list but not in the body of the text. Please cite each reference in the text or, alternatively, delete it. Any reference not dealt with will be retained in this section.

Answer: 

Query: Have we correctly interpreted the following funding source(s) and country names you cited in your article:
Russian Science Foundation, Russia?

Answer: