Chemometrics and Intelligent Laboratory Systems 128 (2013) 101–110

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Multivariate curve resolution — alternating least squares to cope with deviations from data bilinearity in ultrafast time-resolved spectroscopy

B. Debus [a,*], M. Sliwa [a], H. Miyasaka [b], J. Abe [c], C. Ruckebusch [a,*]

[a] Université Lille Nord de France, LASIR — CNRS, F-59655 Villeneuve d'Ascq, France
[b] Division of Frontier Materials Science, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan
[c] Department of Chemistry, School of Science and Engineering, Aoyama Gakuin University, 5-10-1 Fuchinobe, Chuo-ku, Sagamihara, Kanagawa 252-5258, Japan

ARTICLE INFO

Article history:
Received 6 May 2013
Received in revised form 17 July 2013
Accepted 1 August 2013
Available online 13 August 2013

Keywords:
MCR-ALS
Bilinearity
Constraints
Spectroscopy
Time-resolved
Femtosecond

ABSTRACT

Multivariate curve resolution — alternating least squares (MCR-ALS) is a powerful method to infer information about short-lived chemical intermediate states created during ultrafast chemical reaction from a series of time-resolved spectra. However, the application of MCR relates to the fulfillment of a low-rank bilinear model for the decomposition of the experimental data. In ultrafast time-resolved spectroscopy, due to the presence of vibrational relaxation, continuous spectral evolution and band broadening/narrowing are observed on top of spectral variations associated to transitions between excited states. The basic assumption mentioned above may thus sometimes be questioned.

In this paper, a methodology based on partially constrained MCR-ALS where classical constraints such as non-negativity are relaxed for some components is extended to hard- and soft-MCR. These models enable to describe deviations from ideal bilinearity in photoinduced processes. The application of these alternative MCR models is combined with the input of additional information available from the photophysics, both for hard-modeling constraint on the concentration profiles (kinetic rate constants for transitions) and for soft-modeling constraints (selective time domain for vibrational relaxation). Partially constrained MCR models are applied to two types of ultrafast spectroscopy data which show strong shifts and broadening of the signals, UV–visible and infrared femtosecond transient absorption spectra, respectively.

1. Introduction

Multivariate curve resolution — alternating least squares (MCR-ALS) is a widespread methodology to analyze process spectroscopy data [1,2]. The MCR model assumes that the data structure is low-rank bilinear. The data matrix can thus be written as a sum of rank-one pure species components, themselves decomposed in two vectors containing the profiles. Constraints are applied to concentration and spectra profiles to ensure physical meaning of pure-components.

In ideal situation, the bilinear assumption will be valid. However, as a consequence of analytical complexity, MCR developments are going along with situations where the data structure underpinned by the bilinear model can be questioned. Examples can be found in situations where component profiles are not invariant in one mode, when the distribution of the species changes in a continuous way as a function of mixture composition [3], or in the presence of interactions between different individual components [4]. Shifts of the signals and/or changes in the shape of peaks may also hinder the direct application of the MCR model. In the presence of these perturbations, rank estimations and initial estimates of component profiles might sometimes be hardly reliable. Also, criteria for convergence are more difficult to evaluate. Indeed, deviations to ideal case do not preclude acceptable MCR data description (regarding the lack of fit and residuals) providing a sufficient number of components are considered. And this is the main problem because it may lead to misinterpretation of the process data due to "over-modeling" i.e. the consideration of too many components.

Deviations from data bilinearity are encountered in a broad range of analytical domains, as in voltammetry [3,5–9], chromatography [4,10–12] and NMR spectroscopy [13]. Depending on the analytical situation, different strategies based on MCR can be employed to cope with data features causing a loss in bilinearity. In chromatography, combinations of peak alignment preprocessing methods and MCR are often applied to handle retention time shifts. In voltametry, procedures for the correction of shifts along the potential axis are adapted to the features of electrochemical signals. For hyphenated data, problems of handling shifts in elution times and changes in peak shapes are ubiquitous. For GC–MS data [14], the application of MCR-ALS enabled to handle sample-to-sample shifts by unfolding the trilinear data array in the chromatographic direction. In excitation–emission matrix (EEM) fluorescence data analysis, light scatter effects usually result in high-rank data structures which cannot be modeled by a limited number of factors [15]. Among other examples, the reliability of MCR-ALS to decompose sets

* Corresponding authors. Tel.: +33 320436661.
E-mail addresses: bruno.debus@ed.univ-lille1.fr (B. Debus), cyril.ruckebusch@univ-lille1.fr (C. Ruckebusch).

0169-7439/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.chemolab.2013.08.001

of EEMs of fluorescence of CdS quantum dots showing a non-linear spectral behavior was discussed [16]. Also, hyperspectral imaging of multi-fluorophore samples results in non-ideally bilinear data attributed to e.g. varying spectral emissions of quantum dots [17]. Alternative MCR models were thus introduced by relaxing soft-constraints for some components to accommodate the presence of artifacts such as spectral shifts. These so-called partially constrained MCR models [17] were found to provide more reliable description of the investigated systems than classical fully-constrained solutions.

Multivariate curve resolution enables to investigate ultrafast photoinduced processes from a mixture of highly overlapped time-resolved spectra [18–21]. However, the analysis of time-resolved spectroscopy data may sometimes require methods capable of coping with deviations from data bilinearity as, e.g., when studying solvation dynamics where it is unclear if the relaxation proceeds stepwise, continuously or by a combination of both mechanisms [22]. In femtosecond transient absorption spectroscopy, non-linearity of the signals results in the observation of strong spectral shifts, band broadening and continuous spectral evolution. Particular attention will be paid here to the analysis of transient absorption spectroscopy data affected by vibrational relaxation (VR) processes [23–26]. In an oversimplified picture, VR can be described as a two-step process: (i) a rapid intramolecular vibrational redistribution (IVR) of the energy supplied by the electronic excitation pulse over the vibrational or vibronic levels of the molecule (excited or ground sate) followed by (ii) intermolecular energy transfer from the molecule to the solvent [26].

In practice, VR signals are high-rank signals that may contribute significantly to the total signal variation. Their presence may biased spectral and kinetic description of the photoinduced processes investigated, for which a reaction scheme is expected to be provided. To deal with these data, we propose to extend the concept of partially constrained MCR to hybrid MCR models [2,27,28] combining hard-and-soft modeling. As mentioned above, partially constrained MCR-ALS models were initially introduced by Haaland et al. [17] in fluorescence hyperspectral imaging. The main outcome of their work is the following one: model dimensionality and pure component estimations are sometimes improved by relaxing soft-constraints (e.g. non-negativity) for some components that actually describe shifts/broadening of pure species contributions. The key idea is that classical constraints, such as non-negativity, only strictly hold for ideally linear data. This is of particular relevance for a reliable description of photoinduced processes. Indeed, VR features of the signal should not be misinterpreted as the apparition/disappearance of a new transient species. In the presence of artifacts such as spectral shifts, the relaxation of soft-modeling constraints thus permits to describe some components as perturbations of the process pure components. However, relaxing constraints may be considered an opposite approach to good practices in MCR-ALS where it is commonly admitted that the more constraints can be applied the better defined the solution is. For this reason, partially constrained MCR models are in this work considered in situations where additional information from the photophysics can be exploited for the specification of tailored constraints. The work is organized as follows. First, emphasis will be on the description of data features associated to spectral shifts and band broadening/narrowing. Then MCR-ALS basics will be resumed. Partially constrained hard- and soft-MCR models will be introduced on simulated data sets. Then the methods will be applied to two types of ultrafast time-resolved data dealing with photo-switchable systems. At first, IR femtosecond transient absorption data which are intrinsically strongly sensitive to VR will be considered. Secondly, the case of UV–visible femtosecond transient absorption data will be addressed.

2. Data sets

2.1. Spectral features of VR

Physical description of energy relaxation processes in polyatomic molecules in the condensed phase has been the subject of numerous experimental and theoretical investigations [24–26,29,30]. However, modeling the temperature dependence of VR signals is intrinsically complex and is restricted to only few simple chemical systems for which vibrational modes are perfectly quantified such as benzene [31]. A comprehensive modeling of VR processes is beyond the scope of this paper as only a qualitative description of the main spectrokinetic features is required. For this reason, a simplest realistic description of VR signals was built by considering the assumptions below:

(i) absorption bands affected by VR signals are asymmetric, featuring a one-side extended tail toward lower energies; these bands gradually change in position (shift), intensity (increasing) and shape (narrowing) as a function of time;
(ii) bandtail and bandwidth are gradually reduced as VR proceeds, the final spectrum being of symmetric shape;
(iii) the amount of energy involved during VR processes remains constant over time.

An asymmetric parametric function $f(\lambda)$ characterized by a specific Gaussian shape [32] featuring an extended tail in one side of the peak provides a relevant description of the processes investigated, as in Eq. (1),

$$f(\lambda) = A_0 e^{-\ln(2)\left(\frac{1}{T}\ln\frac{1+2T(\lambda-\lambda_{max})}{w_{1/2}}\right)^2} \quad (1)$$

where A_0 and $w_{1/2}$ refer to the magnitude and to the half width of the peak, respectively. The parameter T is related to the extent of asymmetry in the Gaussian shape (usually $-1 \leq T \leq 1$). For the simulated data considered in the current study, initial values for $w_{1/2}$ and λ_{max} were chosen in the range 50–80 nm and 500–540 nm, respectively. Meanwhile, T values were taken in the range 0.1–0.3. Since these parameters are expected to decrease with time increasing (see ii), non-linear increments were calculated on the basis of exponential decaying functions for each parameter.

Fig. 1 provides a typical spectrokinetic feature associated to pure VR signals for a chosen parameter setting (see caption). In practical data analysis, VR processes are time-dependent processes to which a characteristic time can be associated (see Fig. 1b). These signals are characterized by continuous changes in band maximum and band shape, and they result in high-rank data sets, as shown by singular value decomposition (SVD) [33] outcome (see Fig. 1c).

As mentioned above, the area under VR spectral signals remains constant over time. In photochemistry, the so-called band integral (BI) [23] can be calculated from time-resolved spectra, as defined in Eq. (2),

$$\mathrm{BI}(t, \Delta\lambda) = \int_{\Delta\lambda} f(t,\lambda)\frac{d\lambda}{\lambda} \quad (2)$$

where $\Delta\lambda$ refers to the spectral range over which the integral is calculated for the two dimensional transient signal $f(t,\lambda)$. The evaluation of BI is related to the excited-state population dynamic [34] and is therefore of great relevance for the analysis of transient absorption spectroscopy data. Transitions between excited state species will result in the observation of changes in the BI over time whereas for pure VR processes BI remains constant. In this way, time-domain corresponding to pure VR may be detected, and this information may be further used in MCR-ALS analysis.

2.2. Simulated data

Simulated data were considered on the basis of a first order kinetic model involving three species denoted [A], [B] and [C], as in Eq. (3). On top of this, spectrokinetic features of VR signals were added. Two different situations were considered. The first one is concerned with VR occurring from the final species [C], as in Eq. (4). In this equation, the curved arrow marks the changes related to VR, occurring from species [C] and leading to a relaxed form of the species (denoted by a dot).