Analytical Methods

PAPER

View Article Online
View Journal | View Issue

Cite this: *Anal. Methods*, 2014, **6**, 7883

A novel alternating least-squares method based on fixed region scanning evolving factor analysis (FRSEFA) and its application in process monitoring

Peijin Tong, Ting Wu, Xuan Wang, Han Zhang, Yan Kang and Yiping Du*

Received 29th May 2014
Accepted 22nd July 2014

DOI: 10.1039/c4ay01292e

www.rsc.org/methods

A new method called fixed region scanning evolving factor analysis (FRSEFA) is proposed based on fixed-size moving window evolving factor analysis (FSMWEFA). In FRSEFA, a base set matrix X_{bs} that is a submatrix of the entire spectra data matrix X, corresponding to a special region in the time direction, is selected. Then, the base set X_{bs} combines a moving window x_i containing only one time point to construct a submatrix $X_{t,i}$. A series of eigenvalues are obtained to accurately determine the starting time or ending point because the moving window contains only one time point, so that the problem of time shift in FSMWEFA can be solved. The combination of FRSEFA and alternating least squares (ALS) is utilized to determine the precise time of the transformation of components during the reaction and to obtain the profiles of concentration and spectrum of the components of interest in the mixture. Two real experimental data were studied by the developed algorithm. The results showed that the precise time of transformation could be obtained and spectrum and concentration profiles were also estimated, providing more useful information for reaction processes.

1. Introduction

With the development of measuring instruments, several analytical techniques can be used to monitor the entire process of a multicomponent reaction system. Real-time monitoring investigations of the multicomponent reaction system have always attracted various industries. Analysis of the data acquired from the reaction can provide a good interpretation of the changes occurring in the entire reaction, as well as the substances produced or disappeared. In addition, it can further provide the concentration of each component, providing powerful evidence for explaining the reaction mechanism more accurately.

Appropriate data analysis tools, which focus on the interpretation of multivariate responses obtained from sophisticated instrumental techniques, have made significant contributions to the progress of the analysis of multicomponent mixture processes. New monitoring data analysis methods represented by chemometrics have made remarkable progress, providing new ways to optimize various spectra of chemical processes and to analyze chemical measurement data for studying their mechanism.

In recent decades, several chemometric approaches have been proposed and used for experimental design, data preparation and exploration, sample classification, resolution of overlapping signals, calibration and pattern recognition. Their applications have been widespread in spectroscopy.[1] Multivariate resolution,[2–4] as one of the important technique studied in chemometrics, is a technology to extract the response curves (such as spectral curve, pH curve, time curve, elution curve and concentration curve, *etc.*) of the pure substance from unknown mixtures under the premise that there is no need to be aware of the type or composition of the unknown samples in advance. In 1991, Roma Tauler *et al.*[5] used self-modeling curve resolution to study the spectrometric titrations of multi-equilibrium systems. Javier Saurina *et al.*[6] applied multivariate curve resolution for the analysis of stopped-flow kinetic data of binary amino acid mixtures to simultaneously determine phenylalanine and proline. Anna de Juan and Roma Tauler[7] revisited trends in multivariate resolution, showing the application of chemometrics in unravelling information regarding pure compounds from multicomponent processes and mixtures. A study by Li Boyan and Ozaki *et al.*[8] indicated that multiple curve resolution is a powerful tool for analyzing and visualizing spectral data and integrating them with other information, making spectral intensity variations more amenable to interpretation to explore the molecular dynamics of polymers. SMCR has also been used by Slobodan Sasic[9] *et al.* to analyse the online Raman and near infrared reflection (NIR) data. Tanabe and Ozaki *et al.*[10] reported the detailed concentration profile and kinetic sorption profiles of each water component by multivariate curve resolution (MCR) in the study of multicomponent water sorption into a poly film. The aim of our study is to figure out the problem of spectral signal

Shanghai Key Laboratory of Functional Materials Chemistry, and Research Center of Analysis and Test, East China University of Science and Technology, Meilong Rd 130, Shanghai, 200237, P.R. China. E-mail: yipingdu@ecust.edu.cn

overlapping in the monitoring of multicomponent reaction system for obtaining detailed information of each component with multivariate resolution, such that it would improve the signal-to-noise ratio (SNR), accuracy, and expand the application of spectral analysis.

Various kinds of methodologies in SMCR have been developed so far.[11] In general, the determination of the number of significant factors is the first step for resolving a "black" system. The rank of a matrix is a mathematical concept that relates to the number of significant compounds in a dataset; in chemical terms, to the number of compounds in a mixture.[12] To achieve precise resolution results, the accuracy of factor number and iterative initial value are of significant importance. Numerous methods are available in literature such as evolving factor analysis (EFA),[13,14] window factor analysis (WFA),[15] fixed size moving window evolving factor analysis (FSMWEFA),[16] heuristic evolving latent projections (HELP)[17,18] and evolving window orthogonal projection (EWOP).[19] The key feature of these methods mentioned is the use of informative 'window'. Such local-rank analysis methods have successfully solved many problems.[20] These methods contain one critical step that involves calculating the eigenvalues to determine the composition of the regions of evolutionary data. The number of components and time when they appear or disappear in a chemical system are very important parameters. By EFA and/or FSMWEFA techniques, the evolution of chemical contributions can be mathematically estimated without any prior assumption regarding the nature of these contributions or any assumption regarding a chemical model. The results are practically identical to those based on a specific chemical model and the classical least-squares modification. HELP requires suspected selective regions, as well as a zero-component region.[17,18] For complicated chemical reaction systems, such kind of zero-component regions do not exist, such that the application of HELP is somehow limited.

EFA often requires extensive computation time when m, that is, the number of time points in a process, is not too small because data matrix X would be split into many submatrixes, each of which, obtained as the window moves each time, would be larger and larger calculation. In addition, this method requires not only forward but also reverse calculations. Although FSMWEFA, applied in the analysis of complex mixture samples, saves lot of calculation time, it is not easy to accurately determine the starting and ending points of a component of interest during the process because of a time shift caused by the moving window. In the case of large data points, it is easy to find the one-component region (called selective region in HELP[17,18]) of some components, which contains very important information for multivariate resolution. However, for systems that contain a small number of data points or the components of interest only contain a few points in the one-component region, it would be difficult to accurately acquire the region and confirm the starting and ending points. This case very easily leads to the inaccuracy of resolution results.

Surface-enhanced Raman scattering (SERS) has become one of the most valuable tools in chemistry, materials science, and biology because of its fast, ultrahigh sensitivity and rich molecular structural information.[21–23] SERS activity is related to an appropriate interparticle distance of adjacent nanoparticles.[24,25] A new SERS detection called dynamic surface-enhancement Raman spectroscopy,[26,27] in which the interparticle distance of adjacent nanoparticles changes with the reaction, has been demonstrated to enhance the signals of analytes in comparison to the traditional methods by several orders of magnitude. Meanwhile, in actual measurements, the SERS signal is generally mixed with the fluorescence signal and other interference signals, for example noise, causing some difficulties in observing the changes in dynamic SERS. By monitoring the evaporation process and analysing the data, we can obtain the pure SERS signal and its concentration profile to achieve better observation of the system.

Low-molecular-weight thiols, such as cysteine (Cys) mentioned in this work, homo-cysteine (Hcy), *etc.*, play vital roles in physiological and pathological processes. The abnormal levels of these biothiols are related to different diseases, and the discrimination between them is of significant importance. 4-nitro-1,8-naphthalic anhydride (NNA) acting as a fluorescent probe for low-molecular-weight biothiols shows distinct changes when reacting with Cys. By studying the reaction process of NNA with Cys, we may obtain the knowledge of the reaction mechanism.

In this study, a method called fixed region scanning evolving factor analysis (FRSEFA) was developed. In this method, a base set of spectra that is a submatrix of the entire spectra data matrix X, which corresponds to a special region in the time direction, is selected. Then, the base set combines a moving window containing only one time point to construct a submatrix. A series of eigenvalues are then obtained to determine the starting or ending points. The combination of FRSEFA and alternating least squares (ALS) was successfully used in this study to analyze the surface-enhanced Raman scattering (SERS) data to determine evaporation process behaviors, and fluorescence data of fluorescent probe 4-nitro-1,8-naphthalimide (NNA) with cysteine (Cys) to confirm the kinetic reaction system.

2. Theory and algorithm

2.1 Evolving factor analysis (EFA) and fixed size moving window evolving factor analysis (FSMWEFA)

Evolving factor analysis (EFA), proposed by Marcel Maeder,[13] which was first developed for data processing of spectroscopic titration, has been successfully employed for the resolution of chromatography[28] and has been used in several other fields.[29,30] EFA and FSMWEFA, based on principal component analysis (PCA), are two techniques that have shown to be suitable for the study of evolutionary processes, such as those presented in chromatography and chemical reaction processes, for its deriving excellent use of the 'evolving' characteristic of the data and tracking changes in the rank according to the process of time.[31]

In time-based measurements, the spectral data matrix of $\boldsymbol{X}_{m\times n}$ contains n spectral responses of m samples at different

time points, $C_{m\times n_c}$ contains concentrations of n_c pure components, and $S_{n\times n_c}$ contains spectral responses of n_c pure components. The relationship among $X_{m\times n}$, $C_{m\times n_c}$ and $S_{n\times n_c}$ can be illustrated as:

$$X = C \times S^T + E \quad (1)$$

where E is the measured noise matrix, and T denotes the transpose of the matrix.

Several MCR methods rely heavily on the correct determination of factor number and the range of each factor that exists, which can be accomplished by EFA. The basic idea of EFA is to track the change in the rank according to the process of time by eigenvalue curves of the data matrix. A submatrix X_i, $i = 1, 2, 3,\ldots, m$, is extracted from the first to the ith row of X. The eigenvalues of $X_i^T \times X_i$, the covariance matrix of X_i, are calculated and plotted against the time point to obtain eigenvalue curves. When an eigenvalue is above the eigenvalue of noise at a point in time, it indicates that a new species has shown up. The backward EFA calculation is obtained similarly to the forward EFA calculation, but X_i is constructed from the last to the ith row of the raw data matrix X, where $i = m, m-1, m-2,\ldots, 1$. With the help of eigenvalue curves, the number and the areas of pure components can be determined. As mentioned above, EFA often requires extensive computation time when m is not very small.

Considering the drawbacks of EFA, FSMWEFA was developed based on EFA. In contrast to EFA, which works on increasing the size of submatrixes, FSMWEFA conducts a series of PCA runs to same size submatrixes by moving a fixed size window. The results provide a local rank map of the dataset, *i.e.* a representation of how many components are simultaneously present in different zones of the dataset.[7] The resolution ability of FSMWEFA depends on the correct choice of the size of the submatrix ("window") used to examine local regions.[18] A drawback of FSMWEFA is the existence of time shift because of the moving window, which can easily cause inaccurate determinations of starting and ending points of a component of interest, especially when few time points exist in the one-component region or the regions containing more than one component.

2.2 FRSEFA

FRSEFA was developed by considering the idea of the moving window. According to the basic idea of FSMWEFA, the strategies of the fixed size region and moving window are utilized, but the use of these two strategies is revised to construct the FRSEFA algorithm. In FRSEFA, a base set matrix X_{bs}, which is a submatrix of the entire spectra data matrix X, corresponding to a special region in the time direction, is selected. This base set can be a zero-component region, selective region (one-component region), or even more than one component region, but it should be of the same composition. Then, the base set matrix X_{bs} combines a moving window x_i, $i = 1, 2,\ldots, m$, containing only one time point to construct a submatrix $X_{t,i}$, $i = 1, 2,\ldots, m$.

The main difference between FRSEFA and FSMWEFA is that the moving window contains only one time point in FRSEFA, whereas it contains w time points (w should be at least the maximum composition expected in the system[32]) in FSMWEFA. Another difference is the construction of the submatrix. In FSMWEFA, the submatrix X_i is formed by extracting the ith to the $(i + w - 1)$th row of the spectra data matrix X; however, in FRSEFA, the submatrix $X_{t,i}$ is formed by combining the base set X_{bs} with the moving window x_i. Based on these differences, FRSEFA exhibits superiority, *i.e.* higher accuracy of the determination of the starting or ending points because the moving window contains only one time point; thus, the problem of time shift in FSMWEFA can be solved completely.

After the submatrix $X_{t,i}$ is successfully conducted, the eigenvalues of the submatrix $X_{t,i}$ are calculated. The first p eigenvalues of this submatrix recorded are related to time point i. As i increases from 1 to m, all the eigenvalues recorded together form an eigenvalue matrix Z_{eig}, whose dimension is $m \times p$. Z_{eig} is plotted against the time points to gain eigenvalue curves. When an eigenvalue is higher than the eigenvalue of noise at a time point, it indicates that a new component has shown up. The determination of component numbers at any time point is just the same as the classical EFA.

In fact, FRSEFA could supply many kinds of information according to the selection of X_{bs}. When a region of zero-component is selected as X_{bs}, with the help of the calculated eigenvalue matrix Z_{eig}, it is easy to obtain the time point at which a one-component region shows up or disappears. Similarly, when a selective region or a two-component region is selected as X_{bs}, we can obtain the time point at which a two-component or three-component region shows up or disappears. More-component regions can be determined as well, if necessary, according to FRSEFA.

2.3 FRSEFA-ALS

In MCR, alternating least squares (ALS) is a simple, effective and widely used data matrix decomposition method. ALS applies an iterative method on the model of eqn (1) (mentioned in Section 2.1) $X = C \times S^T + E$ to calculate C and S. The first step is to choose a suitable factor number that is the number of principal components, whose chemical significance is the number of chemistry components in the process. Then, the second step is to set the initial value of C or S. In this study, the first loading was taken as the initial value of S. FRSEFA can provide information regarding both the factor number and the initial value, which are required by ALS. The eigenvalue matrix X_{eig} in FRSEFA is used to determine the starting point or appearing time and the ending point or disappearing time of components. With the determined starting and ending points, a submatrix that corresponds to the selected region (zero, one, or two-component region) is easy to obtain and is carried out by ALS calculation.

Therefore, the combination of FRSEFA and ALS can confirm the accurate time of change during the monitoring of mixture reaction and further acquire the spectrum and concentration profiles of the components present in the mixture reaction

process. However, such an application is not limited to reaction monitoring. It can also be useful for several other applications in addition to reaction monitoring, chromatography or pH titration. FRSEFA combined with ALS can settle a system that meets the requirement of continuous change of component concentrations, no matter whether a zero-concentration region exists or not.

Data pretreatments, FRSEFA and alternating least squares in this study were carried out by homemade programs in MATLAB (Ver.7.1:The MATHWORKS, USA).

3. Datasets

3.1 Evaporation process monitoring with SERS spectra

Stock solution of Rhodamine 6G (R6G) at a concentration of 10^{-3} M was prepared with ultrapure water as the solvent. Subsequently, a sample solution at a concentration of 10^{-7} M was obtained by diluting the stock solution with the solvent.

Sample solution of 100 μL and gold colloid of 100 μL were uniformly mixed in a whirlpool oscillator. Then, 10 μL solution was removed from the mixture and was detected as soon as it dropped onto the quartz plate, which was carefully washed by ethanol and ultrapure water before use. SERS spectra were collected every minute as the solution slowly volatilized until the droplet completely dried.

All SERS spectra were collected by a portable Raman instrument (i-Raman, B&W Tek Inc., USA) attached with a microscope (20× objective). The laser excitation wavelength was 785 nm. All SERS spectra were recorded by focusing the laser on the surface of the droplet on the quartz plate with a total accumulation time of 20 s.

The spectra were collected in a matrix $\boldsymbol{X}$ that contained 75 time points in a row and 1838 points (193.97–3205.69 cm^{-1}) in a column. The original spectra at a concentration of 10^{-7} M as well as the blank spectrum that was measured using pure water as a sample are shown in Fig. 1.

3.2 Chemical kinetics process and its fluorescent detection

The reaction mechanism[33] between Cys and NNA involves the nucleophilic substitution of the nitro group of NNA by the mercapto group in Cys (see Fig. 2a). Then, 4-amino substitute (see Fig. 2b), in which intramolecularity brings two reaction partners into close proximity, is formed by further intramolecular aromatic substitution reaction of the alkylthio by the amino group, which occurs through a five-membered ring. In the second step, a hydrogen bond between sulfur and the carboxylic acid group is formed through a six-membered ring, which assists the departure of the alkylthio group from the naphthalene ring.

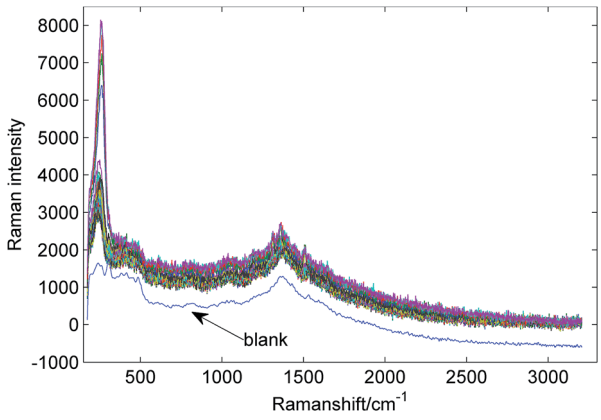

Fig. 1 Blank and original spectra of evaporation process.

An accurately weighed amount of 4-nitro-1,8-naphthalimide (NNA) was dissolved in dimethyl formamide (DMF) to obtain the stock solution of 2.0 mM, and 6.0 mg cysteine (Cys) was dissolved in 5 mL H_2O to obtain a stock solution of 10.0 mM. The DMF solvent was purified using a redistillation procedure. All other commercial reagents were of analytical grade and were used without further purification. The experiments were performed at 50 °C using nondegassed samples.

Sample solution was prepared by adding 9 μL and 300 μL stock solutions of Cys and NNA into 1.5 mL DMF solvent, respectively. All the samples were reacted at 50 °C for 2 hours.

All emission spectra were recorded by a fluorescence spectrometer (Thermo Fisher Scientific Lumina) equipped with a 150 W continuous xenon lamp and 1 cm quartz cells. The extraction wavelength was set as 395 nm. The spectra were collected in a matrix $\boldsymbol{X}$ of 30 time points in a row and 1451 wavelength points (445–590 nm) in a column. The original spectra are shown in Fig. 3.

4. Results and discussions

4.1 SERS spectra data of the evaporation process

As seen in the original SERS spectra in Fig. 1, the influence of noise is very serious. Before the resolution of the spectra, a Savitzky–Golay filter[34] was used for obtaining smooth curves and useful information and avoiding the interference of redundant information. The spectra after smoothing are shown in Fig. 4. From the figure, we can find that the spectra contain some Raman peaks, but very weak signals, mixed with other signals. It is hard to see the change in the signals by the naked eye.

4.1.1 Base set determination. FSMWEFA was first used to investigate the regional information, *e.g.*, zero-component region, one-component region with which the base region is easy to select. Eigenvalues calculated by FSMWEFA (with a window size of 6) are shown in Fig. 5, and the curves are observed to fluctuate over time. There is no zero-component region in the eigenvalue figure throughout the time. The first eigenvalue curve is always over the base level over the entire time region, which appears to indicate that one component exists during the entire process. It might be the background signal (see Section 4.1.4 resolution). It is clear that from the 50th point in Fig. 5 that the second eigenvalue curve becomes higher than the base level. From the 62nd point in Fig. 5 it is clear that the third eigenvalue curve becomes higher, which indicates that two new components show up at the 55th and 67th time points in the evaporation process, respectively.

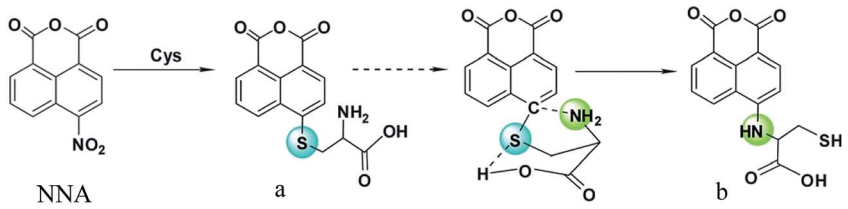

Fig. 2 Reaction mechanism of NNA with Cys.

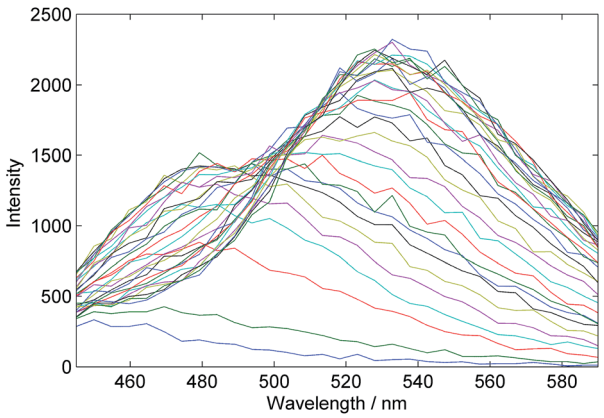

Fig. 3 Time-dependent emission spectra of NNA in the presence of Cys in DMF.

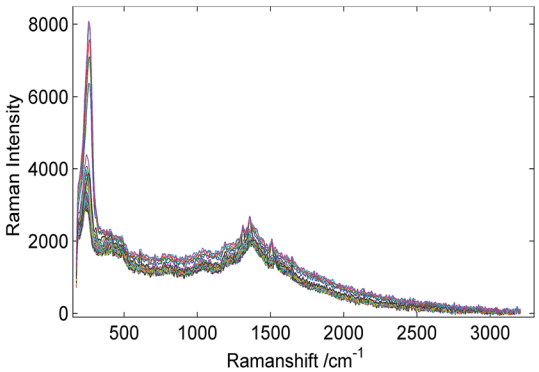

Fig. 4 Smoothing spectra of the evaporation process.

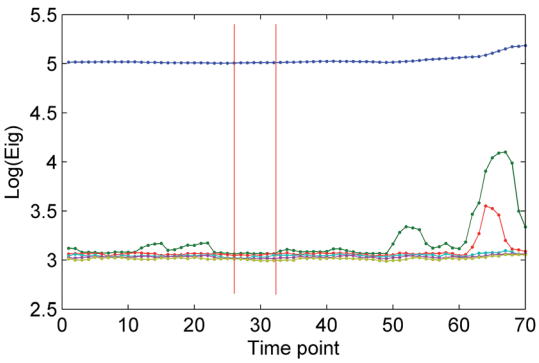

Fig. 5 Eigenvalues calculated by FSMWEFA (with a window size of 6).

4.1.2 FRSEFA based on one-component region selected. All the eigenvalues between the 4th–10th and 25th–32nd time points in Fig. 5 appear to be very smooth and close to the base level despite the highest eigenvalue curve (the first eigenvalue curve), and spectra in these two regions are very nearly the same. They are considered as one-component regions. There is no significant difference between selecting 9th–15th and 30th–37th time points of data (considering the time shift existing in FSMWEFA) as the base set. Thus, the 30th–37th points of the SERS spectra data that is the 30th–37th rows of the spectra matrix $\boldsymbol{X}$ were chosen as the base set for the sake of sensitivity. FRSEFA was applied to obtain the eigenvalue submatrix. The rank map of the first six eigenvalues is illustrated in Fig. 6.

From the rank map obtained by FRSEFA, the eigenvalue curve clearly changes from the 70th point to the end and a small pick appears during the 57th to 69th points. There are only six time points from the 70th point to the end. If FSMWEFA is used, it will be hard to accurately identify this area because of the time shift. Even if the area is identified, it must be very narrow, which makes it hard to carry out further multivariate resolution with ALS. It can be observed from Fig. 6 that one new component appears at the 57th point and clearly changes from the 70th time point, which will be very helpful for further distinguishing its spectrum and concentration.

4.1.3 FRSEFA based on two-component region selected. The 50th–60th time points in Fig. 5 are considered to be two-component regions. The rank map calculated by FRSEFA when considering the 56th–66th rows of the spectra matrix $\boldsymbol{X}$ as the base set is shown in Fig. 7. We can infer that there is no three-component region in the process according to Fig. 7.

4.1.4 Resolution. It can be confirmed that the spectra recorded before the 57th time point are one-component regions. The principal component analysis aims to map the

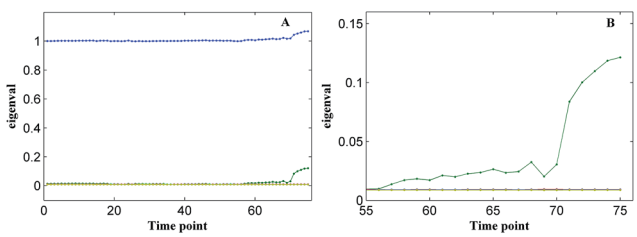

Fig. 6 First six eigenvalues obtained by FRSEFA based on one-component region selected ((A): the eigenvalue curves; (B): enlarged view of A).

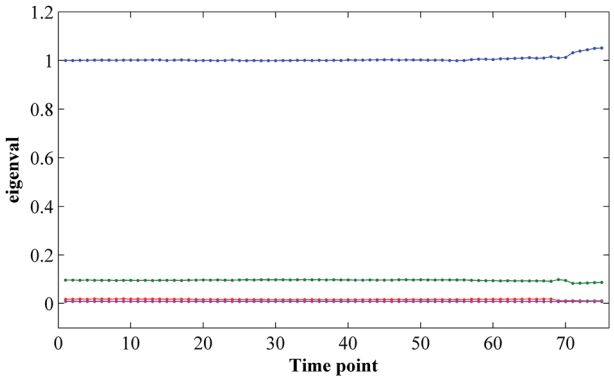

Fig. 7 First six eigenvalues obtained by FRSEFA based on the selected two-component region.

spectroscopy signals onto a low-dimension space with the largest variability. When PCA (singular value decomposition (SVD) adopted in this study) was applied to the spectra data from the 1st to the 56th points, the first loading, shown in Fig. 8, can represent the spectrum of the component within this time area. Its shape is similar to the detected blank spectrum with a correlation coefficient of 98.58%. The spectra collected before the 56th point are blank signals.

The PCA loading of the first principal component was considered to be the initial value into multivariate resolution iterative calculation. A non-negative limit was adopted in both the concentration and spectrum profile. This conditional limit would be used in every iteration to obtain new estimates of $\boldsymbol{S}$ and $\boldsymbol{C}$. Concentrations and spectral curves of physical significance were finally given by repeating iterations until convergence.

The resolution results of the entire process are shown in Fig. 9. Curve a is assigned to the SERS signal of R6G (inset in Fig. 9) and curve b is assigned to the background. SERS bands at 567 and 608 cm^{-1} are attributed to the C–C–C ring in plane bend. The band at 768 cm^{-1} corresponds to the motion of C–H out-of-plane stretching, and the SERS bands at 1123 and 1178 cm^{-1} are assigned to the vibrational mode of the C–H in plane. The bands at 1310, 1359, 1508 and 1650 cm^{-1} are attributed to

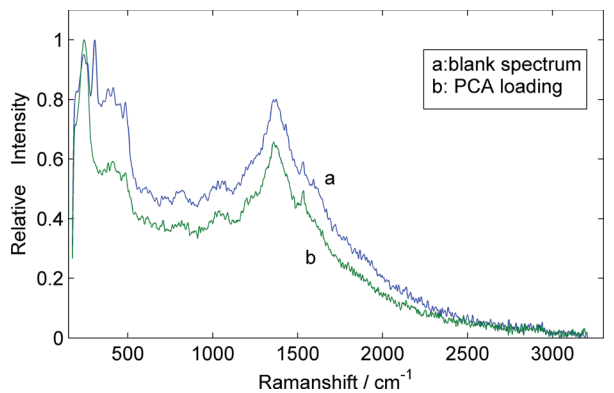

Fig. 8 Comparison of blank spectrum detected and the first loading of the 1st–56th points of the spectra data.

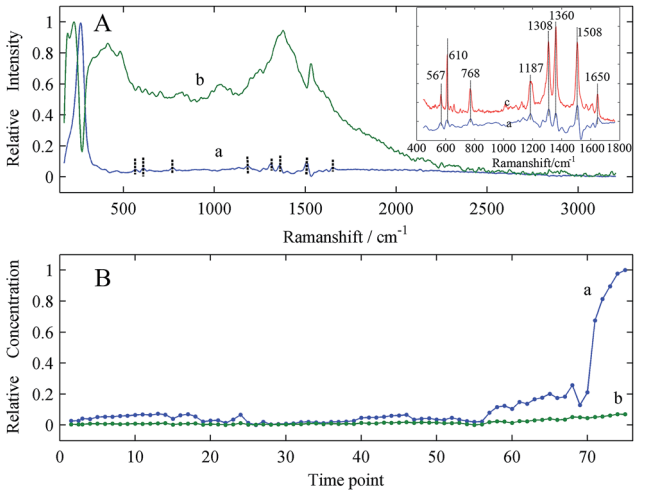

Fig. 9 Resolution results of the evaporation process ((A): spectrum profile; (B): concentration profile; Insert: comparison between the resolved spectrum and pure spectrum; a, b: two components resolved; c: normal SERS spectrum of R6G resolution at a concentration of 10^{-3} M).

the C=C stretching relating to the benzene ring.[35,36] From the resolution results, we find that in the course of the droplet drying, the SERS signal of R6G shows up at the 56th time point and gradually increases, then increases rapidly from the 70th point to the end of the reaction. The background signal remains almost stable throughout the entire evaporation process. The eigenvalue and the kinetic profile of evaporation reactions show extreme matching.

To observe this system in a better manner, the resolution results using FRSEFA and ALS are given in Fig. 10. The spectrum profile shown in Fig. 10A is determined to be a background signal, whose concentration profile remains substantially constant. The curve a in (C) and (E) is related to the SERS signal of the analyte R6G, but the shape of R6G SERS signal resolved is

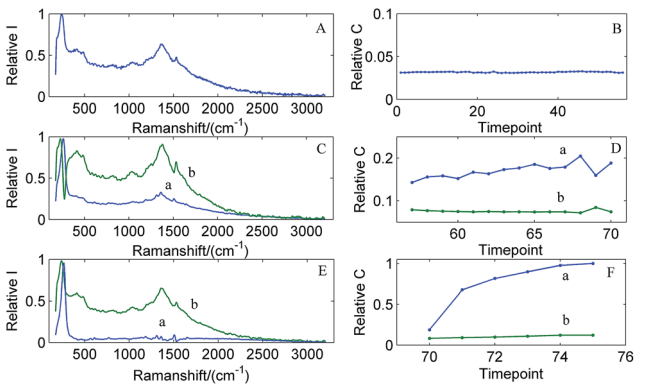

Fig. 10 Resolution results using FRSEFA-ALS ((A) and (B): spectrum and concentration profiles resolved between the 1st and 56th points of data; (C) and (D): spectrum and concentration profiles resolved between the 57th and 69th points of data; (E) and (F): spectrum and concentration profiles resolved between the 70th and 75th points of data; a, b: two components resolved).

clearer in (C) than in (E), and curve b in (C) and (E) is related to the background signal. The R6G SERS signal shows up at the 57th time point and slowly increases during the 57th–69th points, according to (D). For (F) in Fig. 10, we find that the R6G SERS signal increases rapidly from the 70th point till the end. A tentative explanation for this phenomenon could be that the interparticle distance of gold nanoparticles slowly reduced because of capillary force during solvent vaporization. During initial drying, which leads to droplet shrinking, the interparticle distance decreases which is not suitable to form "hot spots",[37,38] such that the R6G SERS signal cannot be observed. As the droplet evaporates continuously, some "hot spots" are generated by decreasing the interparticle distance of gold nanoparticles to a proper distance. Therefore, the R6G SERS signal appears.

Except for one component because of R6G, another resolved component is assigned to the blank sample. Their spectra (not given) are quite similar to each other with a correlation coefficient of 96.72%. The spectrum should be Raman signals contributed by some compounds in the blank sample such as the remaining citrate in gold colloid.[39]

4.2 Fluorescent spectra data of the chemical kinetics process

There are a total of 30 time points in the matrix data and only 6 points in the region of eigenvalue changes that would cause the multivariate resolution results to be even worse. Resolution results using FRSEFA and ALS are given in Fig. 11, showing the obtained spectrum and concentration profiles.

From the concentration profile corresponding to the resolved emission spectra, it can be confirmed that a center of emission peak in 479 nm is formed immediately after adding Cys to NNA solution. It reaches maximum intensity within 15 minutes. After that, fluorescence intensity decreases while a new emission peak center at 530 nm gradually forms and increases steadily. The maximum emission wavelength of two components are 479 and 530 nm, respectively, showing good agreement with the literature data.[33] In kinetic concentration profiles, we find that one of the factors increases regularly, while the other increases rapidly at the beginning of the reaction, and then decreases slowly with the reaction progress. Good results were obtained by this method.

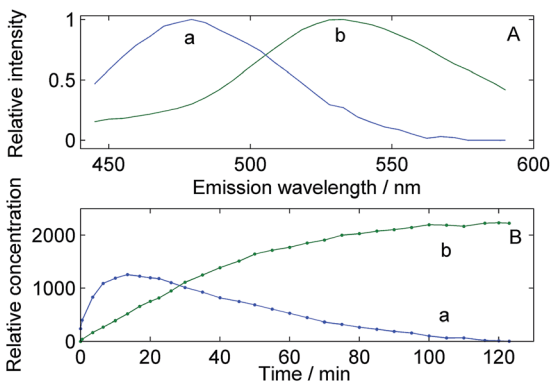

Fig. 11 Resolution results of the reaction of NNA with Cys((A): spectrum profile; (B): concentration profile; a, b: two components resolved).

5. Conclusion

Fixed region scanning evolving factor analysis (FRSEFA) algorithm was constructed in this paper to determine the precise time of the transformation of the components during a reaction. A series of submatrixes were constructed to calculate the local rank map by combining a moving window containing only one time point at a time with a base set that is a submatrix, corresponding to a special region, of the entire spectra matrix in the time direction. Therefore, the composition information at different positions of time can be determined more accurately. FRSEFA combined with ALS was applied to the monitoring of the evaporation process with the SERS technique and the chemical kinetic process of fluorescent probe NNA with Cys with fluorescence detection. Although severe background interference existed, the R6G SERS signal was still resolved and was consistent with the pure spectrum. The exact time points when the R6G SERS signal showed up and then increased were also given. For the chemical kinetic process, the emission spectra of two kinds of products in the reaction of NNA with Cys were well separated and their concentration changes were acquired. The results show that spectrum profiles and concentration profiles of the components in the mixture reaction are well estimated, providing the possibility for monitoring the multicomponent system reaction process and offering powerful evidence for explaining the reaction mechanism.

Acknowledgements

This work was supported by the National Science Foundation of China (grant no: 21205041).

References

1 O. González-García, C. Ariño, J. M. Díaz-Cruz and M. Esteban, *Chemom. Intell. Lab. Syst.*, 2008, **93**, 49–57.
2 M. D. G. García, F. C. Cañada, M. J. Culzoni, L. Vera-Candioti, G. G. Siano, H. C. Goicoechea and M. M. Galera, *J. Chromatogr. A*, 2009, **1216**, 5489–5496.
3 L. Dong and L. Xu, *Chin. J. Anal. Chem.*, 2004, **32**, 741–746.
4 R. Tauler and D. Barceló, *Trends Anal. Chem.*, 1993, **12**, 319–327.
5 R. Tauler, E. Casassas and A. Izquierdo-Ridorsa, *Anal. Chim. Acta*, 1991, **248**, 447–458.
6 J. Saurina, S. Hernández-Cassou and R. Tauler, *Anal. Chem.*, 1997, **69**, 2329–2336.
7 A. de Juan and R. Tauler, *Anal. Chim. Acta*, 2003, **500**, 195–210.
8 B. Li, J. Zhang, Y. Hu, Y. Liang and Y. Ozaki, *Appl. Spectrosc.*, 2006, **60**, 155–161.
9 Y. Ozaki, S. Sasic, J. H. Jiang and H. W. Siesler, *Macromol. Symp.*, 2002, **184**, 229–247.
10 A. Tanabe, S. Morita, M. Tanaka and Y. Ozaki, *Appl. Spectrosc.*, 2008, **62**, 46–50.

11 J.-H. Jiang, Y. Liang and Y. Ozaki, *Chemom. Intell. Lab. Syst.*, 2004, **71**, 1–12.
12 R. G. Brereton, in *Chemometrics*, John Wiley & Sons, Ltd, 2003, pp. 183–269, DOI: 10.1002/0470863242.ch4.
13 H. Gampp, M. Maeder, C. J. Meyer and A. D. Zuberbühler, *Talanta*, 1986, **33**, 943–951.
14 M. Maeder, *Anal. Chem.*, 1987, **59**, 527–530.
15 E. R. Malinowski, *J. Chemom.*, 1992, **6**, 29–40.
16 H. R. Keller and D. L. Massart, *Anal. Chim. Acta*, 1991, **246**, 379–390.
17 O. M. Kvalheim and Y. Z. Liang, *Anal. Chem.*, 1992, **64**, 936–946.
18 Y. Z. Liang, O. M. Kvalheim, H. R. Keller, D. L. Massart, P. Kiechle and F. Erni, *Anal. Chem.*, 1992, **64**, 946–953.
19 C.-J. Xu, J.-H. Jiang and Y.-Z. Liang, *Analyst*, 1999, **124**, 1471–1476.
20 H. A. Gad, S. H. El-Ahmady, M. I. Abou-Shoer and M. M. Al-Azizi, *Phytochem. Anal.*, 2013, **24**, 1–24.
21 X. M. Qian and S. M. Nie, *Chem. Soc. Rev.*, 2008, **37**, 912–920.
22 Y. Wang, B. Yan and L. Chen, *Chem. Rev.*, 2012, **113**, 1391–1428.
23 J. Zheng, Y. Hu, J. Bai, C. Ma, J. Li, Y. Li, M. Shi, W. Tan and R. Yang, *Anal. Chem.*, 2014, DOI: 10.1021/ac404004m.
24 H. Xu, E. J. Bjerneld, J. Aizpurua, P. Apell, L. Gunnarsson, S. Petronis, B. Kasemo, C. Larsson, F. Hook and M. Kall, *Proc. SPIE*, 2001, **4258**, 35.
25 L. Gunnarsson, E. J. Bjerneld, H. Xu, S. Petronis, B. Kasemo and M. Käll, *Appl. Phys. Lett.*, 2001, **78**, 802.
26 L. Yang, H. Liu, J. Wang, F. Zhou, Z. Tian and J. Liu, *Chem. Commun.*, 2011, **47**, 3583–3585.
27 K. Qian, L. Yang, Z. Li and J. Liu, *J. Raman Spectrosc.*, 2013, **44**, 21–28.
28 M. Maeder and A. D. Zuberbuehler, *Anal. Chim. Acta*, 1986, **181**, 287–291.
29 L. Nørgaard, *Anal. Chim. Acta*, 1991, **255**, 143–148.
30 M. G. Trevisan, C. M. Garcia, U. Schuchardt and R. J. Poppi, *Talanta*, 2008, **74**, 971–976.
31 J. Mendieta, M. S. Díaz-Cruz, M. Esteban and R. Tauler, *Biophys. J.*, 1998, **74**, 2876–2888.
32 R. G. Brereton, *Chemometrics*, John Wiley & Sons, Ltd, 2003, p. 378.
33 L. Ma, J. Qian, H. Tian, M. Lan and W. Zhang, *Analyst*, 2012, **137**, 5046–5050.
34 A. Savitzky and M. J. E. Golay, *Anal. Chem.*, 1964, **36**, 1627–1639.
35 V. S. Tiwari, T. Oleg, G. K. Darbha, W. Hardy, J. P. Singh and P. C. Ray, *Chem. Phys. Lett.*, 2007, **446**, 77–82.
36 K. Kneipp, Y. Wang, H. Kneipp, I. Itzkan, R. R. Dasari and M. S. Feld, *Phys. Rev. Lett.*, 1996, **76**, 2444–2447.
37 M. Futamata, *Faraday Discuss.*, 2006, **132**, 45–61.
38 H. Xu, J. Aizpurua, M. Käll and P. Apell, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2000, **62**, 4318–4324.
39 S. Sánchez-Cortés and J. V. García-Ramos, *J. Raman Spectrosc.*, 1998, **29**, 365–371.

