An improvement of window factor analysis for resolution of noisy HPLC-DAD data

SHAO Xueguang (邵学广), SHAO Limin (邵利民), LI Meiqing (李梅青) & LIN Xiangqin (林祥钦)

Department of Chemistry, University of Science and Technology of China, Hefei 230026, China
Correspondence should be addressed to Shao Xueguang (email: xshao@ustc.edu.cn)

Received January 17, 2002

Abstract Window factor analysis (WFA) is a powerful tool in analyzing evolutionary process. However, it was found that window factor analysis is much sensitive to the noise involved in original data matrix. An error analysis was done with the fact that the concentration profiles resolved by the conventional window factor analysis are easily distorted by the noise reserved by the abstract factor analysis (AFA), and a modified algorithm for window factor analysis was proposed. Both simulated and experimental HPLC-DAD data were investigated by the conventional and the improved methods. Results show that the improved method can yield less noise-distorted concentration profiles than the conventional method, and the ability for resolution of noisy data sets can be greatly enhanced.

Keywords: window factor analysis (WFA), error analysis, HPLC-DAD.

The development of modern instrumental analysis results in hyphenated techniques, such as GC-MS, GC-IR and HPLC-DAD, which were designed to reveal more properties of a chemical system. The data sets produced by these modern techniques are often given as two-way matrices, one way is a record of an evolutionary process, the other is a record of a certain property, such as UV, MS, and IR, of the species. The two-way matrix not only provides us with more abundant information, but also brings us a challenge of the analysis of the experimental data.

Window factor analysis (WFA)[1,2] is a self-modeling chemometric technique of multivariate statistical analysis, which is developed to extract concentration profiles of the chemical species from data matrix of an evolutionary process. Since its appearance as a chemometric method in 1980s, many applications of the technique have been reported[2—7] due to its advantage that no *a priori* information concerning the system is required. And by these applications, it was proven that window factor analysis is a powerful tool in analyzing evolutionary process. However, the theory of window factor analysis and its algorithm were developed without consideration of the involved noise in original data matrix. Results by window factor analysis were also found to be very easily affected by noise[8—10]. The signal-to-noise ratio (SNR) of the original data matrix determines the validity of the resolved results in practical use. Some methods[9,10] were designed to solve the problem. In these methods, extra smoothing procedures, such as smoothed principal component analysis (SPCA)[9] or wavelet transform[10], are generally employed to improve the SNR of the

data matrix.

This paper presents a study on the theory of window factor analysis with the noise being taken into consideration, and proposes an improved algorithm. In order to test its performance, three simulated and two experimental data sets of HPLC-DAD with different noise levels are prepared and investigated by the proposed method. It is found that, compared with the conventional window factor analysis, the method could greatly reduce the noise and consequently improve the quality of the results.

1 Theory

1.1 Theory of window factor analysis and conventional algorithm

The theory of the conventional window factor analysis can be summarized as follows[3]. Let $\boldsymbol{D}$ represent the measured data matrix of n-component chemical system, where each column is a spectrum digitally recorded during an evolutionary process, and each row is a record of evolutionary (concentration) profile. Assume that the spectral measurement is a linear sum of each component, then we have

$$\boldsymbol{D} = \sum_{i=1}^{n} \boldsymbol{D}_i = \sum_{i=1}^{n} \boldsymbol{s}_i \boldsymbol{c}_i' = \boldsymbol{SC}, \tag{1}$$

where $\boldsymbol{D}_i$ is a matrix representing the contribution of the ith component to $\boldsymbol{D}$. Vectors $\boldsymbol{c}_i$ and $\boldsymbol{s}_i$ are the spectrum and the concentration profile of the ith component.

Specify a region along the evolutionary axis which exactly fits the concentration profile of the nth component. This region is called the "window" of the nth component, though the concentration profiles of other components may exist inside the window. Let $\boldsymbol{D}^0$ represent a submatrix of $\boldsymbol{D}$ obtained by removing all columns within the window, and perform abstract factor analysis to it, then we have

$$\boldsymbol{D}^0 = \sum_{j=1}^{n-1} \boldsymbol{s}_j^0 \boldsymbol{c}_j^{0\prime} = \boldsymbol{S}^0 \boldsymbol{C}^0, \tag{2}$$

where matrix $\boldsymbol{S}^0$ contains $n-1$ orthonormal spectral vectors $\boldsymbol{s}_j^0$, matrix $\boldsymbol{C}^0$ contains $n-1$ orthogonal non-normalized concentration profile vectors $\boldsymbol{c}_j^0$.

Because the true spectral vectors of $n-1$ components $\boldsymbol{s}_i$ and the abstract spectral vectors $\boldsymbol{s}_j^0$ both lie in an $(n-1)$-dimensional subspace of the overall n-dimensional factor space, $\boldsymbol{s}_i$ can be linearly expressed by $\boldsymbol{s}_j^0$, i.e.,

$$\boldsymbol{s}_i = \sum_{j=1}^{n-1} \beta_{ij} \boldsymbol{s}_j^0. \tag{3}$$

By adding a vector $\boldsymbol{s}_n^0$ which is orthonormal to $\boldsymbol{s}_j^0$, the true spectral vector of the nth component

can be similarly obtained by

$$\boldsymbol{s}_n = \sum_{j=1}^{n-1} \beta_{nj} \boldsymbol{s}_j^0 + \beta_{nn} \boldsymbol{s}_n^0, \tag{4}$$

where $\sum_{j=1}^{n-1} \beta_{nj} \boldsymbol{s}_j^0$ represents the projection onto the hyperplane. β in (3) and (4) is a linear coefficient.

Based on the orthogonality of the abstract spectral vectors $\boldsymbol{s}_j^0$, an expression of window factor analysis can be derived from (1), (3) and (4):

$$\beta_{nn} \boldsymbol{s}_n^0 \boldsymbol{c}_n' = \boldsymbol{D} - \boldsymbol{S}^0 \boldsymbol{S}^{0\prime} \boldsymbol{D} = (\boldsymbol{I} - \boldsymbol{S}^0 \boldsymbol{S}^{0\prime}) \boldsymbol{D} = \boldsymbol{X}_n, \tag{5}$$

where $\boldsymbol{I}$ is the identity matrix.

Eq. (5) shows that $\boldsymbol{X}_n$ is the information of the nth component, and calculable for $\boldsymbol{S}^0$ and $\boldsymbol{D}$ being known. Because rows of $\boldsymbol{X}_n$ are proportional to each other, the average of the row vectors is the uncalibrated concentration profile of the nth component.

1.2 Error analysis of window factor analysis

Obviously, the measurement noise was not taken into consideration in the above theory of the conventional window factor analysis. As a result, the uncalibrated concentration profile obtained by eq. (5) is always distorted by the noise. In some cases, the distortion is so serious that the result of the conventional window factor analysis is meaningless. Therefore, further studies on the error analysis of the method are necessary.

When noise is involved in $\boldsymbol{D}$, the rank of $\boldsymbol{D}^0$ is always greater than $n-1$. According to the previous researches[3,7], some noise can be deleted by abstract factor analysis based on principal component analysis (PCA). But from the theory of PCA, it can be derived that the noise whose standard deviation or variance cannot be neglected will be reserved in the abstract spectral vectors $\boldsymbol{S}^0$. Therefore, for eqs. (3) and (4), the true spectral vectors should be expressed as

$$\boldsymbol{s}_i = \sum_{j=1}^{n-1} \beta_{ij} \boldsymbol{s}_j^0 - \boldsymbol{e}_i, \tag{6}$$

$$\boldsymbol{s}_n = \sum_{j=1}^{n-1} \beta_{nj} \boldsymbol{s}_j^0 + \beta_{nn} \boldsymbol{s}_n^0 - \boldsymbol{e}_n, \tag{7}$$

where $\boldsymbol{e}_i$ and $\boldsymbol{e}_n$ are error vectors corresponding to the reserved noise. Similarly, the added vector $\boldsymbol{e}_n$ is orthogonal to the $(n-1)$-dimensional subspace.

Inserting (6) and (7) into (1) gives

$$\boldsymbol{D} = \sum_{i=1}^{n} \boldsymbol{s}_i \boldsymbol{c}_i' = \sum_{j=1}^{n-1} \boldsymbol{s}_j^0 (\sum_{i=1}^{n} \beta_{ij} \boldsymbol{c}_i') + \beta_{nn} \boldsymbol{s}_n^0 \boldsymbol{c}_n' - \sum_{i=1}^{n} \boldsymbol{e}_i \boldsymbol{c}_i'. \tag{8}$$

Multiplying both sides of eq. (8) by $\boldsymbol{s}_j^{0\prime}$ and recalling that the abstract spectral vectors are mutu-

ally orthonormal lead to

$$\boldsymbol{s}_j^{0\prime}\boldsymbol{D} = \sum_{i=1}^{n}\beta_{ij}\boldsymbol{c}_i' - \boldsymbol{s}_j^{0\prime}\sum_{i=1}^{n}\boldsymbol{e}_i\boldsymbol{c}_i'. \tag{9}$$

Inserting (9) into (8) gives

$$\boldsymbol{D} = \sum_{j=1}^{n-1}\boldsymbol{s}_j^0\boldsymbol{s}_j^{0\prime}\boldsymbol{D} + \sum_{j=1}^{n-1}\boldsymbol{s}_j^0\boldsymbol{s}_j^{0\prime}\left(\sum_{i=1}^{n}\boldsymbol{e}_i\boldsymbol{c}_i'\right) + \beta_{nn}\boldsymbol{s}_n^0\boldsymbol{c}_n' - \sum_{i=1}^{n}\boldsymbol{e}_i\boldsymbol{c}_i'. \tag{10}$$

Eq. (10) can be rearranged as

$$\beta_{nn}\boldsymbol{s}_n^0\boldsymbol{c}_n' + (\boldsymbol{S}^0\boldsymbol{S}^{0\prime} - \boldsymbol{I})\boldsymbol{E}\boldsymbol{C} = (\boldsymbol{I} - \boldsymbol{S}^0\boldsymbol{S}^{0\prime})\boldsymbol{D}, \tag{11}$$

where $\boldsymbol{E}$ is the matrix of the error vectors.

A comparison between (5) and (11) shows clearly that term $(\boldsymbol{I} - \boldsymbol{S}^0\boldsymbol{S}^{0\prime})\boldsymbol{D}$, i.e. $\boldsymbol{X}_n$ in (5), contains two parts. One is the information of the nth component, the other is the contribution of the noise reserved in abstract spectral vectors. Therefore, direct computation from $(\boldsymbol{I} - \boldsymbol{S}^0\boldsymbol{S}^{0\prime})\boldsymbol{D}$, as in the conventional window factor analysis, will have the results noise-distorted.

1.3 A new algorithm of window factor analysis

In order to derive the formula for an improved algorithm of window factor analysis, we can multiply eq. (11) by $\boldsymbol{D}'$. Then we get

$$\boldsymbol{D}'\beta_{nn}\boldsymbol{s}_n^0\boldsymbol{c}_n' + \boldsymbol{D}'(\boldsymbol{S}^0\boldsymbol{S}^{0\prime} - \boldsymbol{I})\boldsymbol{E}\boldsymbol{C} = \boldsymbol{D}'(\boldsymbol{I} - \boldsymbol{S}^0\boldsymbol{S}^{0\prime})\boldsymbol{D}. \tag{12}$$

For the first item on the left hand of eq. (12), with eq. (7) we get

$$\boldsymbol{D}'\beta_{nn}\boldsymbol{s}_n^0\boldsymbol{c}_n' = \left(\sum_{i=1}^{n-1}\boldsymbol{c}_i\sum_{j=1}^{n-1}(\beta_{ij}\boldsymbol{s}_j^{0\prime} - \boldsymbol{e}_i') + \beta_{nn}\boldsymbol{c}_n\boldsymbol{s}_n^{0\prime} - \boldsymbol{c}_n\boldsymbol{e}_n'\right)\beta_{nn}\boldsymbol{s}_n^0\boldsymbol{c}_n'. \tag{13}$$

Based on the fact that $\boldsymbol{s}_n^0$ is orthogonal to the $(n-1)$-dimensional subspace, eq. (13) can be simplified to

$$\boldsymbol{D}'\beta_{nn}\boldsymbol{s}_n^0\boldsymbol{c}_n' = \boldsymbol{c}_n(\beta_{nn}\boldsymbol{s}_n^{0\prime} - \boldsymbol{e}_n')\beta_{nn}\boldsymbol{s}_n^0\boldsymbol{c}_n'. \tag{14}$$

For the second item on the left hand of eq. (12), if we express $\boldsymbol{D}'$ by $\sum_{i=1}^{n}\boldsymbol{c}_i\boldsymbol{s}_i'$, we can get

$$\begin{aligned}\boldsymbol{D}'(\boldsymbol{S}^0\boldsymbol{S}^{0\prime} - \boldsymbol{I})\boldsymbol{E}\boldsymbol{C} &= \left(\sum_{i=1}^{n}\boldsymbol{c}_i\boldsymbol{s}_i'\right)(\boldsymbol{S}^0\boldsymbol{S}^{0\prime} - \boldsymbol{I})\boldsymbol{E}\boldsymbol{C}\\ &= \left(\left(\sum_{i=1}^{n-1}\boldsymbol{c}_i\boldsymbol{s}_i'\right) + \boldsymbol{c}_n\boldsymbol{s}_n'\right)(\boldsymbol{S}^0\boldsymbol{S}^{0\prime} - \boldsymbol{I})\boldsymbol{E}\boldsymbol{C}\\ &= \boldsymbol{D}^{0\prime}\left(\boldsymbol{S}^0\boldsymbol{S}^{0\prime} - \boldsymbol{I}\right)\boldsymbol{E}\boldsymbol{C} + \boldsymbol{c}_n\boldsymbol{s}_n'\left(\boldsymbol{S}^0\boldsymbol{S}^{0\prime} - \boldsymbol{I}\right)\boldsymbol{E}\boldsymbol{C}.\end{aligned} \tag{15}$$

It is obvious that $\boldsymbol{D}^{0\prime}(\boldsymbol{S}^0\boldsymbol{S}^{0\prime} - \boldsymbol{I})$ is the difference between $\boldsymbol{D}^{0\prime}$ and its abstract matrix.

Therefore, when correct factor number is used in the calculation, it should be small enough to be negligible.

Based on the fact that both s_n^0 and e_n are orthogonal to the (n−1)-dimensional subspace, the item $c_n s_n'(S^0 S^{0\prime} - I)EC$ in eq. (15) can be simplified to

$$c_n s_n'(S^0 S^{0\prime} - I)EC = c_n \left(e_n' e_n - \beta_{nn} s_n^{0\prime} e_n\right) c_n'. \tag{16}$$

Therefore, eq. (15) becomes

$$D'(S^0 S^{0\prime} - I)EC = c_n \left(e_n' e_n - \beta_{nn} s_n^{0\prime} e_n\right) c_n'. \tag{17}$$

Inserting eqs. (14) and (17) into (12) gives

$$k c_n c_n' = D'(I - S^0 S^{0\prime}) D = Y_n, \tag{18}$$

where Y_n represents the extracted information of the nth component, k is a constant and equals $\beta_{nn}\beta_{nn} s_n^{0\prime} s_n^0 - \beta_{nn} e_n' s_n^0 + e_n' e_n - \beta_{nn} s_n^{0\prime} e_n$, i.e.,

$$k = (\beta_{nn} s_n^0 - e_n)'(\beta_{nn} s_n^0 - e_n). \tag{19}$$

Eq. (18) is an improved algorithm for window factor analysis. From eq. (19), it is clear that e_n will no longer have contribution to Y_n. Therefore, the uncalibrated concentration profile of the nth component can be easily obtained from matrix Y_n by averaging the row or the column vectors without the interference of the noise.

After all the concentration profiles being calculated, the spectra of all the components can be obtained by least square:

$$S = DC'(CC')^{-1}. \tag{20}$$

2 Experimental

2.1 Data simulation

A four-component HPLC-DAD data matrix was simulated. The spectra used in simulation are shown in fig. 1, and the concentration profiles are generated using Gaussian equation as follows:

$$y = h \exp\left(-4\ln(2)\left(\frac{t - t_0}{W_{1/2}}\right)^2\right). \tag{21}$$

The parameters for the simulation of the four peaks are 6.0, 4.0, 6.0, and 5.0 for h; 4.0, 5.2, 6.2, and 7.5 for t_0, respectively, and all the four values for $W_{1/2}$ are 1.0. Three data sets were prepared with different noise levels of SNR = 50, 20 and 10.

2.2 Chemicals

The stock solutions of the rare earth elements were prepared by dissolving their oxides (99.95%) in HCl to give 1.000 mg • mL^{-1} metal solution in 1.0 mol • L^{-1} HCl. The sample

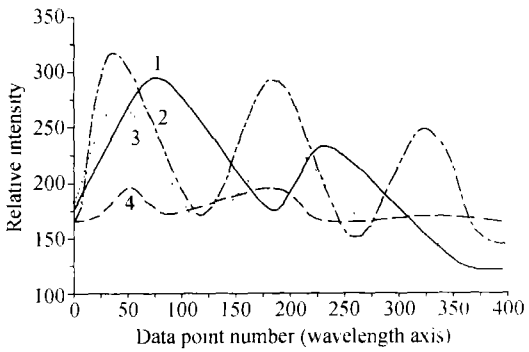

Fig. 1. Spectra used in the simulation of HPLC-DAD data sets. 1, Component 1; 2, component 4; 3, component 3, 4, component 2.

solution was mixed by the stock solution of Yb and Tm, or Yb, Tm and Er. The pH of the sample solution was adjusted to 3.5 with ammonium hydroxide (A.R.). The hydrophobic ion reagent used for pretreating the reversed-phase column was 0.01 mol • L^{-1} 1-dodecanesulphonate. Two mobile phase solutions were prepared containing 0.25 mol • L^{-1} lactic acid and pH = 2.5 (a) and 4.5 (b), respectively. The concentration of post-column reaction reagent of arsenazo III (Fluka Chemie, Switzerland) was prepared 1.0×10^{-4} mol • L^{-1} with redistilled water. All solutions were filtered through a 0.25 μm membrane filter.

A 2-component and a 3-component samples were prepared for measurement. Their constitution and concentration are listed in table 1.

Table 1 Constitution and concentration of the two samples

Sample	Yb	Tm	Er
No.1	501	2.599	—
No.2	20.01	19.99	20.00

Unit: μg • mL^{-1}.

2.3 Equipment and data acquisition

An HPLC system comprising Spectrasystem FL2000 (Spectra-Physics, USA) with the Spectra Focus multi-wavelength UV-Vis detector (Spectra-Physics, USA) and the Spectrasystem workstation was used for the separation. The column was packed with ODS silica (10 μm, 250 mm × 5 mm, Shimadzu, Japan) and the post-column reaction agent was delivered by an LC-6A pumps (Shimadzu, Japan). The experiment was accomplished under the following conditions: the total flow rate was 1.0 mL • min^{-1}; the ratio of the two mobile phase solutions (a) : (b) was 3 : 2 for sample No.1, and 1 : 1 for sample No. 2; the temperature was 20℃; and the flow rate of the post-column reaction agent was 1.0 mL • min^{-1}.

Data points covering the wavelength from 580 to 720 nm, digitized every 5 nm, and the chromatogram between 0 and 12 min, sampled approximately every 0.005 min, were recorded. The size of the data matrix used in calculation is 717 × 29 (from 4.5 to 8.6 min) for the sample No.1, and 932 × 29 (from 4.5 to 9.9 min) for the sample No.2.

3 Results and discussion

3.1 Analysis of the simulated HPLC-DAD data sets

In order to compare the performance of the conventional and improved window factor analy-

sis algorithm, the three simulated data sets with SNR = 50, 20, and 10 were investigated by the two methods respectively. Factor number was chosen 4 based on the principal factor analysis[6], which is in agreement with the factual number of components used in the simulation. The resolved concentration profiles by both window factor analysis methods are shown in figs. 2—4 with different types of lines.

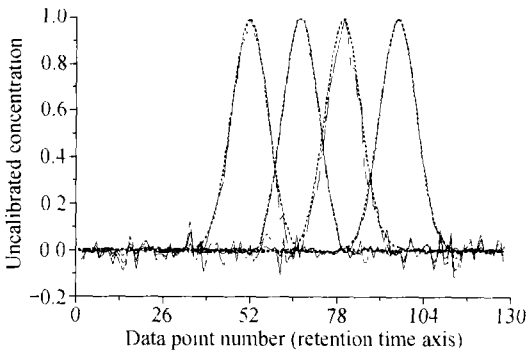

Fig. 2. The resolved concentration profiles by the conventional (solid curves) and improved (dash curves) methods from the simulated data set of SNR = 50.

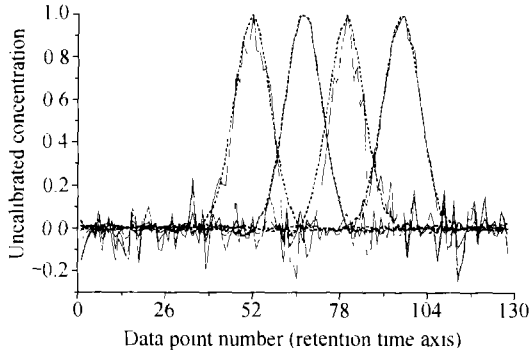

Fig. 3. The resolved concentration profiles by the conventional (solid curves) and improved (dash curves) methods from the simulated data set of SNR = 20.

From the results of the conventional window factor analysis (solid curves in figs. 2—4), it can be found that the concentration profiles are all noise involved. In fig. 2, the noise in each profile is clearly visible even when the SNR is as high as 50. From figs. 2—4, it can be found that the lower the SNR is, the more seriously the results are affected, especially for the concentration profiles of components 1 and 3. When the SNR decreases to 10, the distortion is so serious for components 1 and 3 that it is difficult to obtain reasonable results for further analysis.

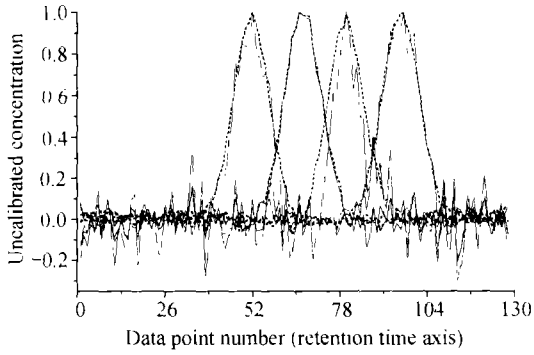

Fig. 4. The resolved concentration profiles by the conventional (solid curves) and improved (dash curves) methods from the simulated data set of SNR = 10.

The dash curves in figs. 2—4 are the resolved results of the same data sets by the improved window factor analysis method. In the case of the SNR being 50, it can be found that all the four concentration profiles are almost in perfect Gaussian shape and smooth enough. With the decrease of SNR, the noise in baseline and the distortion of the resolved peaks also increase as that in conventional method, but the level of noise and the degree of the distortion are very small compared with that in the solid curves. When the SNR is as low as 10, all the resolved concentration profiles are only slightly affected.

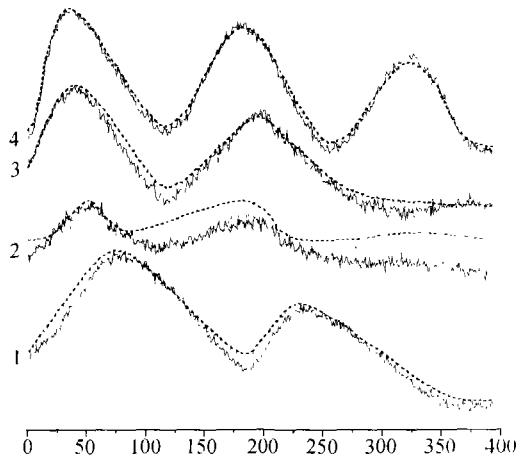

Fig. 5. Comparison between the resolved spectra by the conventional (solid curves) and improved (dot curves) methods and the simulated spectra (dash curves). The data set is the simulated one with SNR = 10.

For further comparison of the conventional and improved method, the resolved spectra for the simulated data matrix with SNR = 10 are shown in fig. 5. It can be found that for the spectra of components 1, 3 and 4, both methods yield satisfactory results. It can also be found that the results of this new method are in better coincidence with the original spectra than those of the conventional method. But for the spectrum of component 2, there is an obvious difference between the calculated spectrum and the original one. That is because the peak of component 2 is the weakest one of all the four peaks in the simulation, and the interference of the noise is the most serious. But the new method gives a better result.

Therefore, the difference between the two window factor analysis methods can be clearly seen by the resolved results of the simulated data sets. Due to the limited ability to eliminate noise, the conventional window factor analysis cannot obtain satisfactory results when the SNR of original data matrix is low. On the contrary, the improved algorithm can exclude the effect of noise and successfully yield satisfactory results from noisy data sets.

3.2 Analysis of the experimental HPLC-DAD data sets

The two experimental HPLC-DAD data sets of the samples in table 1 are analyzed to compare the performance of the two window factor analysis methods for experimental noisy data.

The solid curves in fig. 6 are the best results of many trials by the conventional window factor analysis based on factor number being 3. It can be seen that, for Tm, the component of comparatively high concentration, a reasonable concentration profile was obtained. However, for Yb, the component of low concentration, not only noise is involved in the result, there is also a serious distortion in its resolved profile due to the high level of noise in the original data matrix. Theoretically, the main reason for the distortion should be that the factor number was set to 3, which may result in the loss of some useful information in the abstract spectral vectors. In order to retrieve the lost information, we tried to set the factor number to be 4 or 5. Unfortunately, the noise level in the resolved profiles is so high that it is difficult to obtain any reasonable result.

The dash curves in fig. 6 are the results by the improved window factor analysis method from data set of sample No.1 with factor number being 4. It can be found that, contrary to the conventional method, the new method yields a reasonable profile of Yb. The slight distortion is due to relatively low concentration of Yb. As for the result of Tm, the new method yields a much better

concentration profile than that by the conventional method. Furthermore, we also found that similar result can be obtained with the factor number being 5 or 6.

In order to investigate the performance of the two methods for resolving more complex data set, the data set of sample No.2, which is composed of three components, was also analyzed by the two algorithms of window factor analysis. Unfortunately, we cannot obtain any meaningful result by conventional method. But satisfactory results are obtained by the improved algorithm, which are shown in fig. 7. Although the noise level is still comparatively high, the outline of each resolved profile coincides well with the peak shape of the standard. The tailing of the peaks is caused by the HPLC column and experimental conditions.

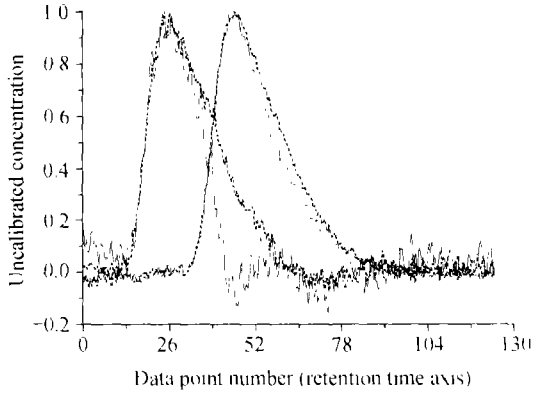

Fig. 6 The resolved concentration profiles by the conventional (solid curves) and improved (dash curves) methods from the two components experimental HPLC DAD data (data set of sample No.1).

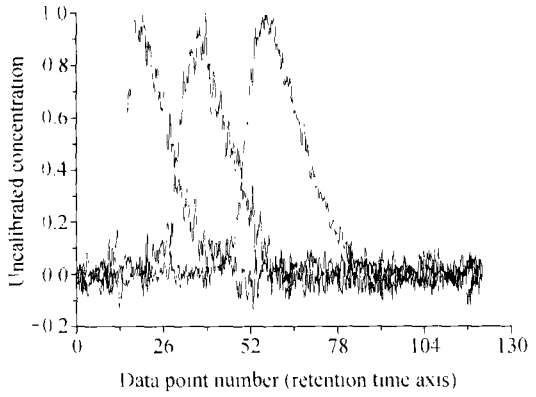

Fig. 7 The resolved concentration profiles by the improved method from the three components experimental HPLC-DAD data (data set of sample No.2).

Spectra of the experimental HPLC-DAD data were calculated. Spectra of pure elements were prepared to serve comparison. Similar to the cases of simulated data, the new method can also produce more coincident spectra than the conventional one.

4 Conclusion

An improved algorithm of window factor analysis was proposed. Theoretical analysis proves that the improved method can prevent the effect of noise in original data matrix from distortion of the resolved results. By investigation of both simulated and experimental data sets, it was proven that, for resolution of data matrix with high level of noise, the improved window factor analysis is superior to the conventional method. The improved method can eliminate the effect of noise and retrieve more useful information from noisy data matrix.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 29975027).

References

1 Maeder, M., Evolving factor analysis for the resolution of overlapping chromatographic peaks, Anal. Chem., 1987, 59(3):

527—530.

2. Malinowski, E. R., Window factor analysis: theoretical derivation and application to flow injection analysis data, J. Chemometrics, 1992, 6(1): 29—40.
3. Den, W., Malinowski, E. R., Investigation of copper(II) ethylenediaminetetraacetate complexation by window factor analysis of ultraviolet spectra, J. Chemometrics, 1993, 7(2): 89—98.
4. Schostack, K. J., Malinowski, E. R., Investigation of window factor analysis and matrix regression analysis in chromatography, Chemometrics Intell. Lab Syst., 1993, 20(2): 173—182
5. Gemperline, P. J., Hamilton, J C., Conditions for detecting overlapped peaks with principal component analysis in hyphenated chromatographic methods, Anal Chem., 1989, 61(20): 2240—2243.
6. Gemperline, P. J., Mixture analysis using factor analysis, I. Calibration and quantitation, J. Chemometrics, 1989, 3(4): 549—568.
7 Malinowski, E. R., Factor Analysis in Chemistry, 2nd ed , New York: Wiley, 1991, Chapter 2.
8 Shao, X. G., Cai, W. S., Resolution of multicomponent chromatograms by window factor analysis with wavelet transform preprocessing, J Chemometrics, 1998, 12(2): 85—93.
9 Chen, Z. P., Jiang J. H., Li, Y. et al., Smoothed window factor analysis, Anal. Chim. Acta, 1999, 381: 233—246.
10. Chen, Z. H., Lin, X Q., Shao, X. G., Smoothed principal component analysis based on the wavelet transform, Chin. J. Anal. Chem., 2000, 28(8): 960—963.