# Trilinear decomposition method applied to removal of three-dimensional background drift in comprehensive two-dimensional separation data

Yan Zhang [a], Hai-Long Wu [a,*], A.-Lin Xia [a], Liang-Hai Hu [b], Han-Fa Zou [b], Ru-Qin Yu [a]

[a] *State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, China*
[b] *National Chromatographic R&A Center, Dalian Institute of Chemical Physics, The Chinese Academy of Sciences, Dalian 116023, China*

## Abstract

A novel technique for removal of three-dimensional background drift in comprehensive two-dimensional (2D) liquid chromatography coupled with diode array detection (LC × LC-DAD) data is proposed. The basic idea is to perform trilinear decomposition on the instrumental response data, which is based on the alternating trilinear decomposition (ATLD) algorithm. In model construction, the background drift is modeled as one component or factor as well as the analytes of interest, hence, the drift is explicitly included into the calibration. The method involves performing trilinear decomposition on the raw data, then extracting the background component and subtracting this background data from the raw data, leaving the analytes' signal on a flat baseline. Simultaneous evaluation of three-dimensional background drift and true signals may improve the quality of the data. This method is applied to the determination and removal of three-dimensional background drifts in simulated multidimensional data as well as experimental comprehensive two-dimensional liquid chromatographic data. It is shown that this technique yield a good removal of background drift, without the need to perform a blank chromatographic run, and required no prior knowledge about the sample composition.
© 2007 Elsevier B.V. All rights reserved.

*Keywords:* Removal; Three-dimensional background drift; LC × LC-DAD; ATLD; Three-way data analysis

## 1. Introduction

Comprehensive two-dimensional (2D) separation instruments coupled with multichannel spectroscopic detectors is ideally suited for the complex sample analysis due to its enhanced peak capacity and selective detection, and it is emerging as powerful tools for chemical analysis. Example of such instruments include LC × LC [1,2], LC × GC [3,4], GC × GC [5,6], LC × CE [7,8], CE × CE [9,10] coupled with multichannel detectors. Background drift is often a problem in this chromatography analysis. The origin of drift is various, e.g. instrumental background noise, gradient liquid chromatography with changing mobile phase composition in function of time, and fluctuations of pressure and temperature can all cause drift. Since the quantitative data evaluation is based on the assumption that the baseline is stable in time. If this prerequisite is violated, major concentration errors can result as drift is evaluated along with true spectroscopic features.

There are two normally used methods to solve the background noise. One is the mean centering. It is necessary for such data pretreatment method that the background is stable. However, problems arise if the background noise changes with time as mean centering is no longer subtracting it completely, a remaining offset would result in errors on the analysis. The other one is to remove the background by subtraction of a blank chromatographic run. But simple subtraction of the eluent response does not always give good results. There are two main reasons for this. The first is the change in response intensity of the eluent spectrum during a chromatography run. The second is that small spectral shape changes of the eluent spectrum might occur.

For this reason, one should therefore use redesign of experiments, pre-treatment of spectral data and some calibration routines to compensate for signal drift. Alternatively, chemometric method is a good choice to determine and correct for drift in instrumental response. With each analysis, 2D separation instruments coupled with multichannel detectors produce a trilinear three-way array of data (3D), composed of the two separation dimensions and the detector channel dimension [11–13]. Background drift in this 3D data is difficult to determine and

---

* Corresponding author. Tel.: +86 731 8821818; fax: +86 731 8821818.
*E-mail address:* hlwu@hnu.cn (H.-L. Wu).

remove. The chemometric method discussed in this paper is to mathematically remove this background drift. The basic idea of this method is to model the drift as one factor in performing the trilinear decomposition on the raw chromatographic data. By subtracting the individual signal representing the background drift from the original raw chromatographic data, the drift along all the three measure dimensions is removed. In this work, the alternating trilinear decomposition (ATLD) [14] algorithm is discussed in detail to remove the background drift in 3D chromatographic data, it is shown that this technique can yield good removal of background drift.

Several multivariate approaches to handing background drift in two-dimensional data are known. Moore et al. discovered a method of removing the low frequency baseline drift in chromatographic data [15], the method involves digitally filtering the chromatographic data using a moving median filter. Andrew et al. [16] applied the Kalman filter to determine drift and correct drift in diode array spectral data sets. Vogt et al. proposed a so-called "pseudo principal components" method for correction of background drifts in optical spectra [17]. Furusjö et al. applied the iterative target testing procedure to yield good rate constant estimated for IR data in the presence of solvent absorption shifts and spectrometer drift [18]. Boelens et al. proposed a method to eliminate the background spectrum (EBS) [19] during analyte elution in column liquid chromatography coupled to spectroscopic techniques. To the best of our knowledge, however, no effort has been devoted to resolve background drift in 3D data, such as those produced by 2D separation analysis coupled with multichannel detectors.

The technique described herein is demonstrated to remove three-dimensional background drift in LC × LC-DAD data, and this technique is applicable to data from all 2D separation instruments coupled with multichannel detectors. The method is tested and demonstrated by means of simulated data as well as experimental data. The LC × LC data was obtained using an in-house LC × LC analyzer that couples a first separation column and a second separation column with a multi-wavelength UV absorbance detector. This method only needs data of a single chromatographic run and does not presuppose knowledge of the analyte spectra. Additional measurements on a blank run would no longer be needed. This would save significant time and effort.

## 2. Chemometric method

Analyzed with of LC × LC-DAD, a given sample produces a three-way trilinear array of data $\underline{\mathbf{X}}$ with dimensions of $I \times J \times K$, where $I$ is the number of elution time data points in the second dimension column, $J$ the number of digitized UV wavelengths, and $K$ is the number of elution time data points in the first dimension column. Mathematically, this instrument response matrix $\underline{\mathbf{X}}$ can be shown as a sum of the signal of individual responsive components and the noise. A pictorial description of this is given in Fig. 1A. $N$ is the total number of responsive components. $\underline{\mathbf{X}}_N$ represents the signal of individual components, and $\underline{\mathbf{E}}$ represents the residual error. As background drift also contributes to the response signal, a certain cube $\underline{\mathbf{X}}_b$ in these cubes $\underline{\mathbf{X}}_N$ represent
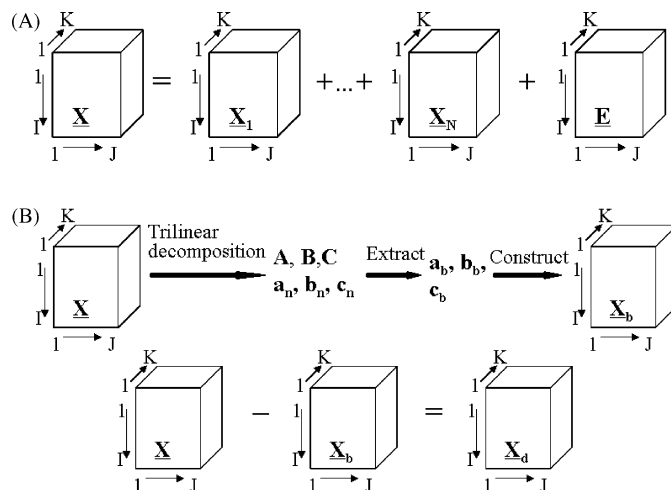


Fig. 1. (A) The 3D data model with $N$ components. (B) Schematic description of the proposed method to remove background drift from 3D data. $N$ is the total number of responsive components, $\underline{\mathbf{X}}$ represents the instrument response signal, $\underline{\mathbf{X}}_N$ represents the signal of individual components, $\underline{\mathbf{X}}_b$ represents the signal of background drift, $\underline{\mathbf{X}}_d$ represents the response signal after background removal, $\underline{\mathbf{E}}$ represents the residual error. $I$, $K$, $J$ are the data points in the second dimension column, first dimension column and spectral dimension, respectively.

the signal of background drift. Alternatively, this 3D data can be written in terms of three vectors for each responsive component: $\mathbf{a}_n$, $\mathbf{b}_n$ and $\mathbf{c}_n$ which collect the profiles for component $n$ in each dimension, respectively. The trilinear decomposition model can be represented as

$$x_{ijk} = \sum_{n=1}^{N} a_{in} b_{jn} c_{kn} + e_{ijk} \tag{1}$$

where $x_{ijk}$ is an element of the $\underline{\mathbf{X}}$. $e_{ijk}$ is the element of array $\underline{\mathbf{E}}$. $\mathbf{a}_n$, $\mathbf{b}_n$ and $\mathbf{c}_n$ represent the second dimension chromatographic profile, the UV spectrum and the first dimension chromatographic profile of the factor $n$ in the model, respectively. $a_{in}$, $b_{jn}$ and $c_{kn}$ are the intensities at the elution time $i$ of the second column, UV wavelength $j$, and elution time $k$ of the first column, respectively. The column vectors $\mathbf{a}_n$, $\mathbf{b}_n$ and $\mathbf{c}_n$ are collected into the corresponding profile matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ ($\mathbf{a}_n$ and $\mathbf{b}_n$ are normalized to unit length).

The model described in Eq. (1) defines a decomposition of $\underline{\mathbf{X}}$ which can provide access to separation profiles ($\mathbf{A}$ and $\mathbf{C}$) and spectral profiles ($\mathbf{B}$) of individual components in the sample. Once vectors that representing the profiles of individual components $\mathbf{a}_n$, $\mathbf{b}_n$ and $\mathbf{c}_n$ are obtained, vectors that corresponding to the background drift (denoted as $\mathbf{a}_b$, $\mathbf{b}_b$ and $\mathbf{c}_b$ in this paper) can be extract and reconstructed into a 3D dataset $\underline{\mathbf{X}}_b$. We then subtract date $\underline{\mathbf{X}}_b$ from the original raw data $\underline{\mathbf{X}}$ to give a difference data $\underline{\mathbf{X}}_d$, thus the background drift is removed from the raw data. The data $\underline{\mathbf{X}}_d$ reproduces the original peaks of the sample, but now on a flat baseline, with no distortion of the original peak shape. Fig. 1B is a schematic description of this method to remove the background drift from three-dimensional instrumental data.

To determine the responsive components, a method known as pseudo-sample extraction and the projection technique (PPT)

[20] was used. Algorithms that based on trilinear decomposition of three-dimensional data can be employed to perform the trilinear decomposition in this paper, e.g. parallel factor analysis (PARAFAC) [21], alternating trilinear decomposition (ATLD) [14], self-weighted alternating trilinear decomposition (SWATLD) [22], and alternating penalty trilinear decomposition (APTLD) [23]. In this paper, the technique to remove background drift is based on the use of ATLD. It is an iterative algorithm which has been proposed for trilinear decomposition of three-way data arrays and second-order linear calibration to a complicated analytical problem [24,25]. It has the properties of fast convergence and being insensitive to the excess of factors used in the calculation. Discussed in this paper is the application of ATLD to removal of three-dimensional background drift in 2D chromatographic analysis, which is for the first time to resolve background drift problems in multi-dimensional chromatography system. The novelty of this strategy for drift removal is to model the drift as a factor in the trilinear decomposition, and to remove the drift from the raw data along the three-dimension directions.

## 3. Experimental

The simulations and all data processing were done in Matlab environment (MathWorks). The MATLAB codes used for ATLD were developed in our laboratory. Experimental LC × LC-DAD data of the traditional Chinese medicines *Rhizoma chuanxiong* was taken. For detail experimental conditions one can refer to literature [26].

## 4. Results and discussion

### 4.1. Simulations

For demonstration and testing of the proposed background removing method, simulated LC × LC-DAD data was first used. For this purpose, a six-compound system was simulated in LC × LC-DAD data structure, a synthetic background shift along all the three dimensions was added. For illustrating, Fig. 2 shows a three-dimensional plot of the data in the first and second column dimensions. As can be seen, the background shifts in this data is obvious.

In order to remove the background shift, chemometric method of ATLD was applied to the data. As background drift is modeled as well as the six compounds of the system, the trilin-
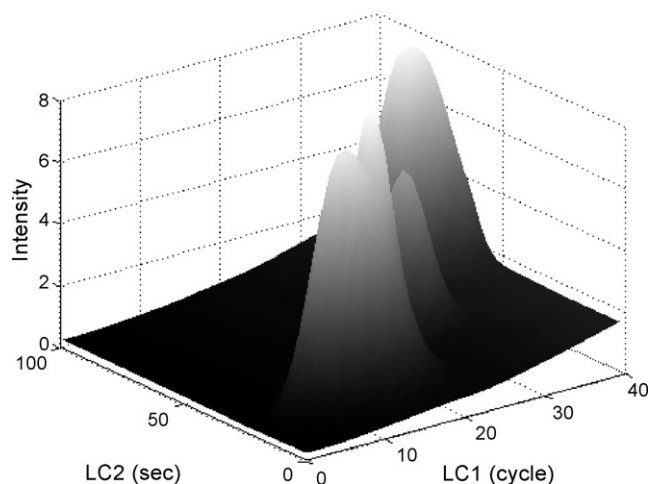


Fig. 2. Three-dimensional chromatogram of simulated data, the absorbance is shown as a function of the first column and the second column.

ear decomposition was applied with a factor number of seven, besides the six compounds, one more factor was used to fit the background. Rank determination method of PPT also indicated that one background component was sufficed to describe the background drift. After decomposition, three profile matrices **A**, **B** and **C** can be obtained. Vectors $a_n$, $b_n$ and $c_n$ are the columns of matrices **A**, **B** and **C**, these vectors represent the signal of individual components in the sample. To extract the vectors that corresponding to the background drift ($a_b$, $b_b$ and $c_b$), one can plot matrices **A**, **B** and **C** against the corresponding number of data points in that dimension. The drift signal can be easy distinguished from that of the real sample components, as their profiles differ widely in shape. In Fig. 3, the extract signal of background drift along the three dimensions is shown. Fig. 3A denotes the drift profile in the second column, Fig. 3B denotes the UV spectral drift, and Fig. 3C denotes the drift in the first column. Once the vectors $a_b$, $b_b$ and $c_b$ that representing the background drift are extracted, they can construct a 3D data $\underline{X}_b$. After subtracting this background data from the original raw data, the noise was removed, leaving the raw data on a flat baseline. In contrast to Fig. 2, a three-dimensional plot of the data after three-dimensional background removal is shown in Fig. 4. It is clear that the peaks are now on a flat baseline. It should be noted that for this brief example we showed, the raw data in the two separation dimensions, actually, the drift in the spectral dimension is serious as it can be seen in the extracted background drift plot in Fig. 3B.
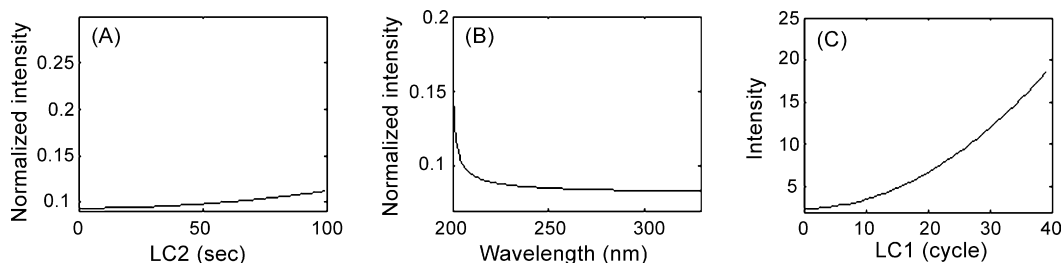


Fig. 3. Background drifts along the three dimensions of the second column (A), the detection (B) and the first column (C).
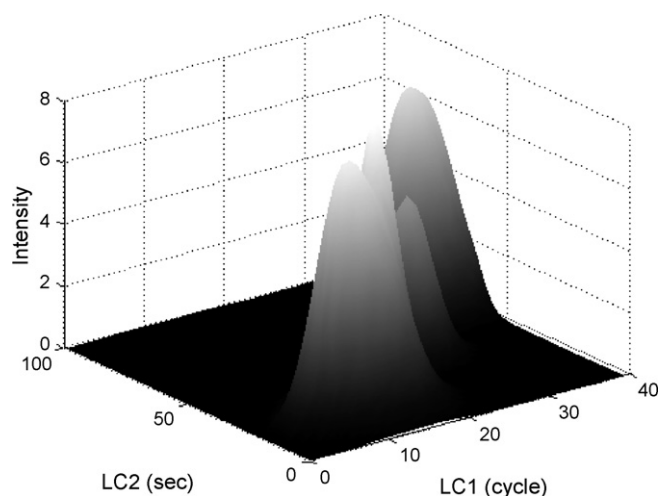
Fig. 4. Three-dimensional chromatogram of simulated data after background drift removal, the absorbance is shown as a function of the first column and the second column.



Fig. 5. Three-dimensional chromatogram of *Rhizoma chuanxiong*, the absorbance is shown as a function of the first column and the second column.

### 4.2. Experimental LC × LC-DAD data

The proposed method was also tested by experimental data. Samples of traditional Chinese medicine *Rhizoma chuanxiong* were separated with LC × LC-DAD. Each LC × LC-DAD analysis produces a 3D array of data as illustrated in Fig. 1A. In this operation, the analysis time reaches as long as 130 min. Due to the complexity of the *chuanxiong* sample and other experimental parameters related to the stability of instrumental response over time, drift characteristic that are not anticipated may appear. A three-dimensional plot of the *Rhizoma chuanxiong* sample is shown in Fig. 5. It shows the absorbance as a function of the first column and the second column. As can be seen, the baseline drift along the first column is clear visible, and little drift appears along the second column. Besides, the drift changes randomly and becomes large with time extended in the first column. As a result, the accuracy of both the qualitative and quantitative analysis may be affected. Method to remove this drift is required.

The ATLD algorithm can be used to remove this background drift. The data processing was done as shown in Fig. 1B. First, this 3D dataset is directly decomposed with ATLD. In model building, the background drift was modeled as one component as well as the constituents of the *Rhizoma chuanxiong* sample. Actually, there are many factors that can cause baseline drift in a chromatographic system, e.g. fluctuations in temperature and
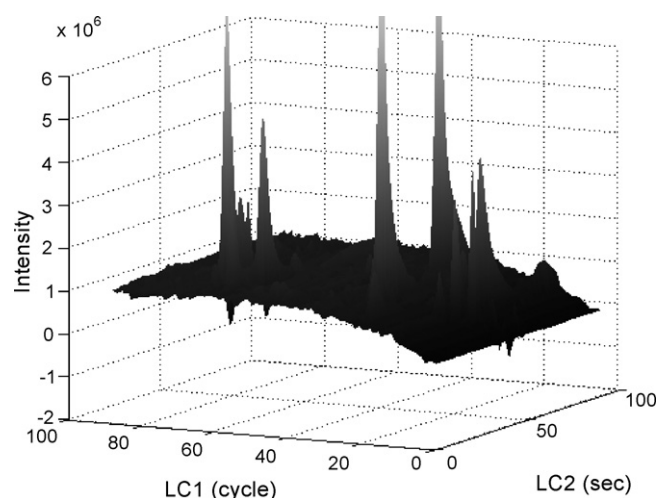
change of the mobile phase in function of time. However, rank determination method of PPT indicates that one background component was generally sufficed to describe the background drift. In the decomposition process, the raw 3D dataset is directly decomposed into three profile matrices related to the second column (**A**), the UV spectrum (**B**), and the first column (**C**). Vectors $a_n$, $b_n$ and $c_n$ that corresponding to the individual components are also obtained. To extract the three vectors $a_b$, $b_b$ and $c_b$ that representing the drift from all these vectors is not a difficult task. One can plot matrices **A**, **B** and **C** against the corresponding number of data points in each dimension, the profile of the drift differs much from that of the real chemical components. In most cases, the background drift are broad time or spectral changes compared to real time or spectral features representing chemical or physical properties of the *Rhizoma chuanxiong* constitutes. In Fig. 6, the profiles of the extracted drift component are shown. Fig. 6A denotes drift profile in the second column, Fig. 6B denotes the UV spectral drift, and Fig. 6C denotes the drift in the first column. As can be seen, drift along the second column is stable at a low intensity, the spectral curve declines gradually from 200 to 250 nm, and then stays stable near a value of zero, but the drift curve in the first column changes continuously with time, moreover, the range it changed is large. It should be noted that the absorbance in the second column and spectral dimension is normalized as described in the theory section. To show the drifts more clearly, three-dimensional plots of the back-
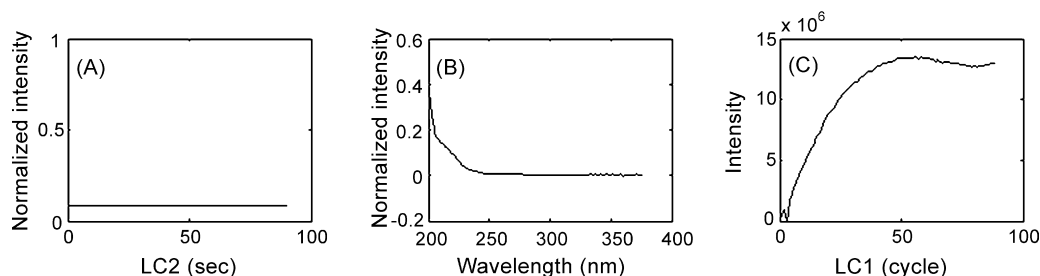


Fig. 6. Background drifts along the three dimensions of the second column (A), the detection (B) and the first column (C).
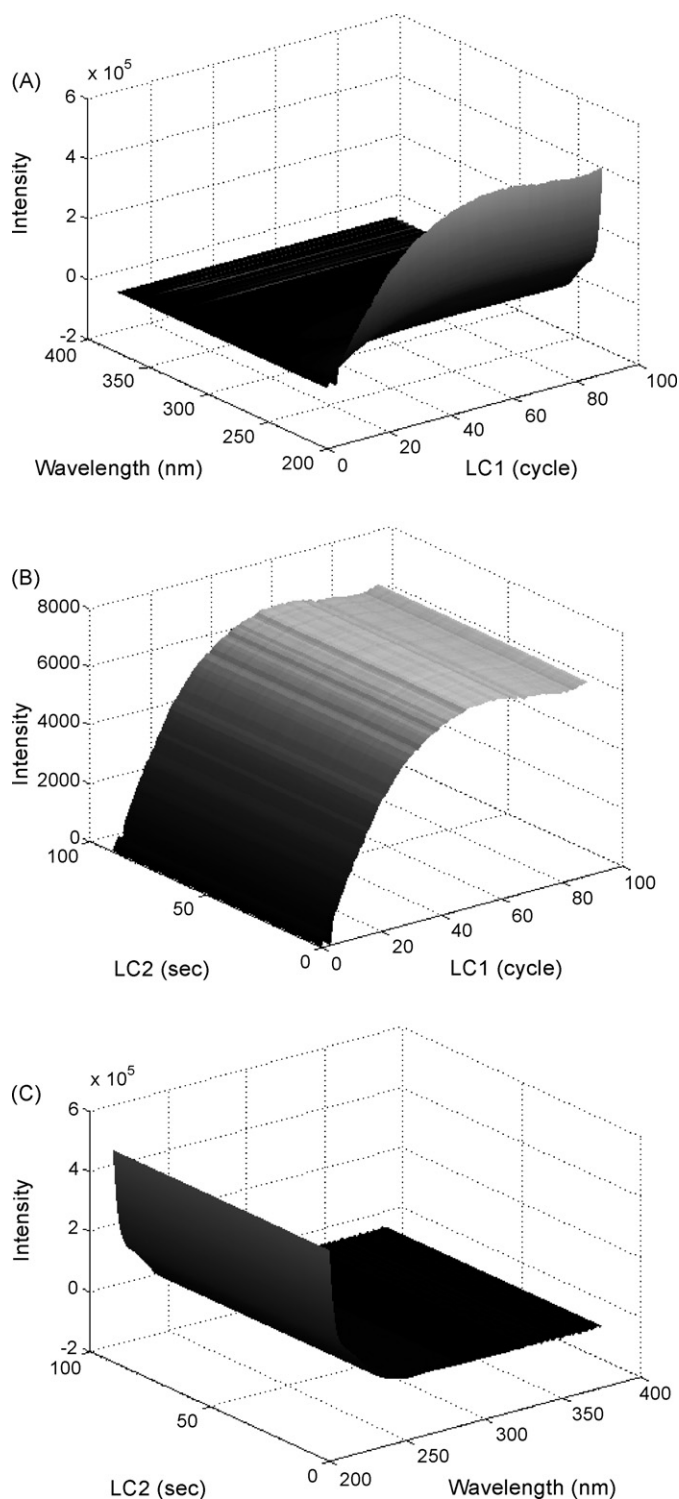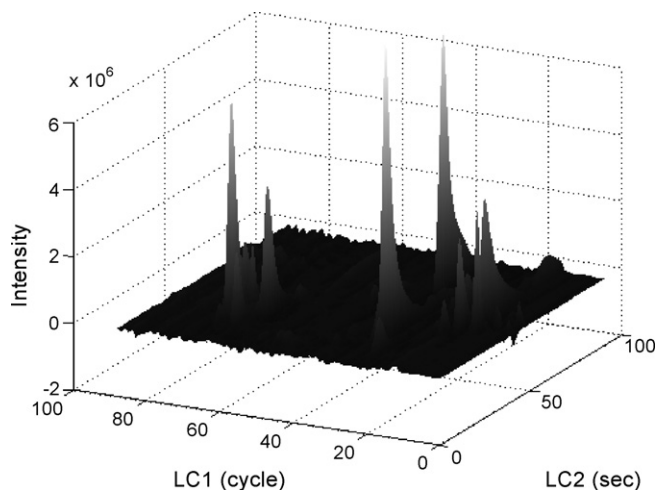
Fig. 8. Three-dimensional chromatogram of *Rhizoma chuanxiong* after background drift removal, the absorbance is shown as a function of the first column and the second column.

shows the absorbance at 49 cycle of the first column as a function of the second column and wavelength. It is clear that the drift in this LC × LC-DAD analysis is serious, especially in the directions of the first column and UV detector.

Once the three vectors that representing the background drift are obtained, the resolved signal for this background drift $\underline{\mathbf{X}}_b$ can be constructed. We then subtract this background data from the original raw data to give the difference data $\underline{\mathbf{X}}_d$, thus the background drift is removed. The difference data reproduces the original peaks of the *chuanxiong* sample, but on a flat baseline, with no distortion of the original peak shape. As an example, a three-dimensional plot of the *chuanxiong* sample after background drift removal is shown in Fig. 8, the absorbance is displayed as a function of the first column and the second column. In comparing with Fig. 5, it visible that the original drift is now effectively removed by the applied chemometric method.

It should be noted that other chemometric methods such as PARAFAC, SWATLD and APTLD can also be used for background drift removal, and similar results as those produced by ATLD can be obtained. However, this paper is not meant to be a comprehensive comparison of all the chemometric methods that can be used for this purpose, but to focus on the very specific issue that has been previously mentioned.

## 5. Conclusions

Background drift is a common phenomenon in comprehensive chromatographic analysis. It poses problems to either qualitative or quantitative evaluation of the analysis. For removing such drift, a novel method based on the use of chemometric trilinear decomposition algorithm is introduced in this paper. The novelty of this strategy is to implement the drift into the calibration model by means of regarding the drift as a component as well as the analytes of interest, thus the drift is simultaneously evaluated with the sample constitutes. The method described here can effectively remove the three-dimensional background drift, while remaining original peak information. It has been



Fig. 7. Three-dimensional plots of the background drift: (A) the absorbance is shown at 63 s of the second column as a function of wavelength and the first column; (B) the absorbance is shown at 270 nm as a function of the second column and the first column; (C) the absorbance is shown at 49 cycle of the first column as a function of the second column and wavelength.

ground drift are shown in Fig. 7. Fig. 7A shows the absorbance at 63 s of the second column as a function of wavelength and the first column, Fig. 7B shows the absorbance at 270 nm as a function of the second column and the first column, and Fig. 7C

demonstrated in drift removal of the simulation and experimental LC × LC-DAD data. Advantages of this method are that it only needs the data of one single chromatographic run and that drift in the two separation dimensions and spectral dimension can be removed simultaneously. The method is easy to program on a personal computer. Also, this technique is applicable to removal of three-dimensional background drift in all comprehensive two-dimensional separation instruments coupled with multichannel detectors.

Furthermore, the technique can be viewed as a good data pretreatment procedure. By stacking numerous of single 2D analysis, in which background drifts were first removed with the described technique, a four-way data array is obtained, thus quantitative analysis can proceed. Our further work will pursue the quantitative study in comprehensive two-dimensional separation instruments coupled with multichannel detectors.

## Acknowledgements

## References

[1] M.M. Bushey, J.W. Jorgenson, Anal. Chem. 62 (1990) 161.
[2] L.A. Holland, J.W. Jorgenson, Anal. Chem. 67 (1995) 3275.
[3] W.C. Quigley, C.G. Fraga, R.E. Synovec, J. Microcol. Sep. 12 (2000) 160.
[4] S. de Koning, H.-G. Janssen, M. van Deursen, U.A.Th. Brinkman, J. Sep. Sci. 27 (2004) 397.
[5] T. Gorecki, J. Harynuk, O. Panic, J. Sep. Sci. 27 (2004) 359.
[6] L. Xie, P.J. Marriott, M. Adams, Anal. Chim. Acta 500 (2003) 211.
[7] M.M. Bushey, J.W. Jorgenson, Anal. Chem. 62 (1990) 978.
[8] J. Zhang, H. Hu, M. Gao, P. Yang, X. Zhang, Electrophoresis 25 (2004) 2374.
[9] D.A. Michels, S. Hu, K.A. Dambrowitz, M.J. Eggertson, K. Lauterbach, N.J. Dovichi, Electrophoresis 25 (2004) 3098.
[10] H. Liu, C. Yang, Q. Yang, W. Zhang, Y. Zhang, J. Chromatogr. B 817 (2005) 119.
[11] B.J. Prazen, R.E. Synovec, B.R. Kowalski, Anal. Chem. 70 (1998) 218.
[12] E. Sanchez, B.R. Kowalski, J. Chemom. 4 (1990) 29.
[13] A.E. Sinha, J.L. Hope, B.J. Prazen, C.G. Fraga, E.J. Nilsson, R.E. Synovec, J. Chromatogr. A 1056 (2004) 145.
[14] H.L. Wu, M. Shibukawa, K. Oguma, J. Chemom. 12 (1998) 1.
[15] A.W. Moore Jr., J.W. Jorgenson, Anal. Chem. 65 (1993) 188.
[16] K.N. Andrew, S.C. Rutan, P.J. Worsfold, Anal. Chim. Acta 388 (1999) 315.
[17] F. Vogt, K. Rebstock, M. Tacke, Chemom. Intell. Lab. Syst. 50 (2000) 175.
[18] E. Furusjö, L.-G. Danielsson, Chemom. Intell. Lab. Syst. 50 (2000) 63.
[19] H.F.M. Boelens, R.J. Dijkstra, P.H.C. Eilers, F. Fitzpatrick, J.A. Westerhuis, J. Chromatogr. A 1057 (2004) 21.
[20] L.Q. Hu, H.L. Wu, J.H. Jiang, Y.J. Ding, A.L. Xia, R.Q. Yu, Anal. Bioanal. Chem. 384 (2006) 1493.
[21] R. Bro, Chemom. Intell. Lab. Syst. 38 (1997) 149.
[22] Z.P. Chen, H.L. Wu, J.H. Jiang, Y. Li, R.Q. Yu, Chemom. Intell. Lab. Syst. 52 (2000) 75.
[23] A.L. Xia, H.L. Wu, D.M. Fang, Y.J. Ding, L.Q. Hu, R.Q. Yu, J. Chemom. 19 (2005) 65.
[24] D.M. Fang, H.L. Wu, Y.J. Ding, L.Q. Hu, A.L. Xia, R.Q. Yu, Talanta 70 (2006) 58.
[25] Y. Zhang, H.L. Wu, Y.J. Ding, A.L. Xia, C. Hui, R.Q. Yu, J. Chromatogr. B 840 (2006) 116.
[26] L.H. Hu, X.G. Chen, L. Kong, X.Y. Su, M.L. Ye, H.F. Zou, J. Chromatogr. A 1092 (2005) 191.