



Determining the number of chemical species in nuclear magnetic resonance data matrix by taking advantage of collinearity and noise



Wanping Wang^a, Limin Shao^{a,*}, Bin Yuan^b, Xu Zhang^b, Maili Liu^b

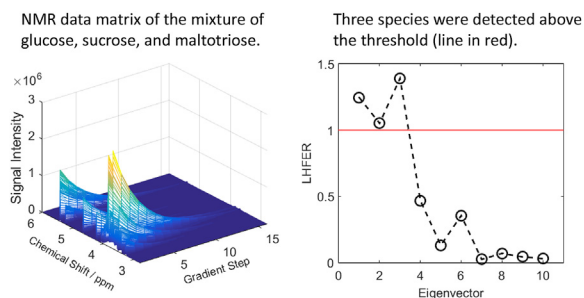
^a Department of Chemistry, University of Science and Technology of China, Hefei, Anhui, 230026, China

^b Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Centre for Magnetic Resonance, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan, Hubei, 430071, China

HIGHLIGHTS

- The method utilizes collinearity and noise rather than eliminates them.
- The method is mathematically rigorous, computationally fast, and readily automated.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 13 October 2017

Received in revised form

14 April 2018

Accepted 17 April 2018

Available online 21 April 2018

Keywords:

Collinearity

Number of chemical components

Pulsed field gradient NMR

ABSTRACT

The number of chemical species is crucial in analyzing pulsed field gradient nuclear magnetic resonance spectral data. Any method to determine the number must handle the obstacles of collinearity and noise. Collinearity in pulsed field gradient NMR data poses a serious challenge to and fails many existing methods. A novel method is proposed by taking advantage of the two obstacles instead of eliminating them. In the proposed method, the determination is based on discriminating decay-profile-dominant eigenvectors from noise-dominant ones, and the discrimination is implemented with a novel low- and high-frequency energy ratio (LHFER). Its performance is validated with both simulated and experimental data. The method is mathematically rigorous, computationally efficient, and readily automated. It also has the potential to be applied to other types of data in which collinearity is fairly severe.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Diffusion-ordered spectroscopy (DOSY) is a widely recognized tool to identify chemical species from complex nuclear magnetic resonance (NMR) spectra [1–5]. DOSY reveals useful information with plots of molecular self-diffusion coefficients versus

chemical shifts. Self-diffusion coefficients are acquired from pulsed field gradient NMR spectra, so it is the acquisition that controls the quality of DOSY. One necessary condition for the acquisition is to determine accurately the number of components in raw NMR data [6].

The determination of the number of components is fairly common in data analysis of complex chemical systems; it is also an essential step for many multivariate analysis methods, such as factor analysis, multivariate curve resolution, multivariate regression, etc [7,8]. Several chemometric methods were developed for

* Corresponding author.

E-mail address: lshao@ustc.edu.cn (L. Shao).

this purpose, including factor indicator function (IND) [9], subspace comparisons [10], RESO [11], Faber-Kowalski F-test [12], Vogt-Mizaikoff F-test [13], and eigenvector comparison [14]. A review article classifies the methods into three categories: (1) empirical, (2) mathematically rigorous, and (3) statistical [15]. From the perspective of application, many methods were developed for chromatographic data, such as LC-NMR [16], LC-DAD [17], GC-MS [18]; a few were for other types of data, such as X-ray absorption spectroscopy [19] and NMR [20]. Most methods are not data-specific, but their performances vary when applied to different types of data. The coexistence of these methods indicates the difficulty of such determination.

Although the aforementioned methods are effective in some cases (and not in others, particularly in experimental data), the common problems they encounter are noise and collinearity. The problem of collinearity is more severe in pulsed field gradient NMR spectral data due to the fact that decay behaviors of molecules are fairly similar. As a result, many methods fail such type of data.

In this paper, a novel method is proposed to determine the number of chemical species in pulsed field gradient NMR spectral data. The proposed method takes advantage of collinearity and noise. Severe collinearity makes linear combinations of decay profiles have similar frequency, which is not the case if significant noise is present in the combinations. Therefore, frequency is a feasible approach to discriminate decay-profile-dominant combinations from noise-dominant ones. For effective discrimination low- and high-frequency energy ratio (LHFER) was designed. After principal component analysis of a pulsed field gradient NMR spectral data matrix, LHFER is calculated for each eigenvector. By counting the number of LHFER values larger than a threshold, one can readily determine the number of decay-profile-dominant eigenvectors, which equals the number of chemical species.

The proposed method was tested with simulated and experimental data, and the results show high accuracy, even in cases of high-level noise or severe collinearity. The proposed method is mathematically rigorous, computationally efficient, and readily automated. With these advantages, the proposed method provides accurate number of chemical species, and improves the reliability of DOSY.

2. Theory

Throughout this paper, bold lower- and upper-case letters denote vectors and matrices, respectively. All vectors are column vectors, the transpose of which are row vectors, indicated with superscript T. The subscript is the matrix size.

Consider that n NMR spectra were measured at different pulsed field gradients, and that each spectrum has m points. By arranging those spectra in a column-wise manner, we obtain an m -by- n matrix, $\mathbf{D}_{m \times n}$. For $\mathbf{D}_{m \times n}$, there exists a bi-linear model,

$$\mathbf{D}_{m \times n} = \mathbf{S}_{m \times p} \mathbf{C}_{n \times p}^T \quad (1)$$

where p is the number of chemical species; $\mathbf{S}_{m \times p}$ and $\mathbf{C}_{n \times p}$ are the NMR spectral and the decay profile matrices, respectively. Each column vector of $\mathbf{S}_{m \times p}$ or $\mathbf{C}_{n \times p}$ is the NMR spectrum or the decay profile of a certain chemical species.

Performing principal component analysis (PCA) on $\mathbf{D}_{m \times n}$ yields

$$\mathbf{D}_{m \times n} = \mathbf{U}_{m \times q} \mathbf{V}_{n \times q}^T + \mathbf{R}_{m \times n} \quad (2)$$

where q is the number of principal components, and theoretically equals p in eq. (1). In eq. (2), $\mathbf{U}_{m \times q}$ and $\mathbf{V}_{n \times q}$ are the principal

component and eigenvector matrices, respectively, which are also known as scores and loadings. Matrix $\mathbf{R}_{m \times n}$ is the residual, which contains information on less significant components such as measurement error, and/or noise. If $\mathbf{R}_{m \times n}$ is negligible, the following equation can be derived from eqs. (1) and (2)

$$\mathbf{S}_{m \times p} \mathbf{C}_{n \times p}^T = \mathbf{U}_{m \times q} \mathbf{V}_{n \times q}^T \quad (3)$$

Equation (3) shows that the column vectors in $\mathbf{C}_{n \times p}$ (i.e. the decay profiles of pure chemical species) span the same linear space as the column vectors in $\mathbf{V}_{n \times q}$ (i.e. the eigenvectors) do. Therefore, an eigenvector, say $\mathbf{v}_{n \times 1}$, can be expressed as a unique linear combination of the decay profiles of pure chemical species by multiplying $\mathbf{C}_{n \times p}$ with a rotation vector, $\mathbf{r}_{p \times 1}$

$$\mathbf{v}_{n \times 1} = \mathbf{C}_{n \times p} \mathbf{r}_{p \times 1} \quad (4)$$

Equation (4) shows that decay profiles of pure chemical species linearly constitute an eigenvector, and thus characterize its frequency. These decay profiles appear alike, which results in severe collinearity on the one hand, on the other hand makes them have similar frequencies. Therefore, when the decay profiles are combined linearly to be eigenvectors, the frequencies of eigenvectors are more or less the same regardless of the linear form. However, the frequency of an eigenvector increases substantially when significant amount of noise is involved. In other words, decay-profile-dominant eigenvectors have low frequency, whereas noise-dominant ones have high frequency.

The above conclusion implies that frequency could be a means to discriminate decay-profile-dominant eigenvectors from noise-dominant ones. The discrimination can ultimately be used to determine the number of chemical species in a NMR spectral data matrix, because the number of chemical species equals that of decay-profile-dominant eigenvectors.

To implement the frequency-based discrimination, a low- and high-frequency energy ratio (LHFER) was designed. LHFER is defined as the following,

$$\text{LHFER} = \frac{E_{\text{LF}}}{E_{\text{HF}}} = \frac{\int_0^{f_t} |A(f)| df}{\int_{f_t}^{f_c} |A(f)| df} \quad (5)$$

where E_{LF} and E_{HF} denote the low- and the high-frequency energy of an eigenvector, respectively; A denotes the Fourier transform of the eigenvector; f_c is the Nyquist frequency, and f_t is the cut-off frequency. The cut-off frequency usually takes a value of one third of the Nyquist frequency. Eigenvectors with higher LHFER than 1 are considered to be decay-profile dominant, and those with lower LHFER than 1 are noise dominant. Therefore, simply by counting how many LHFER are greater than 1, one can determine the number of chemical species. In practice, a plot of LHFER values helps to clarify and confirm the result. A MATLAB program has been developed to do the sophisticated calculation of LHFER, which is available upon request.

3. Experimental

3.1. Simulated data

Pulsed field gradient NMR spectral data matrices were simulated with eq. (1). NMR spectra required in eq. (1) were generated with Lorentzian peaks,

$$s = h \frac{w^2}{4(x-p)^2 + w^2} \quad (6)$$

where h , w , and p are height, width, and position of the Lorentzian peak, respectively. By adjusting values of p , the overlapping degrees of NMR spectra were simulated.

The required decay profiles were generated with an exponential function,

$$c = e^{-0.01dx} \quad (7)$$

where d is the diffusion coefficient; x is the gradient step, ranging from 1 to 32. By adjusting values of d , the similarity degrees of decay profiles were simulated.

Gaussian noise was added to data matrices. Noise was created with various standard deviations to simulate different noise levels.

3.2. Experimental data

Two mixtures were prepared. Mixture 1 is a solution of methanol (3 μ l), ethanol (4 μ l), 1-butanol (8 μ l), sorbitol (15.59 mg), lysine (14.99 mg) and sucrose (21.13 mg) in D₂O (460 μ l). Mixture 2 is a solution of glucose (10.65 mg), sucrose (12.82 mg) and

Table 1

The critical signal-to-noise ratios that ensured 95% accuracy of the proposed method after 100 runs.

Set of simulated NMR spectra	Set of simulated decay profiles	
	c1	c2
s1	324	1468
s2	2592	12076

maltotriose (17.13 mg) in D₂O (460 μ l).

The pulsed field gradient experiments were performed at 25 °C on a Bruker Avance 600 MHz spectrometer. A Bruker pulse sequence was used with diffusion delay 0.18 s and a net diffusion-encoding pulse width (δ) of 2 ms. Water signal was suppressed by pre-saturation. A spectral width of 16 ppm was used, and 16 k complex data points were acquired with 8 scans for each gradient strength and 4 dummy scans, acquisition time of 1.36 s, and relaxation delay of 1.00 s 32 k complex data points were Fourier transformed using an exponential window with a line broadening value of 0.3 Hz. For mixture 1, 32 gradient strengths ranging from 1.445 to 47.187 G/cm were chosen to give linear space in nominal gradient; for mixture 2, 16 gradient strengths ranging from 1.465 to 47.865 G/cm were chosen.

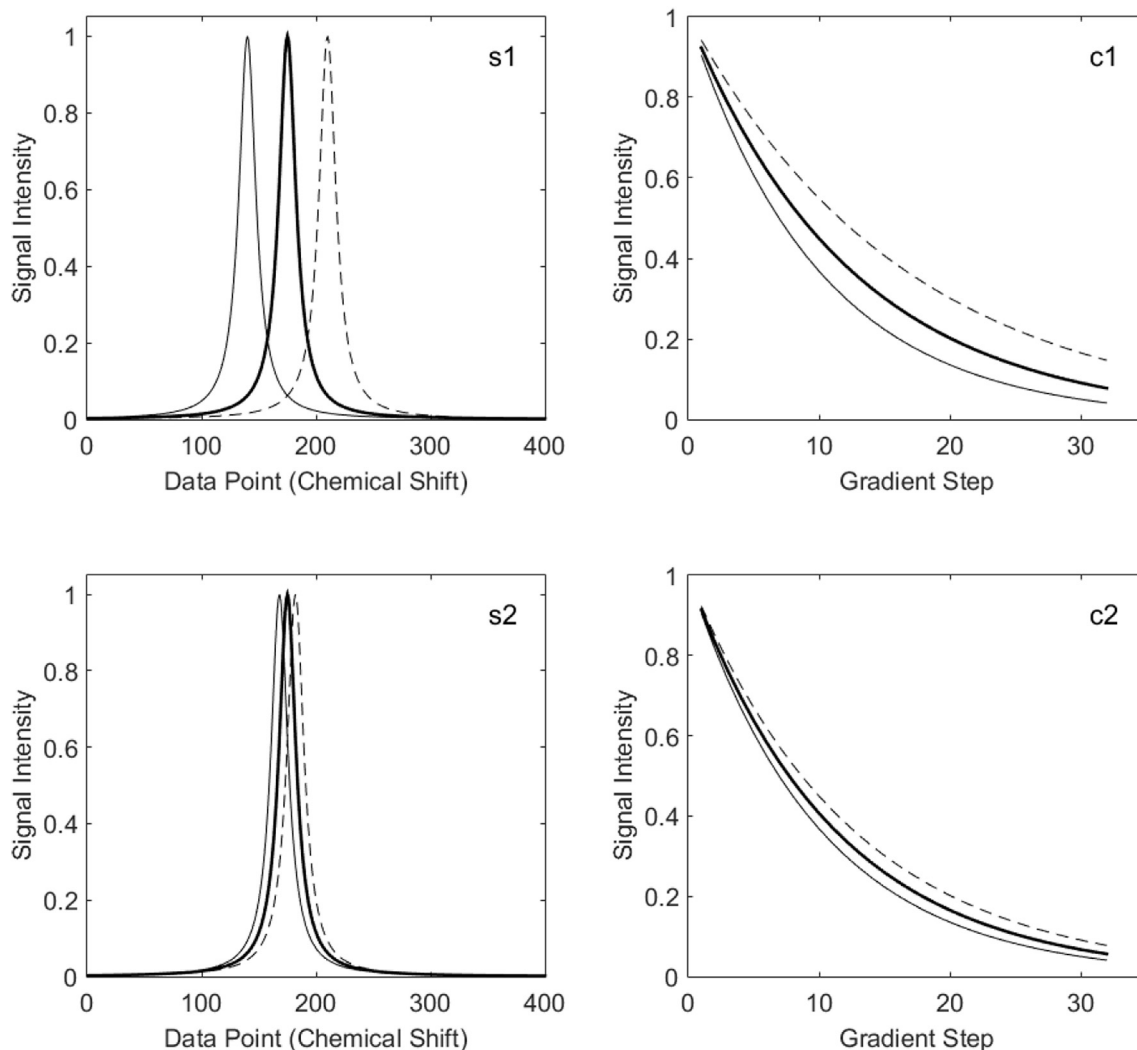


Fig. 1. Two sets of simulated NMR spectra (s1 and s2) and two sets of decay profiles (c1 and c2) for 3 components, in thin solid, thick solid, and thin dashed lines, respectively.

3.3. Data processing

Experimental spectra were phased and baseline-corrected with Topspin 3.2 (Bruker Biospin, Rheinstetten, Germany). To reduce peak shifts, spectra were linearly interpolated by 5 times and aligned with reference to the single analyte peak at 3.0 ppm of the first spectrum. Columns of data matrices were mean-centered in PCA, and eigenvectors were obtained with singular value decomposition (SVD). Manipulations of spectra and data processing were done with self-developed programs in MATLAB 8.5.0/R2015a (The MathWorks Inc., Natick MA) on a Windows 7 platform.

4. Results and discussion

4.1. Simulated data

Two sets of 3-component NMR spectra were simulated with eq. (6); spectra in one set are visibly separated, and those in the other are severely overlapped, shown as s1 and s2 in Fig. 1, respectively. Two sets of 3-component decay profiles were simulated with eq. (7); profiles in one set are relatively different, and those in the other are highly collinear, shown as c1 and c2 in Fig. 1, respectively.

From the two sets of simulated NMR spectra and the two sets of

decay profiles, 4 pulsed field gradient NMR spectral matrices were generated with eq. (1). Of the 4 matrices, the one from s1 and c1 is the easiest to determine the number of components, and the one from s2 and c2 is the hardest. In order to test noise tolerance of the proposed method, Gaussian noise was added to each matrix, and the noise level was gradually increased until the accuracy of determination is below 95% after 100 runs. The critical signal-to-noise ratios are listed in Table 1. In this investigation, signal-to-noise ratio is defined as the ratio of maximum of data matrix to standard deviation of noise.

Decay profiles were simulated with 32 data points. Therefore, in calculating LHFER, the Nyquist frequency is equivalent to 16 data points, and the cut-off frequency, one third of the Nyquist frequency, is equivalent to 5 data points.

In Table 1, the lowest SNR is with the aforementioned “easiest” data matrix. When the overlap among NMR spectra or the collinearity among decay profiles becomes severe, SNR increases to ensure the 95% accuracy of determination. For the “hardest” data matrix, the SNR is the highest.

Fig. 2 shows 3D plots of the aforementioned “easiest” (SNR = 324) and “hardest” (SNR = 12076) matrices and corresponding plots of LHFER values. For the former, features of the 3 components are all distinctive, and even visual inspection of the

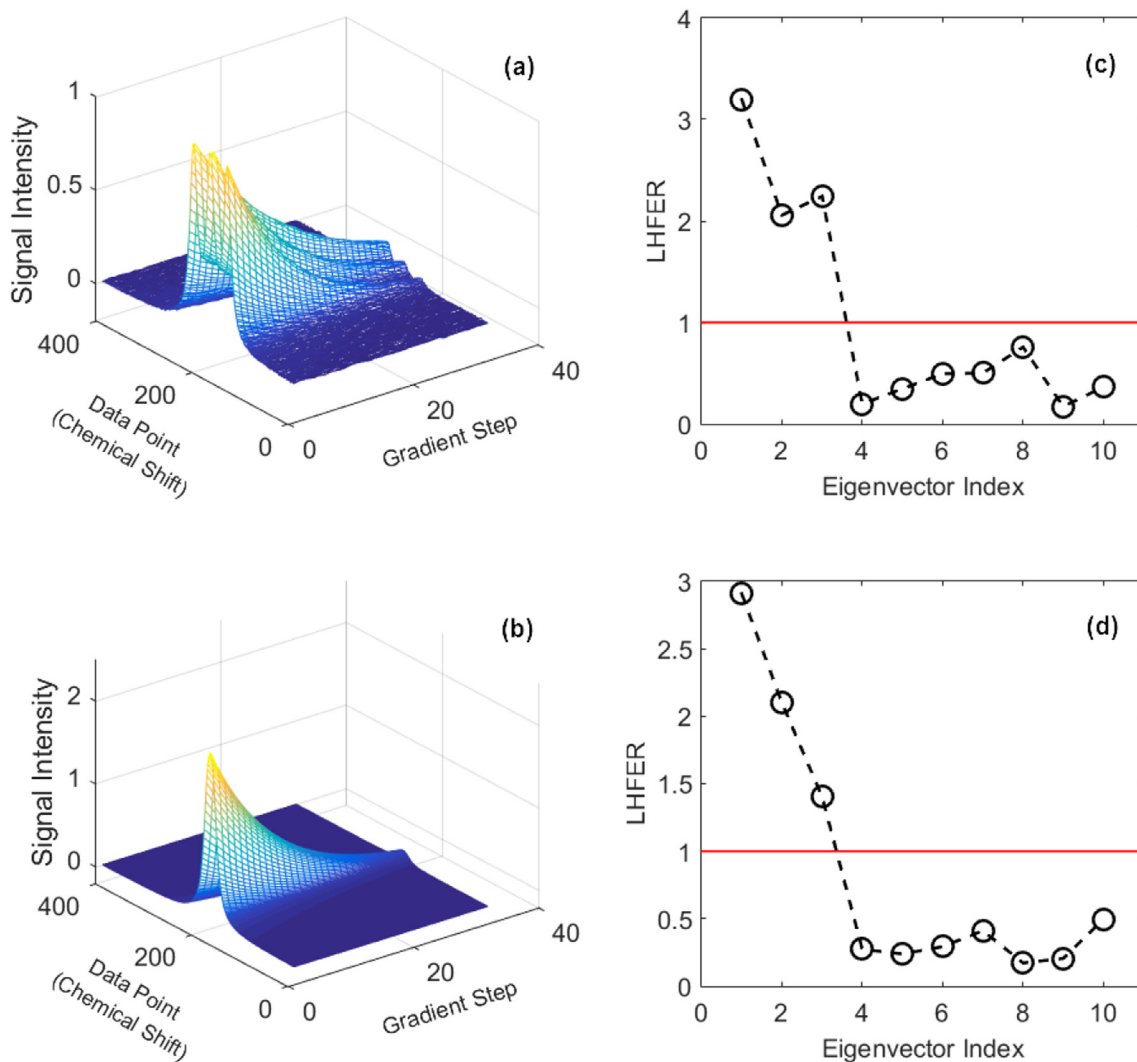


Fig. 2. 3D plots of two matrices with SNR being (a) 324 and (b) 12076, respectively. Plots (c) and (d) are respective results of the method, and the red line indicates the threshold. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

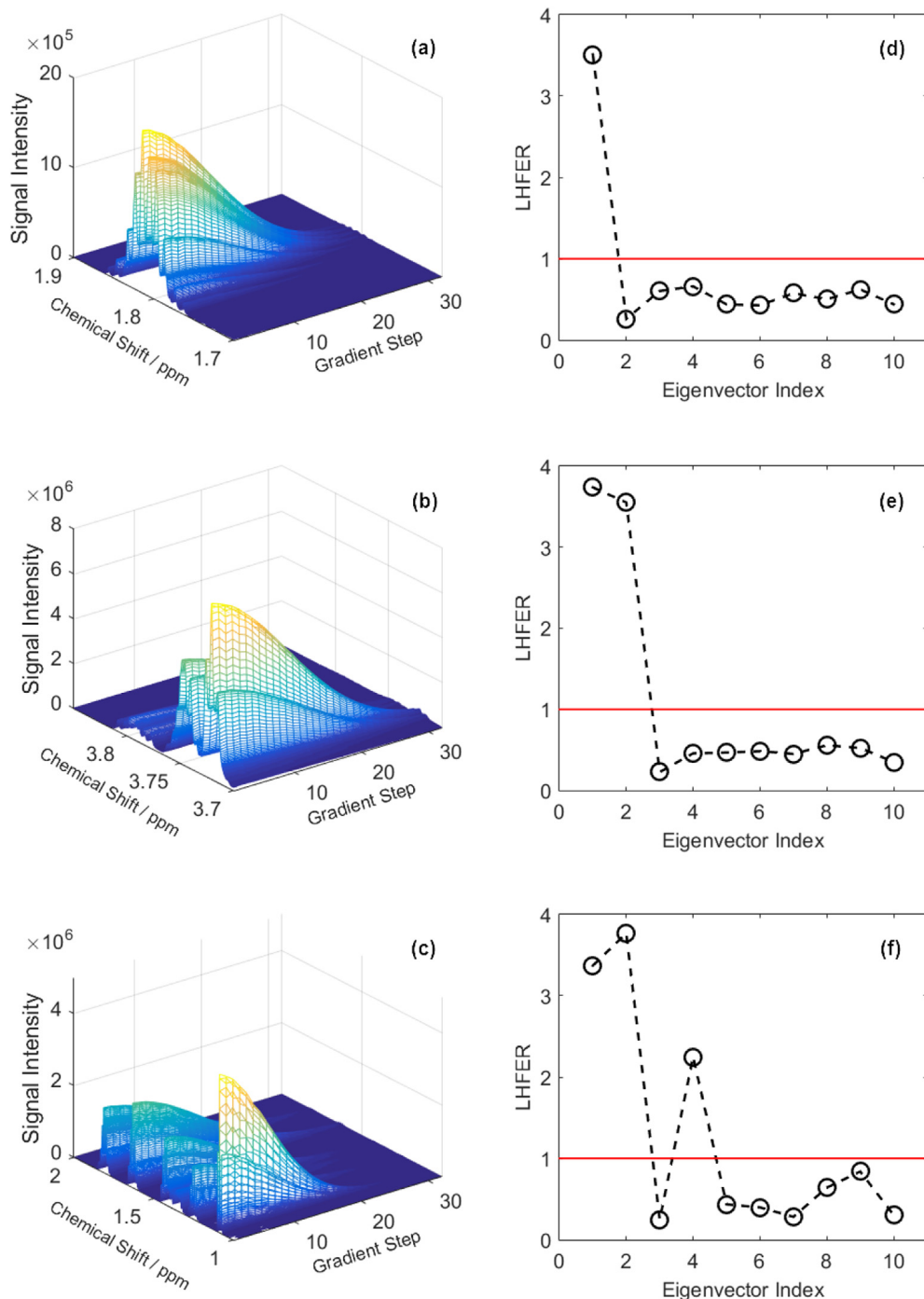


Fig. 3. 3 NMR data matrices constructed with NMR spectra in different regions of chemical shift. The numbers of compounds that contribute NMR information to the 3 matrices are (a) 1, (b) 2, and (c) 3, respectively. Plots (c), (d), and (e) are respective results of the method, and the red line indicates the threshold. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

raw data would yield the correct number of components. For such a data matrix of mild collinearity and overlap, the proposed method demonstrates high tolerance of noise that is fairly apparent in the 3D plot. For the latter matrix, features of individual components are completely indistinguishable due to severe overlap and collinearity. Visual inspection of the 3D plot could determine just one component. In this extreme case, the proposed method still yielded an accurate determination.

4.2. Experimental data of mixture 1

Experimental NMR data contain such interferences as noise and nonlinearity that make the bi-linear model a biased one, and bring much more difficulty for the determination of the number of components. In order to test the proposed method extensively, three data matrices were prepared with NMR spectral segments within 1.9–1.7 ppm, 3.85–3.7 ppm, and 2–1 ppm, respectively. In

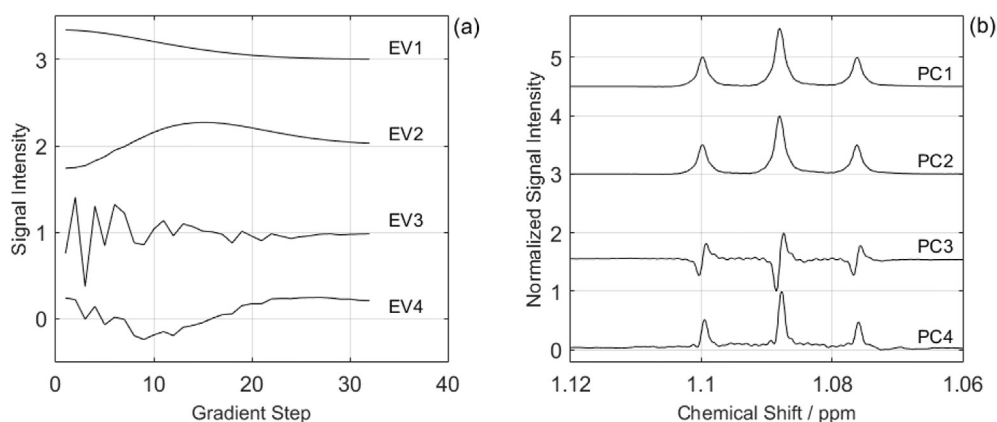


Fig. 4. The first four (a) eigenvectors and (b) parts of principal components of the experimental data matrix composed of NMR spectral segments within 2–1 ppm. In this region of chemical shift, butanol, ethanol, and lysine are the contributors to the NMR spectral information. For clarity, EV1, EV2, and EV3 are displaced by 3, 2, and 1, respectively; each part of PC is normalized to unit length, and PC1, PC2, and PC3 are further displaced by 4.5, 3, and 1, respectively.

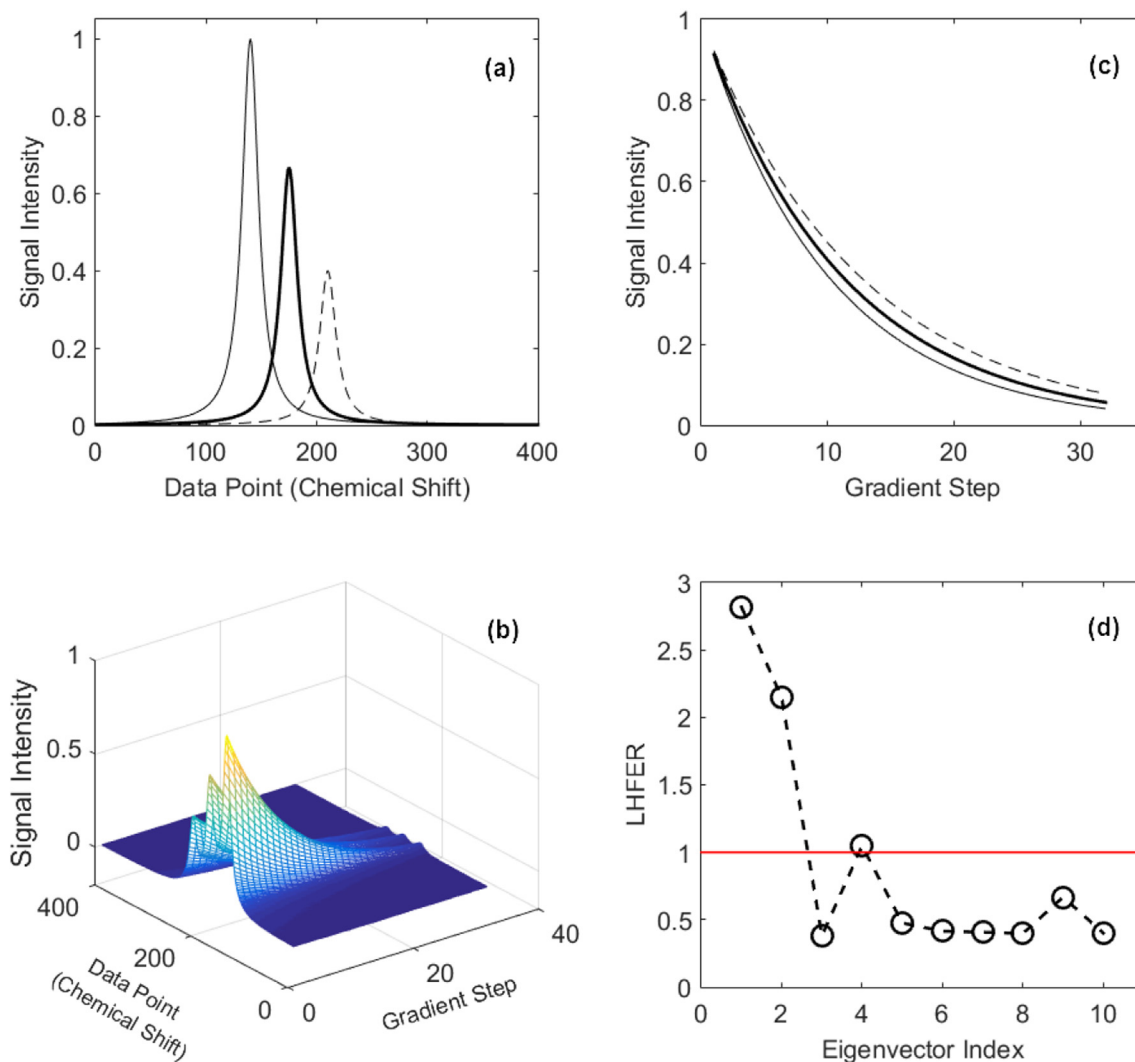


Fig. 5. Simulated (a) NMR spectra and (c) decay profiles for 3 components, in thin solid, thick solid, and thin dashed lines, respectively. Plot (b) is the corresponding data matrix with the SNR being 1187.1. Plot (d) is the result of the proposed method, and red line indicates the threshold. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

the first region of chemical shift, lysine is the only contributor to NMR information; contributors in the second region are sucrose and sorbitol lysine; contributors in the third region are butanol, ethanol, and lysine.

Experimental decay profiles have 32 data points, so the Nyquist and the cut-off frequencies used in calculating LHFER are equivalent to 16 and 5 data points, respectively.

The 3D figures of the three matrices are shown as plots (a), (b), and (c) of Fig. 3, and plots (d), (e), and (f) are the corresponding results of the proposed method. In the 3 plots, the numbers of LHFER values above the threshold are 1, 2, and 3, which is consistent with the numbers of compounds in respective regions. Therefore, the proposed method is validated in all three cases.

From plot (f) of Fig. 3, one could find that the first, the second, and the fourth LHFER values are above the threshold, which means the corresponding eigenvectors are decay-profile dominant, but the third one is not. It is more evident in the plot of the eigenvectors, as shown in Fig. 4(a). In Fig. 4(a), the first and the second eigenvectors are definite combinations of decay profiles, whereas the third eigenvector primarily contains noise; as for the fourth eigenvector, decay-profile feature is clear despite some noise. Fig. 4(a) effectively explains that the first, the second, and the fourth eigenvectors have large LHFER, but the third one has a rather small LHFER in Fig. 3(f). However, this is not a common phenomenon because decay-profile-dominant eigenvectors are generally in consecutive order, e.g. the first 3 eigenvectors of the simulated 3-component data matrix, as shown in Fig. 2. The uncommon phenomenon could be explained by PCA. Fig. 4 shows the first 4 principal components (PC) of the third experimental data matrix; for clarity, only parts within 1.12–1.06 ppm are shown. By comparing the 4 PCs in Fig. 4(b), one can find that PC1, PC2, and PC4 are linear combinations with dominating contributions from the underlying NMR spectra of the species present, but PC3 is characteristic of the first derivative of peak, which reveals NMR peak drifts. Peak drifts make the third PC a significant component, but they are not NMR information, so the corresponding eigenvector is not decay-profile dominant, as found in Fig. 4(a), and its LHFER value is low. Peak drifts in this data matrix are somehow so significant that PCA redistributed them into the third PC due to its variance-maximization nature, ahead of the fourth one that contains NMR information.

Besides NMR peak drifts, relative strong noise and weak signal might also alter the consecutive order of the decay-profile-dominant eigenvectors. In order to investigate such situation, we simulated a 3-component data matrix, and deliberately decreased

NMR signal intensities; the simulated NMR spectra and decay profiles are shown in Fig. 5. In the investigation, the decay-profile-dominant eigenvectors were found to be in consecutive order when noise level was low. When the noise level was increased to be 0.0008 (SNR = 1187.1), the consecutiveness was broken, as shown by plot (d) in Fig. 5. It is the relatively strong noise and weak signal that rendered the third eigenvector to be noise-dominant, and the fourth one to be decay-profile-dominant.

The above results disclosed cases in which the proposed method yielded correct results but the decay-profile-dominant eigenvectors were not in consecutive order. In such cases, non-NMR information in the raw data matrix is somehow significant in terms of variance, be it high noise, severe peak drift, or poor pretreatment of data. So plots of LHFER values are useful not only to confirm results, but also to expose defects of raw data.

4.3. Experimental data of mixture 2

Mixture 2 contains three components, which are glucose, sucrose and maltotriose. The 3D figure of corresponding NMR matrix is shown in Fig. 6; to the right of the 3D figure is the plot of LHFER values. Three LHFER values above the threshold indicate 3 components, which is consistent with the real situation.

The correct result also validates the choice of the cut-off frequency in calculating LHFER, which is one third of the Nyquist frequency. A decay profile in mixture 1 has 32 points, and the one in mixture 2 has 16 points, so the cut-off frequencies used in calculating LHFER for mixtures 1 and 2 are equivalent to 5 points and 2 points, respectively. Although the cut-off frequencies are different for the two independent mixtures, the results were all correct.

The threshold of LHFER is validated too with this data matrix. In other words, value one of LHFER is sufficient to discriminate the decay-profile-dominant eigenvectors from the noise-dominant ones. For data other than NMR, the cut-off frequency and the threshold of LHFER might be different, but can be readily determined with some experimental data.

4.4. Comparison with other methods

The proposed method yielded correct results for the four aforementioned experimental NMR data matrices, in which the numbers of chemical species are 1, 2, 3, and 3, respectively. For comparison, the four data matrices were also processed with seven other methods, namely, determination of rank by augmentation (DRAUG) [8], factor indicator function (IND) [9], ratio of eigenvalues

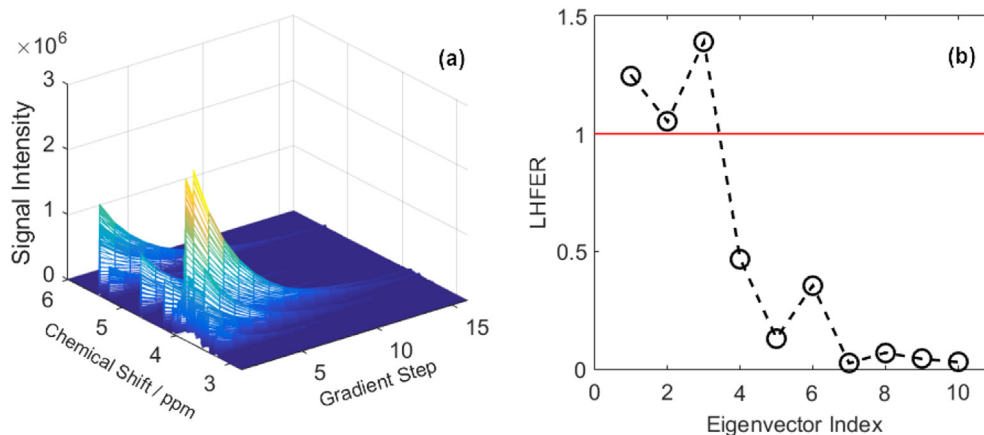


Fig. 6. 3D plot (a) shows the NMR data matrix of mixture 2 (glucose, sucrose and maltotriose). Plot (b) is the result of the method, and red line indicates the threshold. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 2

Numbers of chemical species of the four experimental data matrices determined by other methods.

Experimental Data Matrix	Method						
	DRAUG	IND	RESO	F-test	OPALS	NPFPCA	MS
#1	9	8	1	7	9	1	1
#2	10	15	2	8	5	2	2
#3	10	12	2	10	4	2	2
#4	5	7	1	6	7	3	1

calculated by smoothed principal component analysis and those calculated by ordinary principal component analysis (RESO) [11], F-test [12], orthogonal projection approach and least squares (OPALS) [14], noise perturbation in functional principal component analysis (NPFPCA) [21], and morphological score (MS) [22].

Results of these methods are listed in Table 2. Among the methods, NPFPCA performed the best, yielded 3 correct results; MS and RESO are the second best with 50% correction rate. Methods DRAUG and F-test tend to over-determine numbers of species, which is not uncommon for methods based on statistical principals. The performance of IND and OPALS may be related to collinearity and noise in the experimental data. It should be pointed out that the methods in Table 2 perform satisfactorily for certain types of data, e.g. 2-D chromatographic data matrices, and their performances in this investigation are highly affected by severe collinearity and noise.

5. Conclusion

It is difficult to determine correctly the number of chemical species in a complex system in the presence of collinearity and noise. The difficulty is more acute for pulsed field gradient NMR spectral data because of severe collinearity. The method proposed in this paper takes advantage of the fact that the decay profiles are collinear with a low frequency while noise is random and high-frequency, and implements frequency-based discriminations with

a novel low- and high-frequency energy ratio (LHFER). Its performance is validated with both simulated and experimental data. The method features mathematical rigor, computational efficiency, and easy automation. It also has the potential to be applied to other types of data in which collinearity is fairly severe.

Acknowledgments

This work was funded by the Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT), and the Fundamental Research Funds for the Central Universities (wk2060190040).

References

- [1] A.A. Colbourne, G.A. Morris, M. Nilsson, *J. Am. Chem. Soc.* 133 (2011) 7640.
- [2] M.L. Liu, H.R. Tang, J.K. Nicholson, J.C. Lindon, *Magn. Reson. Chem.* (2002) 40. S83.
- [3] B. Vitorge, D. Jeanneat, *Anal. Chem.* 78 (2006) 5601.
- [4] M. Nilsson, G.A. Morris, *Anal. Chem.* 80 (2008) 3777.
- [5] Y.Y. Du, W.X. Lan, Z.S. Ji, X. Zhang, B. Jiang, X. Zhou, C.G. Li, M.L. Liu, *Anal. Chem.* 85 (2013) 8601.
- [6] K.F. Morris, C.S. Johnson, *J. Am. Chem. Soc.* 115 (1993) 4291.
- [7] E.R. Malinowski, *J. Chemom.* 18 (2004) 387.
- [8] E.R. Malinowski, *J. Chemom.* 25 (2011) 323.
- [9] E.R. Malinowski, *Anal. Chem.* 49 (1977) 612.
- [10] H.L. Shen, Y.Z. Liang, O.M. Kvalheim, R. Manne, *Chemometr. Intell. Lab. Syst.* 51 (2000) 49.
- [11] Z.P. Chen, Y.Z. Liang, J.H. Jiang, Y. Li, J.Y. Qian, R.Q. Yu, *J. Chemom.* 13 (1999) 15.
- [12] K. Faber, B.R. Kowalski, *J. Chemom.* 11 (1997) 53.
- [13] F. Vogt, B. Mizaikoff, *J. Chemom.* 17 (2003) 346.
- [14] S.L. Hao, L.M. Shao, *Chemometr. Intell. Lab. Syst.* 149 (2015) 17.
- [15] W. Lu, L.M. Shao, *Sci. China Chem.* 10 (2014) 1499.
- [16] M. Wasim, R.G. Brereton, *Chemometr. Intell. Lab. Syst.* 72 (2004) 133.
- [17] M. Wasim, R.G. Brereton, *Chemometr. Intell. Lab. Syst.* 81 (2006) 209.
- [18] S. Peters, H.-G. Janssen, G. Vivo-Truyols, *Anal. Chim. Acta* 799 (2013) 29.
- [19] A. Manceau, M. Marcus, T. Lenoir, *J. Synchrotron Radiat.* 21 (2014) 1140.
- [20] B. Yuan, Y.M. Ding, G.M. Kamal, L.M. Shao, Z.M. Zhou, B. Jiang, P. Sun, X. Zhang, M.L. Liu, *J. Magn. Reson.* 278 (2017) 1.
- [21] C.J. Xu, Y.Z. Liang, Y. Li, Y.P. Du, *Analyst* 128 (2003) 75.
- [22] H.L. Shen, L. Stordrange, R. Manne, O.M. Kvalheim, Y.Z. Liang, *Chemometr. Intell. Lab. Syst.* 51 (2000) 37.