NMR IN BIOMEDICINE
NMR Biomed. 2001;**14**:271–277
DOI: 10.1002/nbm.700

NMR spectral quantitation by principal component analysis

R. Stoyanova and T. R. Brown*

Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA

Received 2 January 2001; accepted 2 January 2001

ABSTRACT: The use of principal component analysis (PCA) for simultaneous spectral quantitation of a single resonant peak across a series of spectra has gained popularity among the NMR community. The approach is fast, requires no assumptions regarding the peak lineshape and provides quantitation even for peaks with very low signal-to-noise ratio. PCA produces estimates of all peak parameters: area, frequency, phase and linewidth. If desired, these estimates can be used to correct the original data so that the peak in all spectra has the same lineshape. This ability makes PCA useful not only for direct peak quantitation, but also for processing spectral data prior to application of pattern recognition/classification techniques. This article briefly reviews the theoretical basis of PCA for spectral quantitation, addresses issues of data processing prior to PCA, describes suitable and unsuitable datasets for PCA applications and summarizes the developments and the limitations of the method. Copyright © 2001 John Wiley & Sons, Ltd.

KEYWORDS: peak parameters; spectral estimation; line shape correction

INTRODUCTION

Sophisticated and accurate approaches for quantitation of resonance peaks in a single spectrum (in either the time or frequency domain) are widely available in the NMR community. However, in recent years new instrumental procedures that acquire hundreds of related spectra in a short time have generated a need for approaches which can analyze an entire set of such related spectra simultaneously, taking advantage of the relationships among the spectra to improve the quality of the analysis. Approaches which are able to do this are particularly useful for spectra with low signal-to-noise ratio (SNR) as they utilize the collective power of the data.

This has stimulated new applications of analytical techniques from other fields to spectroscopy data. Some of these are recently developed methods, e.g. artificial neural-networks,[1] Bayesian spectral decomposition[2] and independent component analysis,[3] while others are well known statistical techniques with novel applications to spectroscopy, e.g. principal component analysis,[4] cluster analysis[5] and linear discriminant analysis.[6] In general, the focus of all these methods is spectral classification and pattern recognition, but some can be used to obtain quantitative spectroscopic information.

In this paper we summarize the application of principal component analysis (PCA) to spectral quantitation. PCA was initially introduced for extracting peak intensities using the real part of the frequency spectrum.[7] The potentials and limitations of the approach, and particularly its sensitivity to phase and shift variations among the spectra, were reviewed in detail in a previous publication.[8] The approach has been extended to include estimation of variations in peak positions and phases,[9] which improved peak area quantitation significantly. Elliott *et al.*[10] demonstrated an increased accuracy ($\sqrt{2}$) for estimating the peak areas by analyzing the full complex spectrum. More recent developments have extended the method to quantify all peak characteristics, including the linewidths.[11] Extension of complex PCA to include frequency shifts estimation, as well as comparison between PCA and Hankel total least squares-based methods, are reported in Wang *et al.*[12,15] Below we briefly review the theoretical basis of PCA for spectral quantitation and discuss its applications in light of the above papers.

**Correspondence to:* T. R. Brown, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA.
Contract/grant sponsor: NIH; *contract grant number:* CA41078, CA62556.

Abbreviations used: PCA, principal component analysis.

THEORY

PCA is a well-known statistical technique, which has been used extensively to analyze large multidimensional datasets.[13] It identifies the directions of the largest variations in the data via the principal components ($\vec{\mathbf{P}}_1$, $\vec{\mathbf{P}}_2, \ldots$) and represents the data in a coordinate system defined by the PCs. Typically, the number of coherent changes in large datasets is significantly smaller

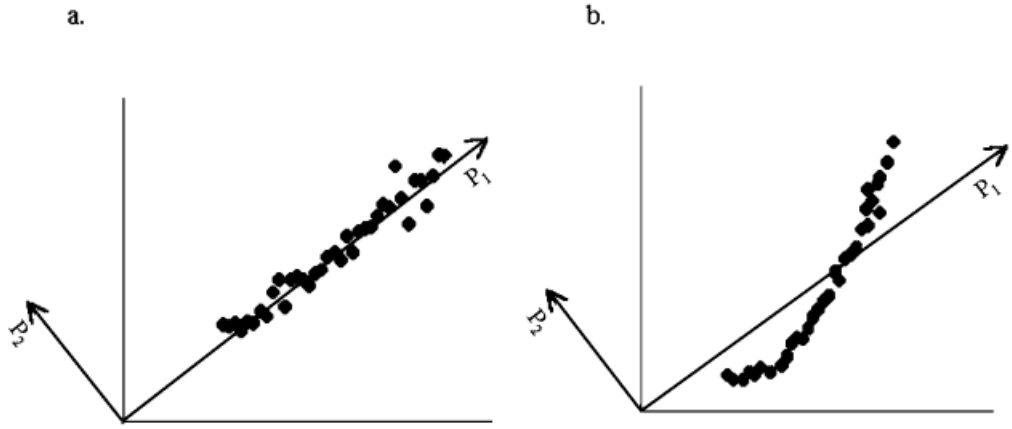

Figure 1. Graphical illustration of the PCA procedure in two dimensions. Principal components of a dataset containing (a) single coherent variation and (b) more than one coherent variation

than the number of original variables and thus PCA provides a representation of the data in a lower-dimensional space of significant variables. To illustrate graphically the PCA procedure, suppose that a two-dimensional dataset **D** of n points is distributed as shown in Fig. 1(a). Then the first PC ($\vec{\mathbf{P}}_1$) will be along the direction of the largest variation, but the second PC ($\vec{\mathbf{P}}_2$), orthogonal to the first, will depend only on the noise in the data and can be ignored. Consequently, the entire dataset can be approximated by the projections of the data points along $\vec{\mathbf{P}}_1$ (called scores) with minimal loss of information, i.e.

$$\mathbf{D} \approx R_1 \vec{\mathbf{P}}_1 \tag{1}$$

This procedure not only reduces the dimensionality of the representation of the data from two to one dimension, but also quantifies, via R_1, the Euclidean distance of the points in the dataset from the origin. The application of PCA for spectral quantitation is completely analogous: if the only coherent variation in a large spectral dataset **D** is the amplitude of a fixed lineshape, then the first PC identifies this shape, regardless of whether it is Lorentzian, Gaussian or a more complex form. Figure 2 presents schematically the PCA decomposition of a dataset of n spectra with m points each, containing a single Lorentzian line $\vec{\mathbf{f}}$ ($\omega_j - \omega_0, \tau_0, \varphi$) ($\omega_0, \tau_0, \varphi$ are the center frequency, decay parameter and phase of the line, $j = 1, \ldots, m$) with varying amplitudes in the presence of noise. Let A_k be the amplitude of the kth spectrum S_k in the dataset ($k = 1, \ldots, n$). Then the peak areas $A = (A_1, A_2, \ldots, A_n)$ can be estimated as:[7]

$$A = \left(\sum_{j=1}^{m} P_{1j} \right) R_1 \tag{2}$$

where P_{1j} is the value of $\vec{\mathbf{P}}_1$ at frequency ω_j.

It can be shown that in this case $\vec{\mathbf{P}}_1$ is the best global estimator of the peak-shape function and thus it can be considered to be an optimum linear filter for the total dataset.[7] PCs are calculated via diagonalization of the covariance matrix C (not mean-centered):

$$C\mathbf{Q} = \frac{1}{m}\mathbf{D}^{\mathrm{T}}\mathbf{D}\mathbf{Q} = \mathbf{Q}\Lambda \tag{3}$$

where Λ is a diagonal matrix of the eigenvalues and **Q** is the matrix containing the eigenvectors of C. The PCs, $\vec{\mathbf{P}}_1, \vec{\mathbf{P}}_2, \ldots$, are the transposed eigenvectors, i.e. $\mathbf{P} = \mathbf{Q}^{\mathrm{T}}$. Since $\mathbf{D} = R\mathbf{P}$ and $\mathbf{P}\mathbf{P}^{\mathrm{T}} = \mathbf{I}$, the scores $R = (R_1, R_2, \ldots)$ are calculated as:

$$R = \mathbf{D}\mathbf{P}^{\mathrm{T}} \tag{4}$$

Any variations in the peak-lineshape $\vec{\mathbf{f}}$, however, interfere with the accurate quantitation of the peak-area.[8] In this case $\vec{\mathbf{P}}_2$ and higher order PCs can no longer be

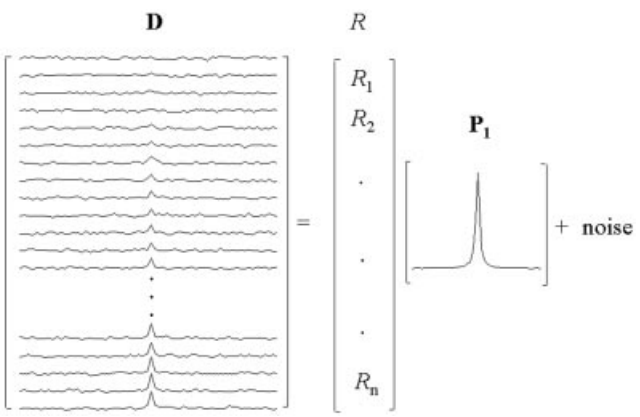

Figure 2. Schematic illustration of the PCA procedure on spectral dataset, containing a single lineshape with varying amplitude

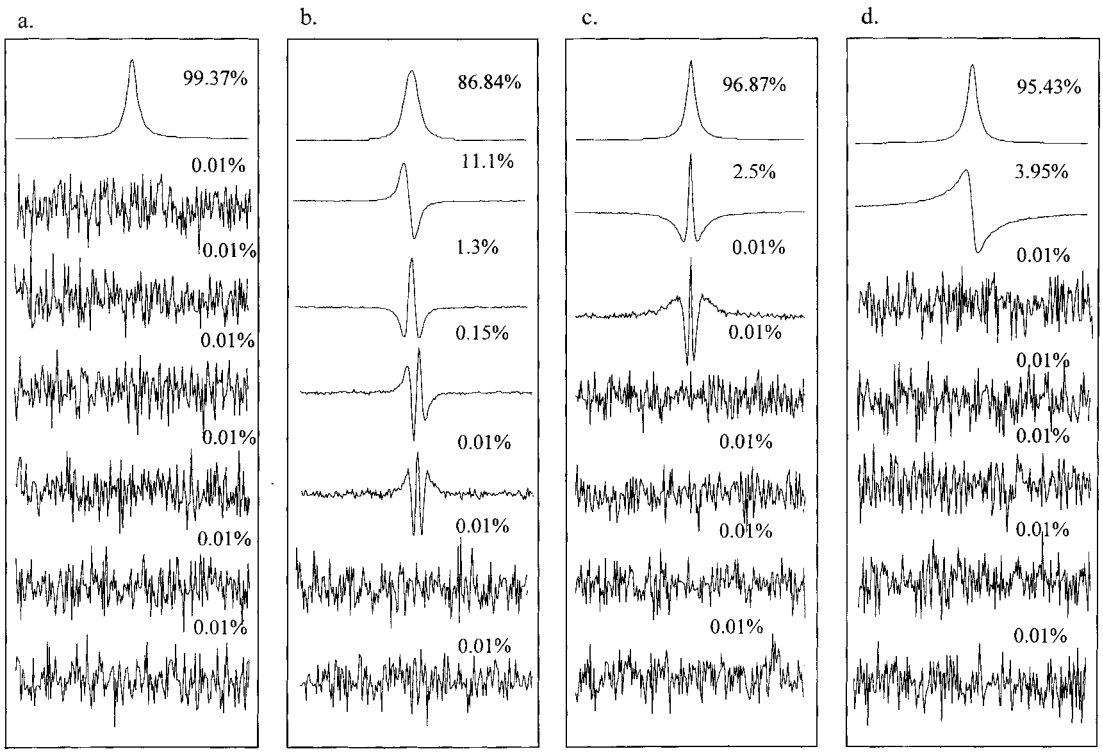

Figure 3. First seven PCs and their corresponding normalized eigenvalues, obtained from spectral data containing a single peak with varying (a) amplitudes; (b) frequency shifts; (c) linewidths; and (d) phases

ignored. A graphical illustration of coherent variations in more than one direction is presented in Fig. 1(b).

In the case of spectroscopy data, the shapes of the higher order PCs indicate what kind of variations (beside the amplitude) are in the data. Experimentally these are usually frequency and/or phase shifts and linewidth variations. We assume that all frequency shifts are within the linewidth of the peak, since it is generally easy to shift the spectra automatically such that the highest point in the peak region in each spectrum is coincident. We simulated four spectral datasets by varying the peak-parameters individually: amplitude, frequency shifts, linewidths and phases. The resultant first seven PCs for each dataset are presented in Fig. 3. For the data with varying amplitude there is only one signal-related PC, as expected. In the presence of frequency and linewidth variations the shapes of $\vec{\mathbf{P}}_2$ and the higher order PCs are related to the derivatives of the signal with respect to the frequency ω and the decay parameter τ. Taylor series expansion of the signal $\vec{\mathbf{f}}$ around ω_0 and τ_0 can be used to model the variations in frequency and linewidth and connect this representation with the PCA decomposition. In the presence of phase variations [Fig. 3(d)] the imaginary part of the signal $\vec{\mathbf{f}}^{\mathrm{I}}$ appears in $\vec{\mathbf{P}}_2$ (here and in the rest of the paper we will use $\vec{\mathbf{f}}^{\mathrm{I}}$ to denote the imaginary part of the signal and $\vec{\mathbf{f}}$ will refer only to the real part). In all cases $\vec{\mathbf{P}}_1$ can be used for initial approximation of $\vec{\mathbf{f}}$.

If only a single parameter were varying then we could connect the PCA representation of the kth spectrum S_k with the individual variations analytically as follows:

$$S_k(\omega) = R_{1k}\,\vec{\mathbf{P}}_1 + R_{2k}\,\vec{\mathbf{P}}_2 + \cdots$$

$$= A_k\left(\vec{\mathbf{f}} + \frac{\partial \vec{\mathbf{f}}}{\partial \omega}\bigg|_{\omega_0}\delta\omega_k + \frac{1}{2}\frac{\partial^2 \vec{\mathbf{f}}}{\partial \omega^2}\bigg|_{\omega_0}\delta{\omega_k}^2 + \cdots\right) \quad (5)$$

$$S_k(\omega) = R_{1k}\,\vec{\mathbf{P}}_1 + R_{2k}\,\vec{\mathbf{P}}_2 + \cdots$$

$$= A_k\left(\vec{\mathbf{f}} + \frac{\partial \vec{\mathbf{f}}}{\partial \tau}\bigg|_{\tau_0}\delta\tau_k + \frac{1}{2}\frac{\partial^2 \vec{\mathbf{f}}}{\partial \tau^2}\bigg|_{\tau_0}\delta{\tau_k}^2 + \cdots\right) \quad (6)$$

$$S_k(\omega) = R_{1k}\,\vec{\mathbf{P}}_1 + R_{2k}\,\vec{\mathbf{P}}_2 + \cdots$$

$$= A_k(\,\vec{\mathbf{f}}\cos\varphi_k - \,\vec{\mathbf{f}}^{\mathrm{I}}\sin\varphi_k)\cdots \quad (7)$$

where A_k, $\delta\omega_k$, $\delta\tau_k$ and φ_k are the peak parameters, i.e. amplitude, frequency and linewidth fluctuations around certain mean frequency and linewidth (ω_0 and τ_0) and the phase dispersion for the kth spectrum in the data.

Connecting the analytical expressions for typical spectral signal variations with the PCA decomposition [eqns (5)–(7)], pointed out by Brown and Stoyanova,[9] provides the theoretical basis for the application of PCA to spectral data.[10–12] Solving these equations with respect to the individual variations A_k, $\delta\omega_k$, $\delta\tau_k$ or φ_k completely describes the individual spectra. If desired the individual spectra can be corrected by $\delta\omega_k$, $\delta\tau_k$ or φ_k, so that the resultant peak possesses the same lineshape in the entire dataset.

In real data the variations appear simultaneously and eqns (5)–(7) must be combined into a single equation. Ignoring the second- order terms in the Taylor series we obtain:

$$\begin{aligned} S_k(\omega) &= R_{1k}\,\vec{\mathbf{P}}_1 + R_{2k}\,\vec{\mathbf{P}}_2 + R_{3k}\,\vec{\mathbf{P}}_3 + R_{4k}\,\vec{\mathbf{P}}_4 + \vec{\varepsilon}_{\mathrm{P}} \\ &= A_k\Big[\vec{\mathbf{f}}\cos\varphi_k + \frac{\partial\vec{\mathbf{f}}}{\partial\omega}\Big|_{\omega_0}\delta\omega_k\cos\varphi_k \\ &\quad + \frac{\partial\vec{\mathbf{f}}}{\partial\tau}\Big|_{\tau_0}\delta\tau_k\cos\varphi_k - \vec{\mathbf{f}}^{\mathrm{I}}\sin\varphi_k\Big] + \vec{\varepsilon}_{\mathrm{f}} \end{aligned} \tag{8}$$

where $\vec{\varepsilon}_{\mathrm{p}}$ and $\vec{\varepsilon}_{\mathrm{f}}$ represent the combined higher order terms in the PCA and Taylor decompositions, respectively. A similar expansion was presented previously[9] without the variable linewidth.

There are variety of methods for solving the above vector equation:

1 The middle and right side of eqn (8) can be projected onto the coordinate system defined by the PCs.
2 The individual spectra can be connected directly to the Taylor series [left and right side of eqn (8)], where PCA results are used to estimate $\vec{\mathbf{f}}, \partial\vec{\mathbf{f}}/\partial\omega$ and $\vec{\mathbf{f}}^{\mathrm{I}}$.[9]
3 The middle and right side of eqn (8) can be projected onto the coordinate system defined by the $(\vec{\mathbf{f}}, \partial\vec{\mathbf{f}}/\partial\omega, \partial\vec{\mathbf{f}}/\partial\tau, \vec{\mathbf{f}}^{\mathrm{I}})$, also called target vectors.

Initially, Brown and Stoyanova,[9] ignoring the linewidth component in eqn (8), undertook the first approach. This procedure performs well when the first three PCs are linear combinations of $\vec{\mathbf{f}}, \partial\vec{\mathbf{f}}/\partial\omega$ and $\vec{\mathbf{f}}^{\mathrm{I}}$. When this is not true, other variations in the data, such as linewidths and baselines, or simply noise in the higher order PCs, can cause misestimation of the parameters. Because of this, when the data are corrected iteratively by these parameters, there is a problem with stopping criteria and convergence.

Witjes *et al.*[11] recently proposed a convergent technique for estimating the peak amplitude, frequency and phase shifts, using the second approach for solving eqn (8). The components of each spectrum along the target vectors $\vec{\mathbf{f}}, \partial\vec{\mathbf{f}}/\partial\omega$ and $\vec{\mathbf{f}}^{\mathrm{I}}$ are obtained using linear regression. They also proposed estimation and correction of the linewidth variations as a separate procedure to be performed once the spectra are frequency and phase shift corrected.

Finally, the third approach for simultaneous estimation of all peak-parameters is to use the results of the PCA analysis and project the middle and right sides of eqn (8) on the functions $\vec{\mathbf{f}}$, $\vec{\mathbf{f}}^{\mathrm{I}}, \partial\vec{\mathbf{f}}/\partial\omega$ and $\partial\vec{\mathbf{f}}/\partial\tau$. The authors are presently implementing this technique and its description and validation will be reported in a forthcoming paper.

The above-mentioned PCA approaches only deal with the real part of the data but can be extended to the analysis of the full complex spectrum by using complex PCA, possibly combined with complex linear regression.[12] This complex PCA procedure was compared with a cross-correlation approach designed to detect large frequency shifts.[12] For large shifts (several times the linewidth) the cross-correlation approach was clearly superior.

PCA procedures,[9,11,12] dealing with estimation of the peak parameters including the frequency shifts need to be iterative, since only the first terms in the Taylor series are retained in the analysis. Once the variations are initially estimated, the data are corrected to remove them. For example, the peaks are aligned by shifting them individually by the estimated amount of frequency variation. Consequently, the variations are smaller at each subsequent iteration. The first PC better represents the true signal in the data and the error associated with ignoring the higher order terms in eqn (8) becomes increasingly negligible. Also, it should be noted that the use of the Taylor series expansion requires that the variation in these parameters should be small, generally less than one linewidth of the peak being measured. On the other hand there are no such limits for the magnitude of phase variations, provided only phase variations are present. In this case, a single step of the PCA procedure [as can be seen from eqn (7)] is sufficient, since the $\sin\varphi$ and $\cos\varphi$ terms completely describe the variations. Elliott *et al.*[10] are able to ignore phase variations of arbitrary magnitude when computing peak areas using complex PCA. In combination with other variations, however, the method does not perform well for phases larger than 60°.[11]

METHODS

Presently used PCA implementations are applied to the signal in the frequency domain and thus routine time-to-frequency domain preprocessing, such as Fourier transform, zero-filling, line broadening and phasing is required.

To demonstrate the estimation and correction procedure, we simulated a dataset by multiplying a single lorentzian line ($m = 512$, $\omega_0 = 0$ Hz, $\tau_0 = 0.008$ s, $\varphi_0 = 0$, height = 1) with 500 uniformly distributed random

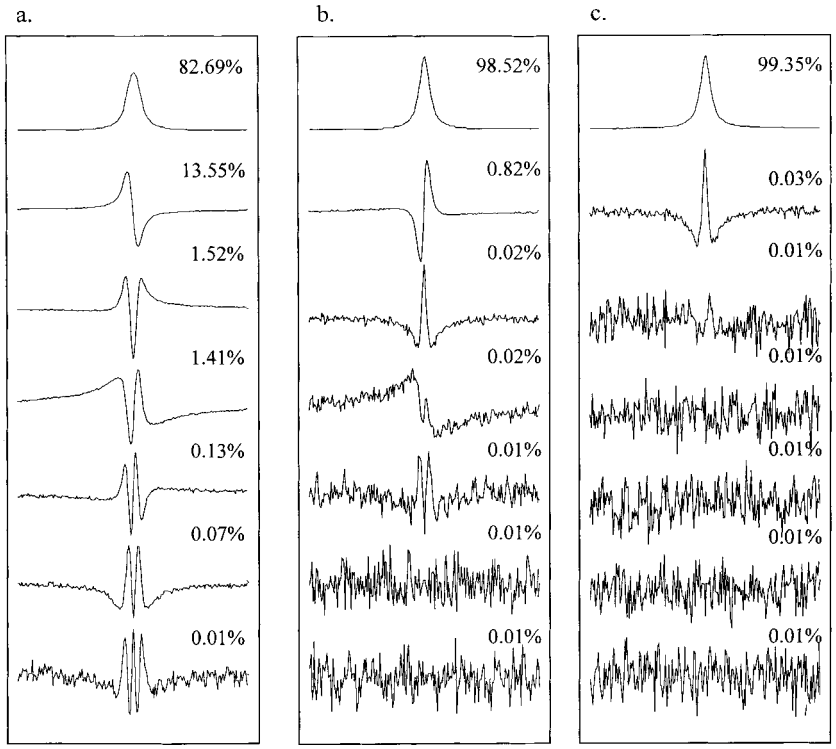

Figure 4. First seven PCs and their corresponding normalized eigenvalues, obtained from (a) spectral data containing a single peak with varying amplitudes, frequencies, linewidths and phases, (b,c) after first and fourth iteration of the PCA procedure and adjusting for phase and frequency shifts

numbers between 20 and 100. After adding 500 sets of Gaussian distributed white noise (mean of 0 and a variance of 1) the SNR of the resultant dataset was between 10 and 50 (SNR is defined as the ratio between the height of the peak and twice the standard deviation of the noise). Further we added random frequency shifts (within one linewidth), linewidth variations (within an eighth of the linewidth) and phase variations (within 25°). The first seven PCs of this dataset are presented in Fig. 4(a). It is apparent that the second and higher order PCs are mixtures of the higher order PCs from Fig. 3(b,c,d). The data were then corrected for phase and frequency errors, as described previously.[9] PCA was applied again, this time to the corrected dataset with the result shown in Fig. 4(b). The first PC now encompasses a much larger fraction of the total variance; it is apparent that the sixth and higher order PCs are now noise-related. Two more iterations of the correction procedure yield the PCs presented in Fig. 4(c). At this point only linewidth variations are left [compare with Fig. 3(c)]. A Java program, performing all functions, presented in this example, can be obtained by request from the authors.

It should be noted that when there are only amplitude variations in the data the average spectrum could serve as an excellent approximation of the first PC [see Fig. 1(a)]. This is the basis of the natural lineshapes approach, proposed by F. E. Heineman *et al.*[14] For data acquired in a well-controlled manner on a stable magnet at high resolution, this approach is completely adequate. However, in the presence of variations other than amplitudes in the data, the average spectrum will not be an accurate estimate of the correct lineshape [see Fig. 1(b)].

To implement PCA for a series of NMR spectra requires a reference peak that is present in all spectra in the dataset whose behavior is related only to the experimental/instrumental artifacts. Correction factors are determined on this reference peak and then applied to the entire spectrum. Once the experimental artifacts are removed from the reference peak, the analysis focuses on the rest of the peaks. The remaining variations in the data can be linked with anatomical or biochemical changes that occur during the experiment.

DISCUSSION

When choosing an analysis method careful consideration must be given to a number of factors. The accuracy and bias of the final estimates is usually of primary concern. For linear estimators it can be shown that PCA, in the presence of amplitude variations, provides the best 'global' estimate for the peak areas in the sense that the sum of squares of the difference between the true and estimated areas is a minimum. Note that techniques such

as HTLStack and HTLSsum, which also use information from the entire dataset, generally approach this optimal linear accuracy[12,15] when applied to Lorentzian lineshapes.

It has been shown also that when quantifying a single resonance, PCA outperforms the other methods for peaks with gaussian lineshape.[15] This brings us to the second important point for choosing the right quantitation procedure: how well the model function of the method fits the experimental signal. One of the great advantages of PCA is that it is a model-free technique—it does not require an assumption for the peak-shape nor any other prior information in terms of frequency positions, etc.

Another criterion is how well the method handles peaks with very low SNR. With PCA they are quantified to an accuracy set by the SNR. This is possible because PCA utilizes the entire dataset to determine an estimated peak shape. Clearly, there is no advantage for using PCA to quantify a single spectrum. In fact, if the true shape is known then a variety of methods for peak estimation such as HTLSstack and HTLSsum[16] are superior to PCA, in particular when the number of spectra is low (since then the use of additional prior knowledge is important).

Finally, the ability of PCA to deal with small experimental variations from spectrum to spectrum is very useful in cases where there is a reference peak, known to be independent of experimental conditions. This peak can be used to align and phase the entire spectrum, removing unwanted variations from other peaks so that the true underlying variations can be detected without being obscured.

PCA is well suited for a variety of closely related spectroscopy data, such as chemical shift imaging data,[17,18] kinetic experiments with muscle or perfused cells,[19,20] high-resolution spectra from tissues and biofluids. On another hand, PCA is not suitable for datasets which contain spectra acquired under very different conditions—different instruments, different acquisition parameters (e.g. number of points in the FID, spectral bandwidth, saturation pulses, decoupling). In such cases the inherent variations in the dataset are very likely to interfere with the analysis.

Since presently used PCA implementations are in the frequency domain, PCA suffers from all complications typical for frequency domain quantitation methods: baseline variations, peak wings-truncation effects and peak overlap. Thus datasets acquired with water or solvent suppression sequences have to be corrected for distortion by the baseline signals prior to PCA. PCA is not able to quantify the individual peaks when they overlap and do not change their relative amplitudes throughout the dataset. Even in cases when the overlapping peaks are changing their relative amplitudes the performance of the procedure is impeded by lack of straightforward methods for recovering the individual peak shapes. Stoyanova *et al.*[7] use an *ad-hoc* procedure for combining the PCs to obtain factors containing only a single peak and hence obtain quantification. Alternatively, the recently developed technique of Bayesian spectral decomposition[2] could be used to calculate the individual positive peak shapes, which can be quantified as above. Wang *et al.*[15] suggest the use of frequency selective filters prior to PCA for isolating the signal just at a given frequency.

PCA is also useful for processing data prior to application of high level classification or pattern recognition techniques. Typically, in *in vivo* NMR experiments changes in the peaks are subtle and often masked by the instrumental/experimental variations. This may seriously impede the process of pattern discovery and identification. In our opinion, PCA provides an excellent nearly automatic method to analyze large numbers of NMR spectra.

Acknowledgements

This work was funded under NIH grants CA41078 and CA62556.

REFERENCES

1. Ala-Korpela M, Changani KK, Hiltunen Y, Bell JD, Fuller BJ, Bryant DJ, Taylor-Robinson SD, Davidson BR. Assessment of quantitative artificial neural network analysis in a metabolically dynamic *ex vivo* ^{31}P NMR pig liver study. *Magn. Reson. Med.* 1997; **38**: 840–844.
2. Ochs MF, Stoyanova RS, Arias-Mendoza F, Brown TR. A new method for spectral decomposition using a bilinear bayesian approach. *J. Magn. Reson.*, 1999; **137**: 161–176.
3. Nuzillard D, Bourg S, Nuzillard J-M. Model-free analysis of mixtures by NMR using blind source separation. *J. Magn. Reson.* 1998; **133**: 358–363.
4. Hadberg G. From magnetic resonance spectroscopy to classification of tumors. A review of pattern recognition methods. *NMR Biomed.* 1998; **11**: 148–156.
5. Howells SL, Maxwell RJ, Griffiths JR. An investigation of tumour ^{1}H nuclear magnetic resonance spectra by the application of chemometric techniques. *Magn. Reson. Med.* 1992; **28**: 214–236.
6. Preul MC, Caramanos Z, Colins LD, Villemure J-G, Leblanc R, Olivier A, Pokrupa R, Arnold DL. Accurate, non-invasive diagnosis of tumours by using proton magnetic resonance spectroscopy. *Nature Med.* 1996; **2**: 323–325.
7. Stoyanova R, Kuesel AC, Brown TR. Applications of principal-component analysis for NMR spectral quantitation. *J. Magn. Reson. Ser. A* 1995; **115**: 265–269.
8. Kuesel AC, Stoyanova RS, Aiken NR, Li C-W, Szwergold BS, Shaller C, Brown TR. Quantitation of resonances in biological ^{31}P NMR spectra via principal component analysis: potential and limitations. *NMR Biomed.* 1996; **9**: 93–104.
9. Brown TR, Stoyanova RS. NMR spectral quantitation by principal-component analysis. II. Determination of frequency and phase shifts. *J. Magn. Reson. Ser. B* 1996; **112**: 32–43.
10. Elliott MA, Walter GA, Swift A, Vandenborne K, Schotland JC, Leigh JS. Spectral quantitation by principal component analysis using complex singular value decomposition. *Magn. Reson. Med.* 1999; **41**: 450–455.
11. Witjes H, Melssen WJ, in 't Zandt HJA, van der Graaf M, Heerschap A, Buydens LMC. Automatic correction for phase shifts, frequency shifts, and lineshape distortions across a series of single resonance lines in large spectral data sets. *J. Magn. Reson.* 2000; **144**: 35–44.
12. Wang Y, Van Huffel S, Heyvaert E, Vanhamme L, Mastronardi N,

Van Hecke P, Magnetic resonance spectroscopic quantification via complex principal component analysis. In Proceedings of the 2000 5th International Conference on Signal Processing, World Computer Congress 2000, Beijing, China, Baozong Y, Xiaofang T, (eds), **Vol. III**. IEEE: New York, 2000; 2074–2077.
13. Anderson TW. An Introduction to Multivariate Statistical Analysis, 2nd edn. Wiley, New York 1971.
14. Heinman FW, Eng J, Berkowitz BA, Balaban RS. NMR spectral analysis of kinetic data using natural lineshapes. *Magn. Reson. Med.* 1990; **13**: 490–497.
15. Wang Y, Van Huffel S, Vanhamme L, Van Hecke P. Magnetic resonance spectroscopic quantification via complex principal component analysis and Hankel total least squares based methods. ESAT-SISTA report TR 2000-81, ESAT Laboratory, Katholieke Universiteit Leuven, Belgium, July 2000.
16. Vanhamme L, Van Huffel S. Multichannel quantification of biomedical magnetic resonance spectroscopic signals, invited paper for the session on 'Sensor array signal processing'. In Proceedings SPIE Conference on Advanced Signal Processing: Algorithms, Architectures, and Implementations VIII, Luk FT (eds). SPIE: San Diego, 22–24 July, 1998; 237–248.
17. Stoyanova R, Arias-Mendoza F, Brown TR. Making multi-dimensional spectroscopy practical: application of principal-component analysis to CSI data. In High-power Gradient MR-imaging, Oudkerk M, Edelman RR (eds). Blackwell: Berlin, 1997.
18. Gonen O, Murdoch JB, Stoyanova R, Goelman G. 3D multivoxel proton spectroscopy of human brain using a hybrid of 8th-order Hadamard encoding with 2D chemical shift imaging. *Magn. Reson. Med.* 1998; **39**: 34–40.
19. Walter G, Vandenborne K, Elliott M, Leigh JS. In vivo ATP synthesis rates in single human muscles during high intensity exercise. *J. Physiol.* 1999; **519**(Pt 3): 901–910.
20. Aiken NR, Szwergold BS, Kappler F, Stoyanova R, Kuesel AC, Shaller C, Brown T. Metabolism of phosphonium choline by rat-2 fibroblasts: effects of mitogenic stimulation studied using ^{31}P NMR spectroscopy. *Anticancer Res.* 1996; **16**: 1357–1364.