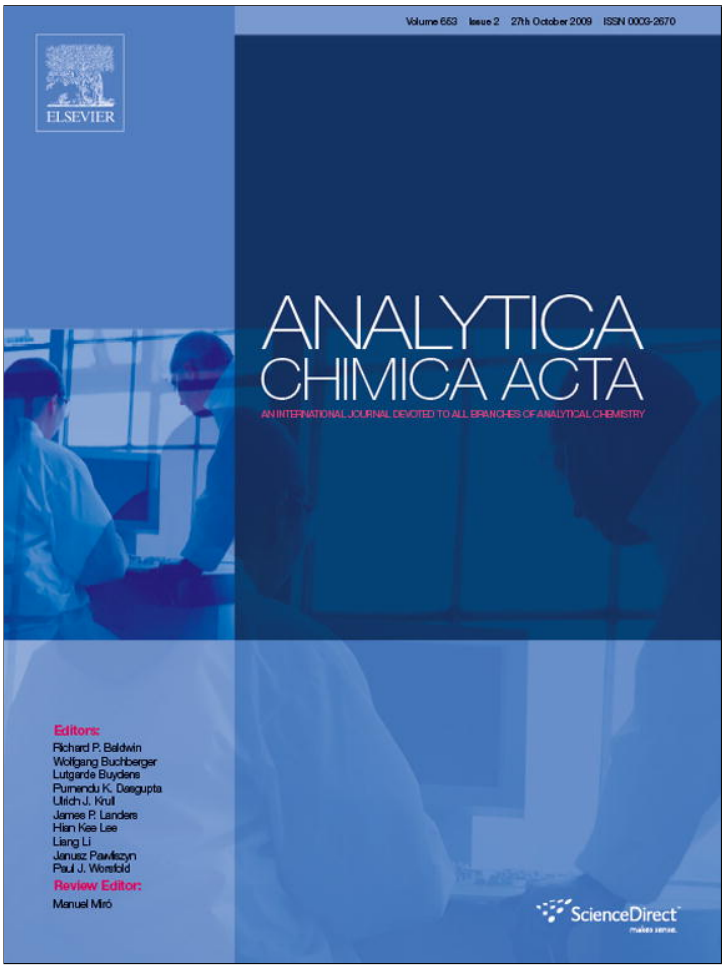

Analytica Chimica Acta 653 (2009) 143–153

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Extraction of multiple pure component ^{1}H and ^{13}C NMR spectra from two mixtures: Novel solution obtained by sparse component analysis-based blind decomposition☆

Ivica Kopriva[a,*], Ivanka Jerić[b], Vilko Smrečki[c]

[a] Division of Laser and Atomic Research and Development, Ruđer Bošković Institute, Bijenička cesta 54, HR-10000, Zagreb, Croatia
[b] Division of Organic Chemistry and Biochemistry, Ruđer Bošković Institute, Bijenička cesta 54, HR-10000, Zagreb, Croatia
[c] NMR Center, Ruđer Bošković Institute, Bijenička cesta 54, HR-10000, Zagreb, Croatia

ARTICLE INFO

Article history:
Received 29 May 2009
Received in revised form
11 September 2009
Accepted 14 September 2009
Available online 18 September 2009

Keywords:
Chemometrics
Nuclear magnetic resonance spectroscopy
Underdetermined blind source separation
Sparse component analysis
Wavelet transform

ABSTRACT

Sparse component analysis (SCA) is demonstrated for blind extraction of *three* pure component spectra from *only two* measured mixed spectra in ^{13}C and ^{1}H nuclear magnetic resonance (NMR) spectroscopy. This appears to be the first time to report such results and that is the first novelty of the paper. Presented concept is general and directly applicable to experimental scenarios that possibly would require use of more than two mixtures. However, it is important to emphasize that number of required mixtures is always *less* than number of components present in these mixtures. The second novelty is formulation of blind NMR spectra decomposition exploiting sparseness of the pure components in the wavelet basis defined by either Morlet or Mexican hat wavelet. This enabled accurate estimation of the concentration matrix and number of pure components by means of data clustering algorithm and pure components spectra by means of linear programming with constraints from both ^{1}H and ^{13}C NMR experimental data. The third novelty is capability of proposed method to estimate number of pure components in demanding underdetermined blind source separation (uBSS) scenario. This is in contrast to majority of the BSS algorithms that assume this information to be known in advance. Presented results are important for the NMR spectroscopy-associated data analysis in pharmaceutical industry, medicine diagnostics and natural products research.

1. Introduction

NMR spectroscopy is of undoubted importance in pharmaceutical and natural products research. It is widely used for the structure elucidation, identification and quantification of impurities or metabolites. NMR spectroscopy is a powerful tool in drug discovery, especially in fragment-based drug design as an alternative to high-throughput screening [1], and in instrumental diagnostics, one of the most developing areas of current medicine [2]. It is clear that NMR spectroscopy represents a forefront of progress in post-genomic era. Particularly, ^{1}H NMR spectroscopy is used for the structure determination of complex molecules and for the quantitative analysis of the most important biological fluids (urine, blood plasma, cerebrospinal fluid, bile, tears, etc.) [2]. The ^{1}H NMR spectrum of any biological fluid is a superposition of the spectra of great number of compounds. Quantification and identification of the components present in the mixture is a traditional problem not only in NMR spectroscopy [3–5] but also in infrared (IR) spectroscopy [6,12], EPR spectroscopy [7,8], mass spectrometry [9,10,12], Raman spectroscopy [11], etc. Identification of the spectra of mixtures proceeds in majority of the cases by matching the mixture's spectra with a library reference compounds [3,6–8]. This approach is ineffective with the accuracy strongly dependent on the library's content of the pure component spectra. Provided that available number of linearly independent mixtures is equal or greater than the number of components, it is possible to separate mixture's spectra into component spectra using only the measurements of the mixture's spectra. This problem is generally known as blind source separation (BSS) and is for described case (more measured mixture's spectra than component spectra) solved by algorithms of independent component analysis (ICA) [11–18]. ICA assumes that pure components are statistically independent and that at most one is normally distributed. The two requirements: to have more linearly independent mixtures than pure components and to have statistically independent pure components seem to be most critical for the success of the BSS approach to blind extraction of the pure components [6,8,9,12]. Significant amount of efforts has

☆ Patent pending under the number PCT/HR2008/000037.

* Corresponding author. Tel.: +385 1 4571 286; fax: +385 1 4680 104.
E-mail address: ikopriva@irb.hr (I. Kopriva).

0003-2670/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2009.09.019

been devoted to relax statistical independence assumption: (i) raw data preprocessing technique by first or second order derivative has been proposed in [6,12] to reduce level of statistical dependence among pure components. This technique belongs to the generalization of the ICA known as dependent component analysis (DCA) [16,21–23]; (ii) an algorithm for blind decomposition of EPR spectra has been derived in [8] minimizing contrast function that exploits sparseness rather than statistical independence among the pure components; (iii) the so called mean field ICA has been proposed in [9] to cope with statistically dependent components. However, all discussed algorithms still require the number of linearly independent measurements to be greater or equal to the unknown number of pure components. Linear independence requirement can be found questionable because it implies concentrations of the pure components to be different in different mixtures. This does not have to be always fulfilled. Thus, a BSS method capable to extract pure components from reduced number of mixtures (that is less than the number of pure components) appears to be of great importance. This leads to underdetermined BSS (uBSS) problem that is not solvable under standard ICA assumptions [24–27]. Additional *a priori* information about pure components such as sparseness is required to solve it. Here, we propose a sparse component analysis (SCA) approach to solve related uBSS problem by having at disposal *two mixtures only* [24–27]. It combines geometric approach known as clustering in wavelet domain to estimate concentration or mixing matrix and linear programming in Fourier/frequency domain to estimate pure components. As opposed to the majority of the ICA/SCA-based BSS algorithms, no *a priori* information about the number of pure components is required because it is also estimated from data during the clustering phase in the time-scale domain. The algorithm assumes that in average only one pure component exist at the each time-scale. This assumption is satisfied with high probability for both ^{1}H and ^{13}C NMR spectra when either Morlet of Mexican hat wavelet is chosen as the basis function. Therefore, it is believed that proposed SCA-based approach to blind extraction of the pure components is practically important.

2. Theory

2.1. Underdetermined blind source separation

The time domain BSS problem is modeled as

$$\boldsymbol{X} = \boldsymbol{AS} \tag{1}$$

where $\mathbf{X} \in R^{N\times T}$ represents matrix of N measured mixtures across T variables, $\mathbf{A} \in R^{N\times M}$ represents the matrix of basis vectors also called the mixing matrix and matrix $\mathbf{S} \in R^{M\times T}$ contains M pure components. We have neglected the additive noise term in (1) due to the fact that used experimental NMR data contain low noise as well as that they were de-noised before blind spectra decomposition is performed. For this purpose we have multiplied recorded time domain NMR data with an exponentially decaying window which is a standard procedure used for de-noising of NMR data [3,35]. In related problem of blind NMR spectra decomposition mixing matrix **A** is also called concentration matrix. That is because coefficients $\{a_{nm}\}_{n=1}^{N}$ represent amount of concentration of the pure component s_m in the mixtures $\{\boldsymbol{X}_n\}_{n=1}^{N}$, where $\mathbf{X}_n$ denotes row vector of **X**. Note that number of pure components M is in principle unknown although many ICA/SCA algorithms assume that it is either known in advance or can be easily estimated. This does not seem to be true in practice. Here, we shall treat M as unknown parameter that will be estimated together with the concentration matrix in wavelet domain by data clustering algorithm. The BSS problem consists of finding the pure component matrix **S** using mixtures matrix **X** only, i.e., mixing matrix **A** is assumed to be unknown. ICA algorithms solve the BSS problem provided that source signals or pure components are statistically independent and non-Gaussian, as well as that $N \geq M$ [14–20]. Then, a solution of the unsupervised decomposition problem (1) is obtained with scale and permutation indeterminacy:

$$\hat{\mathbf{S}} = \mathbf{WX} \tag{2}$$

with $\mathbf{WA} = \mathbf{P\Lambda}$, where **W** represents the de-mixing matrix, **P** is a general permutation matrix and **Λ** is a diagonal matrix. This implies that ICA-based solution of the unsupervised decomposition problem is unique up to the ordering, scale and sign. ICA algorithms find de-mixing matrix **W** though minimization or maximization of the related contrast function $I(\mathbf{W},\mathbf{X})$ that represent statistical (in)dependence measure between $\{\hat{\boldsymbol{s}}_m\}_{m=1}^{M}$. Thus, learning of **W** is achieved by minimizing mutual information between $\{\hat{\boldsymbol{s}}_m\}_{m=1}^{M}$. We recommend [15,16] for detailed description of how the ICA algorithms are constructed. Brief description of the ICA algorithms commonly used in analytical chemistry is also given in [12]. When $N < M$ the BSS problem is underdetermined because there are less measured mixtures N available than unknown pure components M. In order to solve the uBSS problem additional *a priori* information about the pure components must be available [24–27]. *A priori* information that is used most often is sparseness of the source signals in suitably chosen basis [24–27].

2.2. Sparse component analysis

SCA exploits sparseness of the pure components in some *a priori* basis or representation domain. A sparse signal is a signal whose most samples are nearly zero, and just few percent take significant values. While NMR data are not sparse in time domain they are sparse in frequency (Fourier) domain or time-scale domain. For example if pure component would be harmonic (sine or cosine) signal with frequency ω it would contain many non-zero values in time domain but would be perfectly sparse in frequency domain, i.e., there would be only one non-zero component at frequency ω. Signal that has at least $k \ll T$ zero components is called k-sparse. For the solution of related uBSS problem it is however important that pure component signals s_m are mutually sparse. This assumption is satisfied with high probability when each pure component signal is sparse for itself i.e. in such situation it is very likely that only few (one or two) pure component signals will be active (non-zero) at each coordinate in the chosen representation domain. Thus, pure components with no (or only very few) overlaps in spectral domain are considered sparse enough to enable solution of related uBSS problem: blind extraction of pure components from smaller number of mixtures. The SCA approach proposed here differs from the SCA method proposed in [8] by the fact that it maximally exploits redundancy of the linear data model (1) in the chosen basis. The SCA method in [8] solves BSS problem by finding de-mixing matrix **W** by minimizing cost function that measures sparseness of the sources, however it still requires $N = M$. On the other side the SCA approach used here, and referred in [24–27], breaks down BSS problem into two separate problems: estimation of the mixing or concentration matrix **A** using geometric concept known as data clustering [24–29], and estimation of the magnitude spectra of the pure components (based on estimated **A**) by solving resulting underdetermined system of linear equations through linear programming [24,25,30,31], ℓ_1-regularized least square problem [32,33] or ℓ_2-regularized linear problem [34]. In the case of the NMR spectroscopy it is customary to assume that Fourier basis yields sparse representation, however wavelet basis with properly chosen wavelet function can yield even sparser representation. This is because the time domain NMR signals are not pure sinusoids but harmonic signals with amplitude decaying exponentially with

some time constant τ [35]. The real part of the time domain NMR signal can be written as

$$s(t) \cong \sin(\omega t + \varphi)\mathrm{e}^{-t/\tau} \tag{3}$$

where φ represents some arbitrary phase. Provided that wavelet function can be found to resemble structure of the time domain NMR signal (3), real continuous wavelet transform (CWT) [36], at the proper scale a and time shift b

$$S(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s(t)\psi\left(\frac{t-b}{a}\right) dt \tag{4}$$

will yield sparser representation of the NMR signal than Fourier transform. It is the Morlet wavelet that has this property. The real part of the Morlet mother wavelet (defined at scale $a=1$ and shift $b=0$) is of the form [36]

$$\psi(t) = \frac{1}{\pi^{1/4}} \cos(\omega t)\mathrm{e}^{-t^2/2} = \frac{1}{\pi^{1/4}} \sin(\omega t + (\pi/2))\mathrm{e}^{-t^2/2} \tag{5}$$

The other wavelet that resembles the structure of the NMR signal (3) is the Mexican hat wavelet [36]

$$\psi(t) = (1-t^2)\,\mathrm{e}^{-t^2/2} \tag{6}$$

To support statement that Morlet wavelet resembles waveform structure of the time-domain NMR signal and yields its sparse representation we show in Fig. 1a real part of the time domain experimental ^{1}H NMR signal that represents one pure component. Fig. 1b shows Morlet wavelet at the scale that corresponds with the time domain ^{1}H NMR signal shown in Fig. 1a. Fig. 1c and d shows absolute value of the wavelet coefficients in the time-scale domain, scaled by 10^4, obtained by transforming time domain ^{1}H NMR signal shown in Fig. 1a by means of CWT and Morlet wavelet. Fig. 1c shows wavelet coefficients as a function of scale a for the time shift index set to $b=0$. The resolution level or scale a^* at which CWT yields maximally sparse representation is determined by finding the maximal absolute value of the wavelet coefficients at the time shift value frozen to $b=0$. This yields scale value of $a^* \approx 110$. The CWT of the same pure component shown in Fig. 1a with the resolution level frozen to a^* yields a transform that is very sparse in time shift, b, domain, i.e. only few coefficients are non-zero when time shift index b is close to zero. That is shown in Fig. 1d. We would like to point out that we have checked many other types of wavelets in blind extraction of pure components. However, the basis with Morlet and Mexican hat wavelet was the only one that enabled successful extraction of three pure components from two mixtures of ^{1}H NMR spectra in demanding experiment reported in Section 4.

Model (1) is written in time-scale domain as

$$\boldsymbol{x}(a,b) = \boldsymbol{A}\boldsymbol{s}(a,b) \tag{7}$$

where $\mathbf{x}(a,b)$ and $\mathbf{s}(a,b)$ denote column vectors comprised of transformed individual components $\{x_n(t)\}_{n=1}^{N}$ and $\{s_m(t)\}_{m=1}^{M}$. For the solution of the related uBSS problem with $N<M$, majority of clustering algorithms require that signal $\mathbf{s}(a,b)$ is $(M-N+1)$-sparse with $M-N+1$ zero components or with no more than $N/2$ non-zero components. By setting the number of mixtures to be $N=2$ this implies $k=M-1$, i.e., the assumption is that pure components do not overlap in transformed time-scale domain. This assumption is recently relaxed by a concept known as k-plane clustering [28,29]. Robustness with respect to noise and outliers is achieved by assuming that pure components are in average $(M-N+1)$-sparse. Hence, it is allowed that pure components at certain number of time-scale coordinates violate $(M-N+1)$-sparseness assumption. We shall

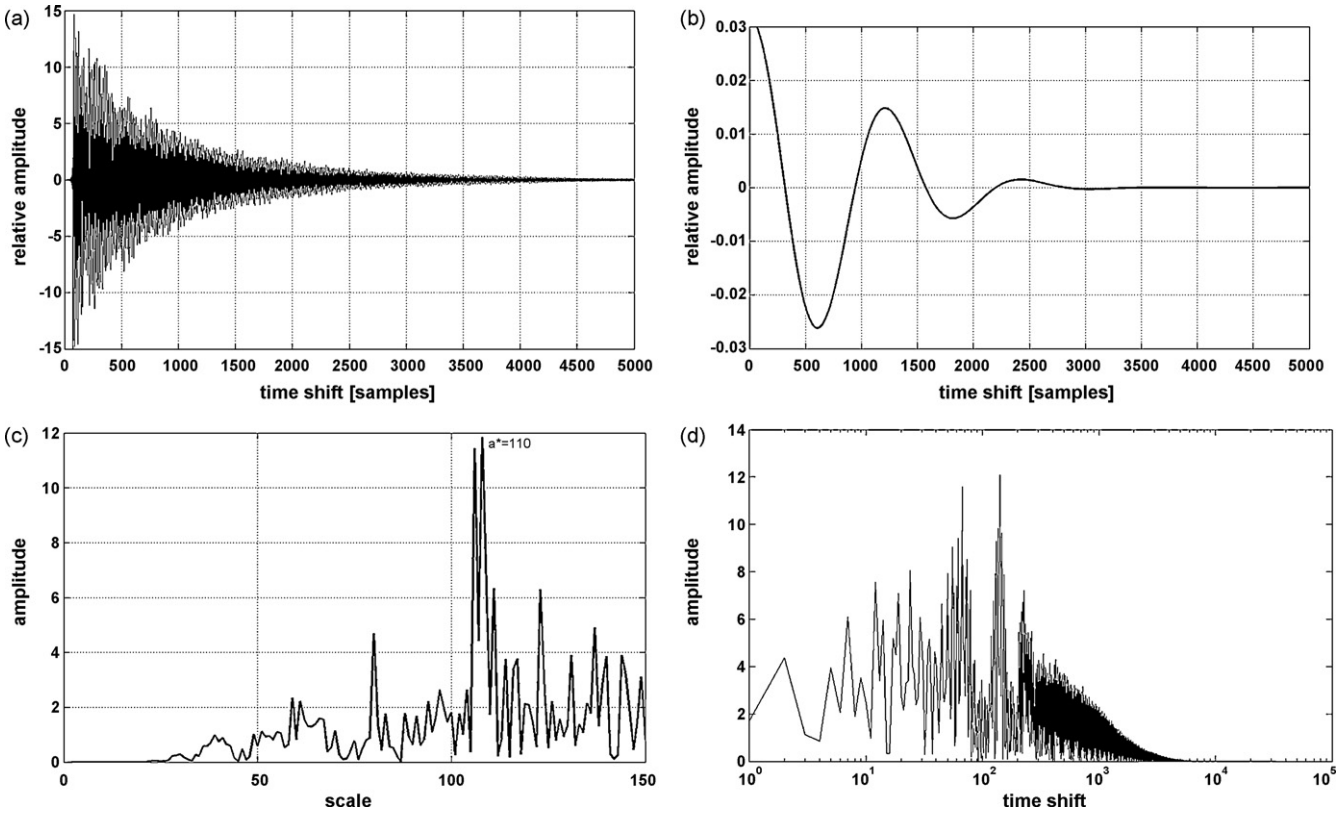

Fig. 1. (a) Real part of the experimental time domain ^{1}H NMR signal that represents one pure component; (b) Morlet wavelet at the scale that corresponds with the time domain ^{1}H NMR signal shown in a; (c) absolute value of the wavelet coefficients, scaled by 10^4, as a function of scale a, obtained by transforming time domain ^{1}H NMR signal by means of CWT and Morlet wavelet and time shift index set to $b=0$. Maximal value occurs at resolution level $a^*=110$; (d) absolute value of the wavelet coefficients, scaled by 10^4, as a function of time shift b, obtained by transforming time domain ^{1}H NMR signal by means of CWT and Morlet wavelet and scale set to $a^*=110$. The abscissa is in logarithmic scale.

assume that pure components in transformed time-scale domain are in average $k = M - 1$ sparse. This implies that at the majority of time-scale coordinates (a,b) only one pure component $s_m(a,b)$ out of M will be non-zero. It is however important to say that $k = M - 1$ sparseness assumption will probably be violated when pure components represent complex chemical compounds with high degree of similarity, such as biomolecules from proteomic research or isolated from natural sources or structurally related compounds usually obtained during natural product synthesis. In such scenarios it will be necessary to increase number of mixtures N from two to three or more. This number depends on level of overlap between pure components in the representation (wavelet) domain. For example, if $J \leq M$ pure components are non-zero at the majority of (a,b) coordinates, than the number of mixtures should satisfy: $N = J + 1$. Nevertheless, it will be demonstrated on demanding experimental problem that $k = M - 1$ sparseness-based concept works not only for ^{13}C NMR data but also for the ^{1}H NMR data. Because the number of coordinates in time-scale domain that deviates from the $k = M - 1$ sparseness assumption was reasonably small, in comparison to the overall number of coordinates, it did not significantly influence accuracy of the clustering-based estimation of the mixing matrix.

2.3. Data clustering algorithm

By assuming the average number of active sources to be 1 (i.e. that in majority of time-scale coordinates only one out of M pure components is non-zero), estimation of the mixing matrix $\mathbf{A}$ can be organized in K-means clustering fashion [37]. Provided that at each time-scale coordinate only one pure component is non-zero the following applies:

$$\mathbf{x}(a, b) = \mathbf{a}_m s_m(a, b) \tag{8}$$

where $\mathbf{a}_m$ represents the column vector of the mixing matrix that corresponds with the concentration profiles of the pure component s_m across the N mixtures. Hence, if we assume that column vectors $\mathbf{a}_m$ are normalized to ℓ_2-unit norm they can be estimated from mixtures data. K-means clustering algorithm assumes that number of clusters M in data set (this corresponds with the number of pure components) is known. It is also assumed that initial values of M cluster centers $\mathbf{a}_m$ are given. Then set of binary indicators $r_{tm} \in \{0, 1\}$ is assigned to each data point $\mathbf{x}_t$ in accordance with:

$$r_{tm} = \begin{cases} 1 & \text{if } \; m = arg\min_j ||\boldsymbol{x}_t - \boldsymbol{a}_j||^2 \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The cluster centers $\mathbf{a}_m$ are updated according to

$$\boldsymbol{a}_m = \frac{\sum_t r_{tm} \boldsymbol{x}_t}{\sum_t r_{tm}} \tag{10}$$

The two phases (9) and (10) of re-assigning data points to clusters and re-computing the cluster means are repeated until there is no further change in the assignment or some predefined number of iterations is exceeded.

However, in the sequel we shall adopt clustering algorithm described in [28] due to its robustness to outliers (data points that violate $k = M - 1$ sparseness assumption) and small number of *a priori* information required by the algorithm. Because we have discussed that solution of the BSS problem by means of ICA algorithms is characterized by scale indeterminacy we shall assume the unit norm constraint (in the sense of ℓ_2 norm) on the columns of the mixing matrix $\mathbf{A}$, i.e., $\{|\boldsymbol{a}_m|_2 = 1\}_{m=1}^{M}$. Since we have assumed the number of mixtures to be $N = 2$ the normalized mixing vectors $\{\boldsymbol{a}_m\}_{m=1}^{M}$ lie in a plane on the unit circle, i.e., they are parameterized as $\mathbf{a}_m = [\cos(\varphi_m)\sin(\varphi_m)]^T$. The clustering algorithm is outlined by the following steps:

(1) In the time-scale domain we find scale a^* that yields the maximal amplitude of the coefficients. We transform data with the fixed scale a^* along the time shift index b.
(2) We remove all data points close to the origin for which applies: $|\mathbf{x}(a^*,b)|_2 \leq \varepsilon$, where ε represents some predefined threshold. This corresponds with the case when all pure components are close to zero i.e. when no pure component exists. It is clear from (8) that such points are irregular from the point of view of determination of the vector of concentration profiles because it would imply pure component is not contained in any mixture at all. It is evident that proper setting of ε requires some *a priori* information related to signal-to-noise ratio. Thus, de-noising of recorded NMR data enables to reduce ε and not to loose weak pure components. Hence, *a priori* information about the ratio between strongest and weakest pure component contained in the mixture is also useful for proper setting of ε. It is however also possible to look at the elimination of small data points as de-noising procedure itself.
(3) Normalize to unit ℓ_2 norm remaining data points $\mathbf{x}(a^*,b)$, i.e., $\mathbf{x}(a^*,b) \rightarrow \mathbf{x}(a^*,b)/|\mathbf{x}(a^*,b)|_2$. This step does not influence the quality of the results due to the fact that in any BSS problem uniqueness of the estimation of the mixing matrix and pure components is possible up to the scale only.
(4) Calculate function $f(\mathbf{a})$, $\mathbf{a} = [\cos(\varphi)\ \sin(\varphi)]^T$

$$f(\boldsymbol{a}) = \sum_{t=1}^{\tilde{T}} \exp\left(-\frac{d^2(\boldsymbol{x}(a^*, b_t) \times \boldsymbol{a})}{2\sigma^2}\right) \tag{11}$$

where $d(\boldsymbol{x}(a^*, b_t), \boldsymbol{a}) = \sqrt{1 - (\boldsymbol{x}(a^*, b_t) \times \boldsymbol{a})^2}$ and $(\mathbf{x}(a^*, b_t) \times \mathbf{a})$ denotes inner product. $\tilde{T} \leq T$ denote number of data points that are remained after elimination process in step 2. Parameter σ in (11) is called dispersion. If set to sufficiently small value, in our experiments this turned out to be $\sigma \approx 0.05$, the value of the function $f(\mathbf{a})$ will approximately equal the number of data points close to $\mathbf{a}$. Thus by varying mixing angle $0 \leq \varphi \leq \pi/2$ we effectively cluster data. The mixing angle is confined in the interval $[0, \pi/2]$ due to the fact that mixing vectors have chemical interpretation as concentration that is a positive quantity. Thus, they must stay in the first quadrant.
(5) Number of peaks of the function $f(\mathbf{a})$ on interval $[0, \pi/2]$ corresponds with the estimated number of pure components $\hat{M}$. Locations of the peaks correspond with the estimates of the mixing angles $\{\hat{\varphi}_m\}_{m=1}^{\hat{M}}$, i.e., mixing vectors $\{\hat{\boldsymbol{a}}_m\}_{m=1}^{\hat{M}}$, where $\hat{\boldsymbol{a}}_m = [\cos(\hat{\varphi}_m)\sin(\hat{\varphi}_m)]^T$. The hat sign introduced here is used to denote estimate of the related quantity. Given statements are based on the fact that mixing vectors are actually cluster centers (see description given previously for the application of K-means data clustering algorithm to estimation of the mixing matrix). Parameterization of the mixing vector in terms of mixing angle helps to determine the positions of local maxima of the clustering function (11). In order for this approach to estimation of mixing matrix and number of pure components to work it is understood that: (i) all pure components are presents in each mixture in some amount; (ii) there are no two pure components with the same concentration profiles because in such situation the two corresponding mixing vectors would be parallel and the two pure components would be indistinguishable i.e. they would be represented by their linear combination.

Hence, at the end of data clustering phase estimated number of pure components $\hat{M}$ is obtained. This is an important contribution because estimation of the number of pure components is very complex issue and it is related to what in computer science is known as intrinsic dimensionality problem [38]. Several methods for estimating the number of pure components exist [39–41]. How-

ever, all of them assume $N \geq M$. Thus, they are not (at least directly) applicable to uBSS problem considered here. To estimate number of pure components robustly we used the root-mean-squared-error (RMSE) criterion between original and reconstructed data [9,10,12]

$$\mathrm{RMSE}(\hat{M}) = \sqrt{\frac{\sum_{n=1}^{N}\sum_{t=1}^{T}\left(x_n(t) - \sum_{m=1}^{\hat{M}}\hat{a}_{nm}\hat{s}_m(t)\right)^2}{NT}} \tag{12}$$

where $\hat{a}_{nm}$ denotes estimated coefficient of the mixing matrix and $\hat{s}_m$ denotes estimated pure component. Thus, by slightly varying dispersion parameter σ in (11) we obtain different values for the estimated number of pure components. We have to allow that some of them will not correspond with the true components but can be outliers caused by chemical noise or other types of imperfections that exist in experimental world. Hence, we propose information-theoretic criteria called negentropy [15], to measure information content of to be estimated pure components and rank them according to estimated negentropy measure. Negentropy is differential entropy defined relatively to the entropy of the Gaussian process. Approximation of negentropy for random process x is obtained as

$$J(x) \approx \frac{(C_3(x))^2}{12} + \frac{(C_4(x))^2}{48} \tag{13}$$

where $C_3(x)$ and $C_4(x)$ are third order and fourth order cumulants of the random process x [42]. Because we shall calculate negentropy of the magnitude spectra of the estimated pure components in frequency domain we use the definition for the cumulants for the non-zero mean random process x

$$\begin{aligned} C_3(x) &= E[x^3] - 3E[x]E[x^2] + 2E^3[x] \\ C_4(x) &= E[x^4] - 4E[x]E[x^3] + 12E^2[x]E[x^2] - 6E^4[x] \end{aligned} \tag{14}$$

where $E[x]$ in (14) denotes mathematical expectation of x. The Gaussian random process is the least informative among the random processes with unbounded support and has highest entropy. Hence, random processes that are informative are non-Gaussian with the non-zero negentropy measure. We intuitively expect that pure components are informative. Thus, the estimated pure components that correspond to the true pure components are expected to have significantly larger negentropy than the negentropy of the outliers.

2.4. Linear programming-based solution of the underdetermined system of linear equations

Unlike the case of even- or over-determined BSS problems that are solved by finding the mixing matrix **A** by means of ICA algorithms, solution of the uBSS problems is considerably more difficult. The reason is that even when **A** is known solution of the linear system of Eqs. (1) or (7) is not unique, because there are more unknowns (M) than equations (N). If pure components are $(M-N+1)$-sparse, a unique solution is obtained at the minimum of the ℓ_1 norm of **s** [24–27,30–34]. We could formulate

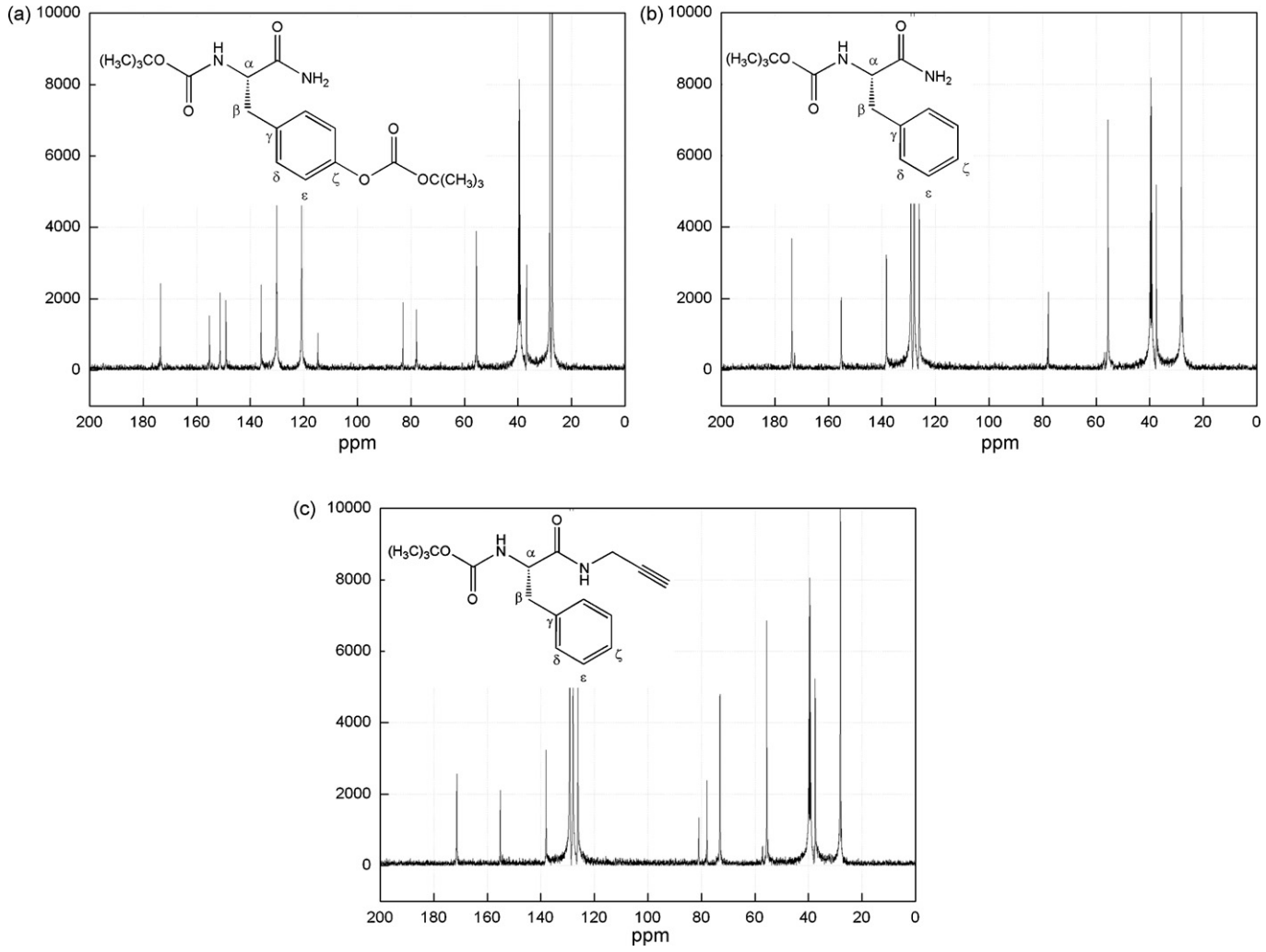

Fig. 2. (a–c) ^{13}C NMR magnitude spectra of the true pure components.

linear programming-based solution in the time-scale basis (7). However, the results of the NMR data analysis are customary presented in frequency domain in which case the pure components estimated in the time-scale basis ought to be inverse-transformed back to time domain and then to frequency domain. In order to reduce computational complexity of the proposed blind spectra decomposition algorithm we take advantage of the result presented in [30] and estimate the pure components directly in frequency domain. The result in [30] states that minimum of the ℓ_1 norm yields accurate solution of the uBSS problem even if pure components are $(M-N)$-sparse. It means that it is allowed that two pure components (or N pure components in the general case) can co-exist at each frequency. We now write model (1) in frequency domain because it is of actual interest in the NMR data analysis. Provided that at the majority of frequencies only one pure component is present the following relation between magnitude spectrum of the mixtures and pure components holds

$$\boldsymbol{x}_a(\omega) = \boldsymbol{A}\boldsymbol{s}_a(\omega) \tag{15}$$

where $\mathbf{x}_a(\omega) = [|x_1(\omega)|_2 \ldots |x_N(\omega)|_2]^T$, $\mathbf{s}_a(\omega) = [|s_1(\omega)|_2 \ldots |s_M(\omega)|_2]^T$ and $|x|_2$ denotes ℓ_2 norm of 'x'. Assuming $\mathbf{A}$ is estimated in time-scale domain by means of described data clustering algorithm we obtain $\mathbf{s}_a(\omega)$ in (15) as the solution of linear program

$$\hat{\boldsymbol{s}}_a(\omega) = \underset{\boldsymbol{s}_a(\omega)}{arg\,min} \sum_{m=1}^{\hat{M}} s_m(\omega) \text{ s.t. } \hat{\boldsymbol{A}}\boldsymbol{s}_a(\omega) = \boldsymbol{x}_a(\omega) \quad \boldsymbol{s}_a(\omega) \geq \boldsymbol{0} \tag{16}$$

Problem (16) can be solved by several methods but linear programming is known to yield unique solution due to the convexity of the linear program [24,25]. Algorithms were suggested in [26,30,33,34] as substitutes for the linear programming in the case of large scale problems or when the noise can not be neglected. Representative for such a case is ℓ_1-regularized least square problem:

$$\hat{\boldsymbol{s}}_a(\omega) = \underset{\boldsymbol{s}_a(\omega)}{arg\,min} \frac{1}{2}||\hat{\boldsymbol{A}}\boldsymbol{s}_a(\omega) - \boldsymbol{x}_a(\omega)||_2^2 + \lambda||\boldsymbol{s}_a(t)||_1 \tag{17}$$

that can be solved by interior point method [33]. We have tested both linear programming method (16) and interior point method used to solve (17). The two algorithms yielded results with basically similar quality implying that noise level in de-noised experimental data was low. Therefore, reported experimental results were obtained by means of linear programming.

As explained in the concluding paragraph in Section 2.2 the assumption that only one out of M pure components is present at each frequency will probably be violated when pure components represent complex chemical compounds. In such scenarios

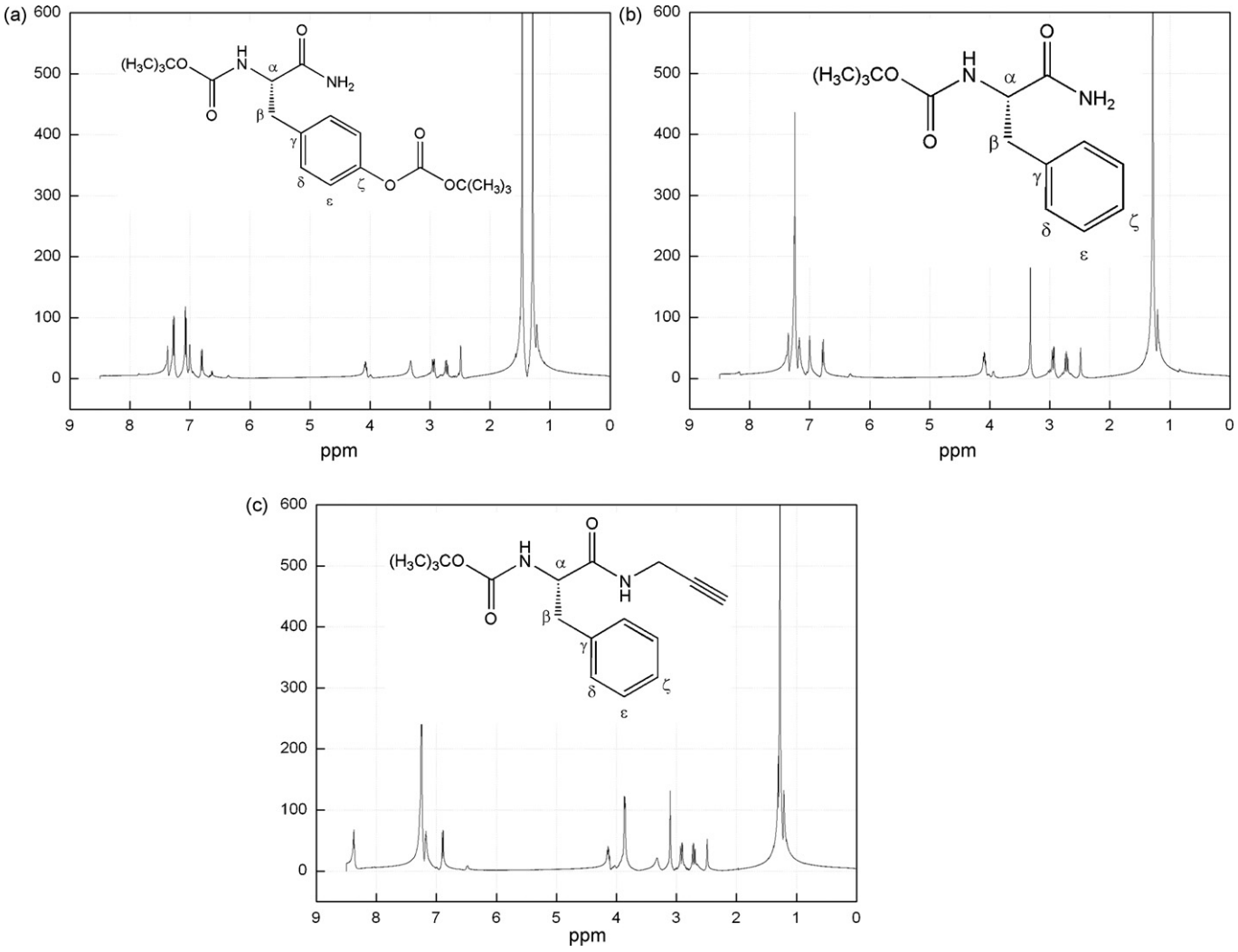

Fig. 3. (a–c) ^{1}H NMR magnitude spectra of the true pure components.

it will be necessary to increase number of mixtures N from two to three or more. This number depends on how many pure components are expected to co-exist at the majority of frequency coordinates. In such situation the relation (15) between magnitude frequency responses does not hold any more. When multiple pure components occupy each frequency point ω we notice relation between real and imaginary part of $\mathbf{x}$ as: $R\{\mathbf{x}(\omega)\} = \mathbf{A}R\{\mathbf{s}(\omega)\}$ and $I\{\mathbf{x}(\omega)\} = \mathbf{A}I\{\mathbf{s}(\omega)\}$. Written in matrix formulation it reads as

$$\begin{bmatrix} R\{\boldsymbol{x}(\omega)\} \\ I\{\boldsymbol{x}(\omega)\} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A} \end{bmatrix} \begin{bmatrix} R\{\boldsymbol{s}(\omega)\} \\ I\{\boldsymbol{s}(\omega)\} \end{bmatrix} \tag{18a}$$

or

$$\tilde{\boldsymbol{x}}(\omega) = \tilde{\boldsymbol{A}}\tilde{\boldsymbol{s}}(\omega) \tag{18b}$$

To satisfy nonnegativity constraints on variables that is required by linear program we introduce dummy variables $\mathbf{u}, \mathbf{v} \geq \mathbf{0}$ such that $\tilde{\boldsymbol{s}} = \boldsymbol{u} - \boldsymbol{v}$, $\boldsymbol{z} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix}$ and $\tilde{\boldsymbol{A}} = [\bar{\boldsymbol{A}} - \bar{\boldsymbol{A}}]$. Linear programming-based solution with equality constrains, that is equivalence of (16) when pure components do not overlap, is obtained as

$$\begin{aligned} &\hat{\boldsymbol{z}}(\omega) = \underset{\boldsymbol{z}(\omega)}{arg\,min} \sum_{m=1}^{2M} z_m(\omega) \text{ s.t.} \tilde{\boldsymbol{A}}\boldsymbol{z}(\omega) = \bar{\boldsymbol{x}}(\omega) \\ &\boldsymbol{z}(\omega) \geq \boldsymbol{0} \end{aligned} \tag{19}$$

Pure components are obtained from the solution of linear program (19) as $\tilde{\boldsymbol{s}}(\omega) = \boldsymbol{u}(\omega) - \boldsymbol{v}(\omega)$. Equivalent formulation of the noise robust solution (17) is obtained as:

$$\hat{\boldsymbol{z}}(\omega) = \underset{\boldsymbol{z}(\omega)}{arg\,min} \frac{1}{2} ||\tilde{\boldsymbol{A}}\boldsymbol{z}(\omega) - \bar{\boldsymbol{x}}(\omega)||_2^2 + \lambda ||\boldsymbol{z}(\omega)||_1 \tag{20}$$

3. Experimental

3.1. Software environment

Described SCA-based approach for blind decomposition of ^{1}H and ^{13}C NMR spectra that includes data clustering and linear programming algorithm was tested using custom scripts in MATLAB programming language (version 7.1; The MathWorks, Natick, MA). The linear programming part of the SCA algorithm has been implemented using `linprog` command from the Optimization toolbox. Continuous wavelet transform, Eqs. (4)–(7), were implemented using `cwt` command from the Wavelet toolbox. All programs were executed on desktop personal computer running under the Windows XP operating system using Intel Core 2 Quad Processor Q6600 operating with clock speed of 2.4 GHz and 4GB of RAM installed.

3.2. NMR measurements

Compounds Boc$_2$-Tyr-NH$_2$ (pure component **1**), Boc-Phe-NH$_2$ (pure component **2**) and Boc-Phe-NH–CH$_2$–C≡CH (pure component **3**) were used for the preparation of two mixtures: $\mathbf{X}_1$ (**1**:**2**:**3** = 20 mg:20 mg:7 mg) and $\mathbf{X}_2$ (**1**:**2**:**3** = 10 mg:25 mg:15 mg). Mixtures were dissolved in 600 μL of DMSO-d$_6$. NMR experiments were carried out on a Bruker AV600 spectrometer equipped with a 5 mm BBO probe with z-gradient. The liquid-state ^{1}H and ^{13}C NMR spectra (600.13 MHz for ^{1}H and 150.90 MHz for ^{13}C) were measured in DMSO-d$_6$ at 298 K using standard ^{1}H and APT techniques. Assignments of the NMR spectra of pure components **1**–**3** are given bellow.

3.2.1. Boc$_2$-Tyr-NH$_2$ (pure component **1**)

^{13}C NMR (Fig. 2a): 27.2 (CH$_3$ OBoc), 28.1 (CH$_3$ NHBoc), 36.8 (β Tyr), 55.5 (α Tyr), 77.9 (C NHBoc), 82.9 (C OBoc), 120.8 (ε Tyr), 130.1 (δ Tyr), 135.9 (γ Tyr), 149.1 151.3 (CO Boc), 155.2 (ζ Tyr), 173.5 (CO Tyr).

^{1}H NMR (Fig. 3a): 1.31 (s, 9H, CH$_3$ NHBoc), 1.48 (s, 9H, CH$_3$ OBoc), 2.74, 2.95 (dd, 2H, β,β′ Tyr, $^3J_{\alpha,\beta}$ = 10.4 Hz, $^3J_{\alpha,\beta'}$ = 4.3 Hz, $^2J_{\beta,\beta'}$ = 13.8 Hz), 4.08 (m, 1H, α Tyr), 6.81 (d, 1H, NH Tyr, $^3J_{\alpha,NH}$ = 8.9 Hz), 7.01, 7.38 (br s, 2H, NH$_2$ Tyr), 7.07 (d, 2H, ε Tyr, $^3J_{\delta,\varepsilon}$ = 8.4 Hz), 7.28 (d, 2H, δ Tyr, $^3J_{\delta,\varepsilon}$ = 8.4 Hz).

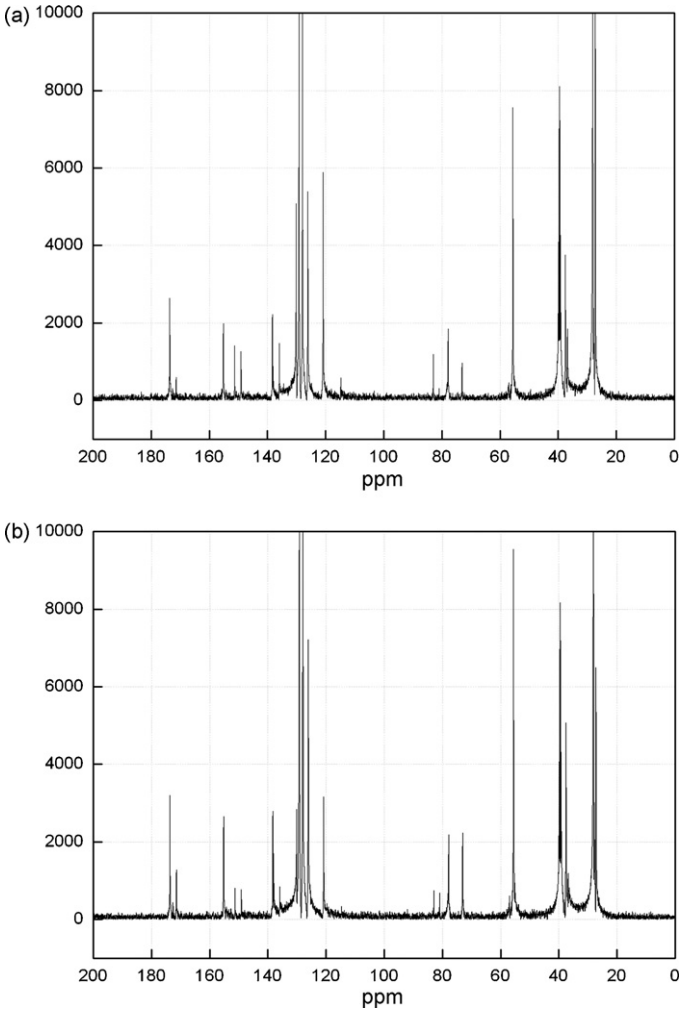

Fig. 4. (a) ^{13}C NMR magnitude spectra of the mixture X_1; (b) ^{13}C NMR magnitude spectra of the mixture X_2.

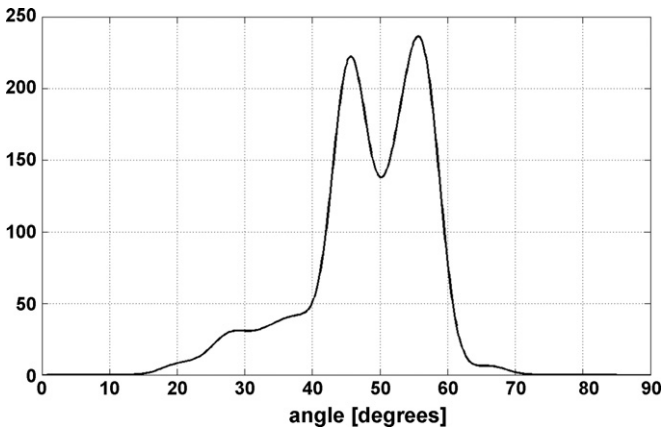

Fig. 5. Clustering function in the mixing angle domain for ^{13}C NMR mixtures.

*3.2.2. Boc-Phe-NH_2 (pure component **2**)*

^{13}C NMR (Fig. 2b): 28.1 (CH_3 Boc), 37.5 (β Phe), 55.5 (α he), 77.9 (C Boc), 126.1 (ζ Phe), 127.9 (ε Phe), 129.1 (δ Phe), 138.3 (γ Phe), 149.1, 151.3 (CO Boc), 173.6 (CO Phe).

^{1}H NMR (Fig. 3b): 1.30 (s, 9H, CH_3 Boc), 2.74, 2.96 (dd, 2H, β,β′ Phe, $^3J_{\alpha,\beta}$ = 10.3 Hz, $^3J_{\alpha,\beta'}$ = 4.1 Hz, $^2J_{\beta,\beta'}$ = 13.6 Hz), 4.10 (m, 1H, α Phe), 6.78 (d, 1H, NH Phe, $^3J_{\alpha,NH}$ = 8.7 Hz), 7.01, 7.36 (br s, 2H, NH_2 Phe), 7.25 (m, 5H, arom Phe).

*3.2.3. Boc-Phe-NH–CH_2–C≡CH (pure component **3**)*

^{13}C NMR (Fig. 2c): 28.0 (CH_2 propargyl), 28.1 (CH_3 Boc), 37.5 (β Phe), 55.5 (α Phe), 73.0 (CH propargyl), 77.9 (C Boc), 81.0 (C propargyl), 126.1 (ζ Phe), 127.9 (ε Phe), 129.2 (δ Phe), 138.0 (γ Phe), 155.2 (CO Boc), 171.4 (CO Phe).

^{1}H NMR (Fig. 3c): 1.29 (s, 9H, CH_3 Boc), 2.72, 2.93 (dd, 2H, β,β′ Phe, $^3J_{\alpha,\beta}$ = 10.3 Hz, $^3J_{\alpha,\beta'}$ = 4.2 Hz, $^2J_{\beta,\beta'}$ = 13.6 Hz), 3.12 (t, 1H, CH propargyl, $^4J_{H,H}$ = 2.4 Hz), 3.87 (m, 2H, CH_2 propargyl), 4.14 (m, 1H, α Phe), 6.90 (d, 1H, NH Phe, $^3J_{\alpha,NH}$ = 9.1 Hz), 7.25 (m, 5H, arom Phe), 8.38 (t, 1H, NH propargyl, $^3J_{NH,H}$ = 5.2 Hz).

4. Results and discussion

To test the above described approach, structurally similar amino acid derivatives Boc_2-Tyr-NH_2 (pure component **1**), Boc-Phe-NH_2 (pure component **2**) and Boc-Phe-NH–CH_2–C≡CH (pure component **3**) were chosen in this study. Although ^{13}C NMR spectra, shown in Fig. 2a–c, are relatively easy to distinguish, there are certain overlapping present in the region 20–40 ppm and 120–140 ppm (see also assignments in the Section 3.2). As clearly seen in Fig. 3a–c ^{1}H NMR spectra reflect high degree of similarity, thus providing sufficiently challenging experimental ground for mathematical algorithm. Both ^{13}C and ^{1}H NMR spectra were included in the experimental performance evaluation to demonstrate versatility of the approach proposed for blind decomposition of the NMR spectra.

4.1. Case 1: ^{13}C NMR spectra

Mathematically less demanding case of ^{13}C NMR spectra was carried out first. ^{13}C NMR spectra of the three pure components are shown in Fig. 2a–c. Fig. 4a and b show ^{13}C NMR spectra of the two mixtures, Eq. (15). Mixtures were designed to fulfill requirements important from both experimental (chemistry/spectroscopy) and mathematical point of view. First, one component (**3** in $\mathbf{X}_1$) is present in low concentration, while two are equally distributed and second, all three components are present in nearly the same concentration ($\mathbf{X}_2$). Presence of component in low concentration is often the case in NMR analysis of mixtures obtained by chemical synthesis. Separation of pure components that are present in the similar concentrations in the mixtures is a challenge for blind decomposition algorithms due to the fact that unknown concentration matrix becomes ill-conditioned. It is evident that region

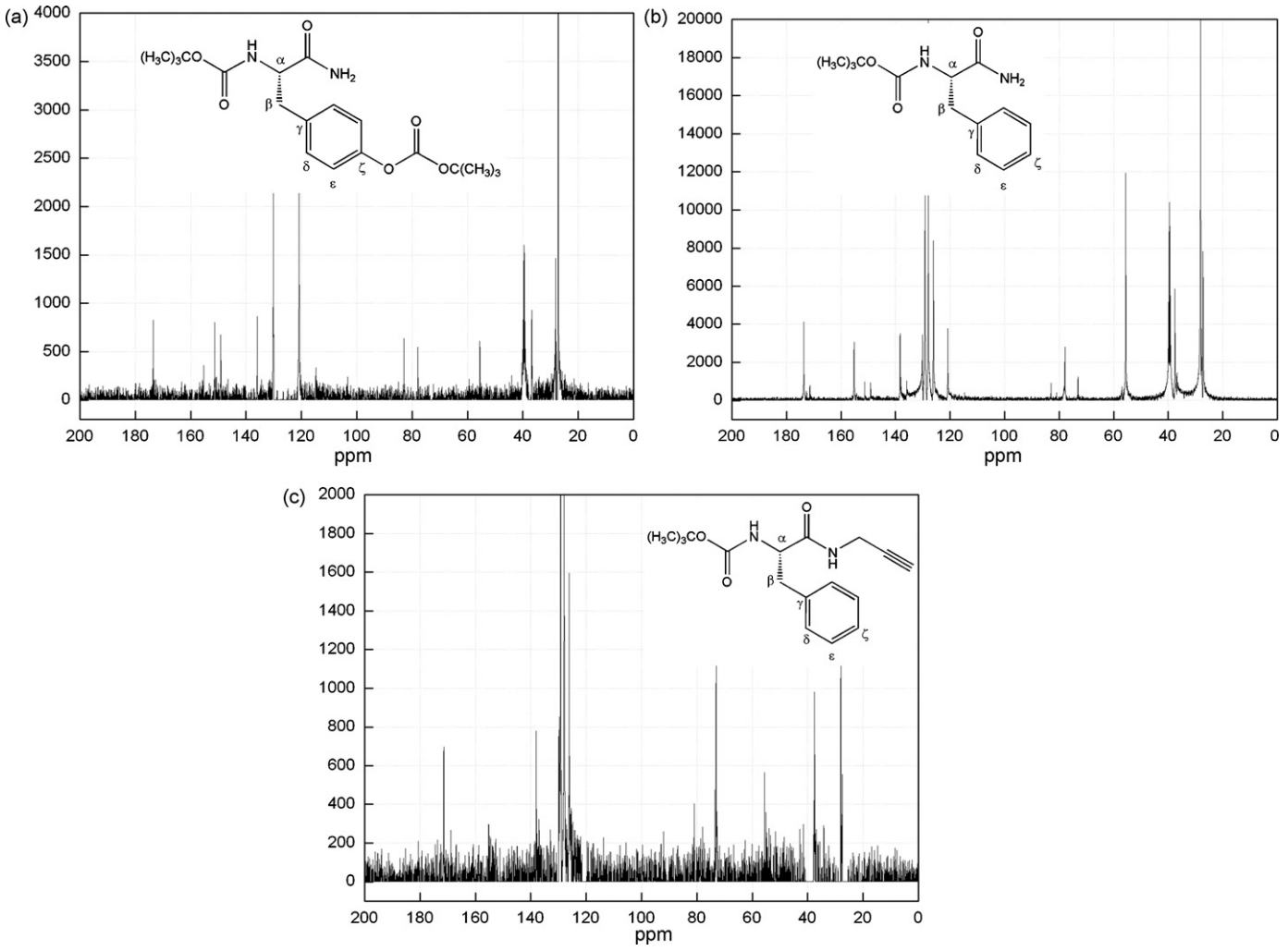

Fig. 6. ^{13}C NMR magnitude spectra of the estimated pure components: (a) PC_1; (b) PC_2; (c) PC_3.

around 30 ppm (Boc groups) and aromatic region (120–140 ppm) are the most "signal crowded" parts of the spectra. Fig. 5 shows clustering function, Eq. (11), in the mixing angle domain, wherein continuous wavelet transform, Eq. (4), with the Morlet wavelet, Eq. (5), has been used to transform mixtures from time to time-scale domain. When dispersion factor is set to $\sigma = 0.0425$ the number of the pure components is estimated as 3 with the data reconstruction error, Eq. (9), RMSE = 2.5. The clustering function shown in Fig. 5 illustrates this case. Two peaks that are well distinguished corresponds with the pure components **1** and **2** that were present in higher concentration in the mixtures, while less distinguished peak corresponds with the pure component **3** that is present in the lower concentration in the mixture. Numbers at the ordinate indicate the overall number of (a^*,b) points clustered at each of the peaks. Evidently, this number is more than four times greater for the pure components **1** and **2** than pure component **3**. The magnitude spectra of the estimated pure components that correspond to the three true pure components (Fig. 2a–c) are shown in Fig. 6a–c. Comparison with spectra of true pure components proves high degree of matching. Larger discrepancy is only found between the true third pure component, Fig. 2c, and its estimate, Fig. 6c. This is a consequence of high spectral similarity between the second and the third pure component but also of significantly lower concentration of the third pure component in the mixture $\mathbf{X}_1$. Normalized correlation coefficients between true and estimated pure components spectra are respectively given as 0.871, 0.954 and 0.819. Clearly, the accuracy of the estimation of pure components would be improved if three instead of two mixtures would be used. This would however increase the computational and experimental complexity of blind decomposition procedure. Nevertheless, quality of the results obtained from two mixtures only, as indicated by the values of normalized correlation coefficients, can be considered satisfactorily.

4.2. Case 2: ^1H NMR spectra

Blind decomposition of the ^{1}H NMR spectra has not been considered as experimental ground for BSS analysis so far, owing to the significant overlapping between the pure components spectra. Therefore, we assume that blind decomposition of more than two pure components that are structurally similar from two mixtures only should be of great practical importance in ^{1}H NMR spectroscopy. ^{1}H NMR magnitude spectra of the three pure components are shown in Fig. 3a–c. Negentropy measures, Eq. (13) and (14), calculated on the magnitude spectra of the three pure components were: 1.955×10^{17}, 2.793×10^{16} and 2.627×10^{16}. The magnitude spectra of the two mixtures, Eq. (15), are shown in Fig. 7a and b. Fig. 8 shows clustering function, Eq. (11), in the mixing angle domain wherein continuous wavelet transform, Eq. (4), with the Mexican hat wavelet, Eq. (6), has been used to transform two mixtures from time to time-scale domain. When dispersion factor is set to $\sigma = 0.04$ the number of the pure components is estimated as 4 with the data reconstruction error, Eq. (9), RMSE = 1.32×10^{-11}. When dispersion factor is set to $\sigma = 0.035$ the number of the pure components is estimated as 5 with the data reconstruction error RMSE = 8.1×10^{-13}. The clustering function shown in Fig. 8 illustrates this later case. In direct comparison with Fig. 5, that shows the clustering function for ^{13}C NMR spectra, it is evident that it becomes more difficult to distinguish peaks in the case of ^{1}H NMR spectra. The reason for this is more often violation of the $k = M - 1$ sparseness assumption on pure components in the wavelet domain i.e. ^{1}H spectra of the pure components overlap more often than ^{13}C spectra. This influences directly accuracy of data clustering methods described in Section 2.3. Again, accuracy of data clustering could be increased by using three or four mixtures instead of two only. That, however, would increase experimental and computational complexity of blind spectra decomposition process. Nevertheless,

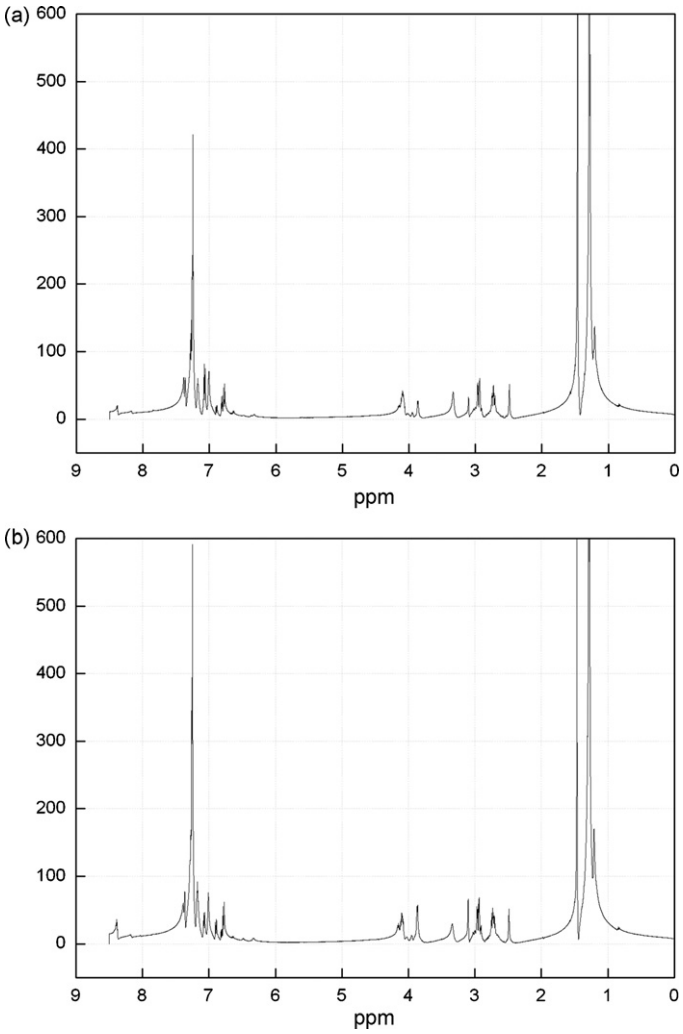

Fig. 7. ^{1}H NMR magnitude spectra of the two mixtures: (a) X_1 and (b) X_2.

results obtained for estimation of the pure components spectra, as indicated by the values of normalized correlation coefficients, can be considered satisfactorily. The magnitude spectra of the estimated pure components that correspond to the three true pure components are shown in Fig. 9a–c. Again, despite of the structural similarities and NMR peak overlapping, estimated ^{1}H spectra of three pure components exhibit good matching with the true component spectra. Normalized correlation coefficients between

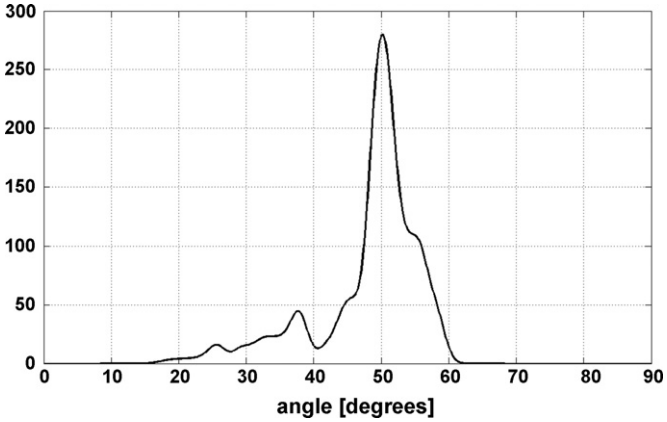

Fig. 8. Clustering function in the mixing angle domain for ^{1}H NMR mixtures.

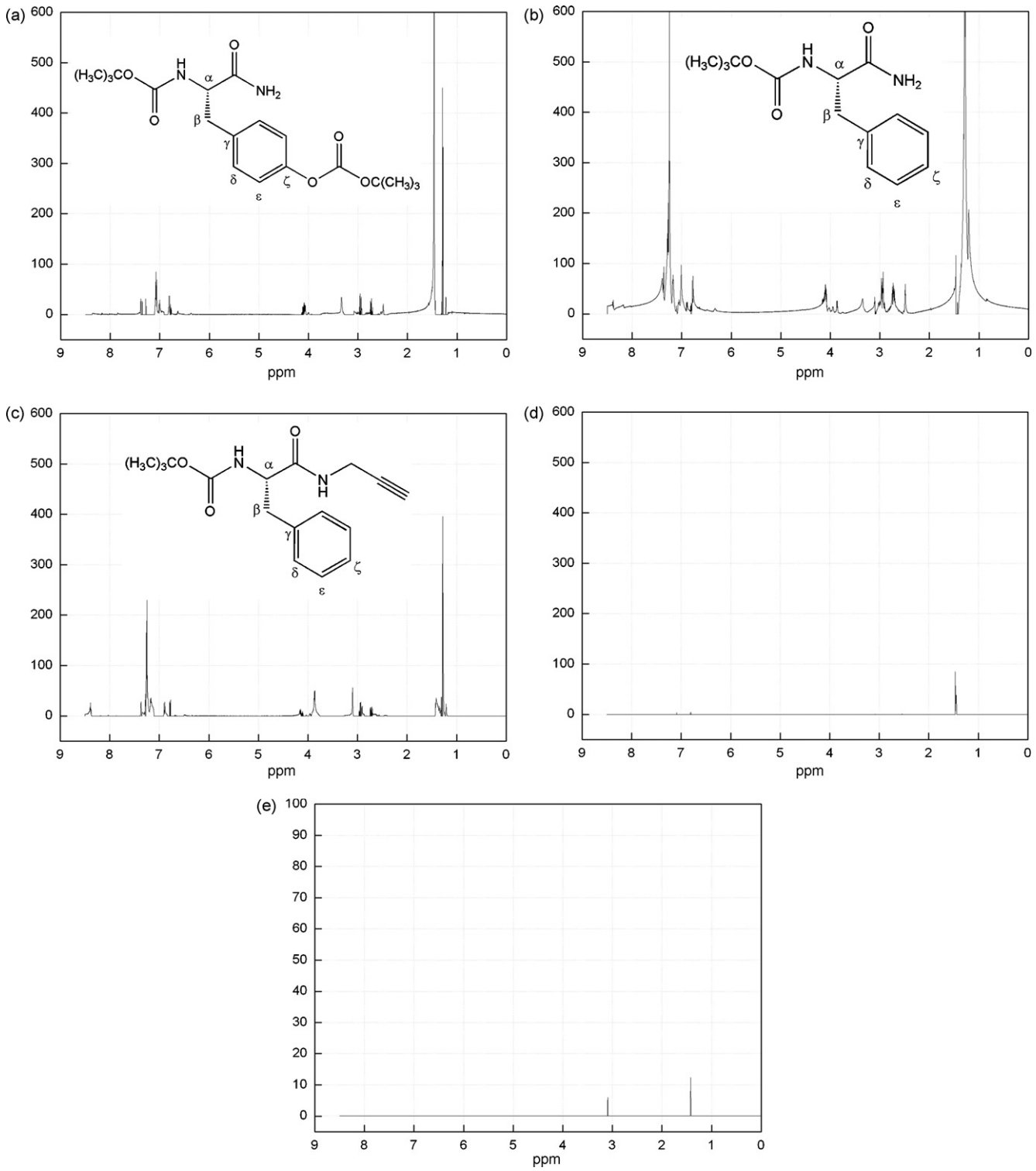

Fig. 9. ^{1}H NMR magnitude spectra of the estimated pure components: (a) PC_1; (b) PC_2; (c) PC_3 and two outliers: (d and e).

true and estimated pure components spectra are respectively given as 0.906, 0.938 and 0.818. In addition to violation of the $k = M - 1$ sparseness assumption, artifacts present in the experimental data, including chemical noise, will contribute to the presence of artificial or dummy peaks in the clustering function. This problem is especially emphasized when some of the pure components are contained in the low concentration in the mixtures. Thus it becomes very difficult to choose the proper value of dispersion constant σ in the clustering function (11) and presence of dummy peaks in the clustering function is very likely to occur. As already mentioned, such case is illustrated in Fig. 8. Therefore, selection criterion such as information-theoretic one called negentropy, is of great importance to distinguish estimated pure components that correspond to the true ones from those that correspond to artifacts. We call these later pure components the outliers. Negentropy measures, Eqs. (13) and (14), calculated on the magnitude spectra of the estimated pure components shown in Fig. 9a–c were: 1.542×10^{16}, 6.602×10^{16} and 1.379×10^{12}. Fig. 9d and e show magnitude spec-

tra of two components that are classified as outliers. As it is seen their magnitudes are between one and two orders of magnitudes smaller than magnitudes of the estimates of the true pure components. More importantly, their negentropies were: 1.536×10^6 and 1.89 and that is 10 orders of magnitude or more different that the negentropies of the estimated pure components that correspond to the true pure components. Thus, negentropy criterion can serve as a basis to discriminate estimates that correspond to the true pure components from those that ought to be classified as outliers. Note also relatively large discrepancy like in ^{13}C spectra between the third true pure component, Fig. 3c, and its estimate, Fig. 9c. This can be rationalized by the same wording like for ^{13}C spectra.

5. Conclusions

SCA-based approach has been proposed for blind extraction of *more than two* pure components spectra in ^{1}H and ^{13}C NMR spectroscopy measuring *two mixtures only*. However, presented concept is general and directly applicable to experimental scenarios that possibly would require use of more than two mixtures. This appears to be the first time to report such results, because other blind decomposition methods require the number of mixtures to be equal or greater than the unknown number of pure components. Proposed SCA-based approach solves the resulting underdetermined BSS problem by splitting it into two problems: blind estimation of the number of pure components and the mixing or concentration matrix by means of data clustering in the time-scale domain, and estimation of the magnitude pure components spectra in frequency domain by means of linear programming. This is enabled by exploiting sparseness among the pure components in time-scale domain that is achieved owing to the use of CWT with the Morlet and Mexican hat wavelets as the basis functions. To cope with the presence of outliers caused by chemical noise or other types of imperfections that exist in experimental data, information-theoretic based criteria called negenetropy has been introduced to rank the estimated pure components in term of their information content. Having in mind the importance of NMR spectroscopy for structure determination in natural products research as well as for quantification in pharmaceutical industry and medicine diagnostics, reported results present considerable contribution to further development of data analysis approaches in many areas of natural sciences.

Acknowledgment

The work of I. Kopriva, I. Jerić and V. Smrečki was respectively supported by the Ministry of Science, Education and Sports, Republic of Croatia under Grants 098-0982903-2558, 098-0982933-2936 and 098-0982929-2917.

References

[1] M. Pellecchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giralt, W. Jahnke, T.L. James, S.W. Homans, H. Kessler, C. Luchinat, B. Meyer, H. Oschkinat, J. Peng, H. Schwalbe, G. Siegal, Nat. Rev. Drug. Discov. 7 (2008) 738–745.
[2] T.N. Kolokolova, O.Y. Savel'ev, N.M. Sergeev, J. Anal. Chem. 63 (2008) 104–120.
[3] D. Nuzillard, S. Bourg, J.-M. Nuzillard, J. Magn. Reson. 133 (1998) 358–363.
[4] D. Nuzillard, J.M. Nuzillard, IEEE Sig. Proc. Lett. 5 (1998) 209–211.
[5] W. Naanaa, J.M. Nuzillard, Sig. Proc. 85 (2005) 1711–1722.
[6] J. Chen, X.Z. Wang, J. Chem. Inf. Comput. Sci. 41 (2001) 992–1001.
[7] J.Y. Ren, C.Q. Chang, P.C.W. Fung, J.G. Shen, F.H.Y. Chan, J. Magn. Reson. 166 (2004) 82–91.
[8] Ch. Chang, J. Ren, P.C. Fung, Y.S. Hung, J.G. Shen, F.H.Y. Chan, J. Magn. Reson. 175 (2005) 242–255.
[9] G. Wang, W. Cai, X. Shao, Chemom. Intell. Lab. Syst. 82 (2006) 137–144.
[10] X. Shao, G. Wang, S. Wang, Q. Su, Anal. Chem. 76 (2004) 5143–5148.
[11] V.A. Shashilov, M. Xu, V.V. Ermolenkov, I.K. Lednev, J. Quant. Spectrosc. Radiat. Transfer 102 (2006) 46–61.
[12] G. Wang, Q. Ding, Z. Hou, Trends Anal. Chem. 27 (2008) 368–376.
[13] S.A. Astakhov, H. Stögbauer, A. Kraskov, P. Grassberger, Anal. Chem. 78 (2006) 1620–1627.
[14] P. Comon, Sig. Proc. 36 (1994) 287–314.
[15] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley Interscience, 2001.
[16] A. Cichocki, S.I. Amari, Adaptive Blind Signal and Image Processing, John Wiley, New York, 2002.
[17] A. Belouchrani, K. Abed-Meraim, J.F. Cardoso, E. Moulines, IEEE Trans. Sig. Proc. 45 (1997) 434–444.
[18] J.F. Cardoso, A. Souloumiac, IEE-Proc. F. Radar Sig. Proc. 140 (1993) 362–370.
[19] Z. Koldovský, P. Tichavský, E. Oja, IEEE Trans. Neural Net. 17 (2006) 1265–1277.
[20] A. Bell, T.J. Sejnowski, Neural Comput. 7 (1995) 1129–1159.
[21] A. Hyvärinen, Proceeding of the International Conference on Artificial Neural Networks (ICANN'98), 1998, pp. 541–546.
[22] I. Kopriva, D. Seršić, Neurocomputing 71 (2008) 1642–1655.
[23] A. Cichocki, P. Georgiev, IEICE Trans. Fundam. Electron. Commun. Comput. Sci. E86-A (2003) 522–531.
[24] Y. Li, A. Cichocki, S. Amari, Neural Comput. 16 (2004) 1193–1234.
[25] P. Georgiev, F. Theis, A. Cichocki, IEEE Trans. Neural Net. 16 (2005) 992–996.
[26] P. Bofill, M. Zibulevsky, Sig. Proc. 81 (2001) 2353–2362.
[27] D. Luengo, I. Santamaria, L. Vielva, Neurocomputing 69 (2005) 198–215.
[28] F.M. Naini, G.H. Mohimani, M. Babaie-Zadeh, Ch. Jutten, Neurocomputing 71 (2008) 2330–2343.
[29] Y. Washizava, A. Cichocki, Proceedings of ICASSP'06, Toulouse, France, 2006, pp. 681–684.
[30] I. Takigawa, M. Kudo, J. Toyama, IEEE Trans. Sig. Proc. 52 (2004) 582–591.
[31] D.L. Donoho, M. Elad, Proc. Nat. Acad. Sci. 100 (2003) 2197–2202.
[32] J.A. Tropp, A.C. Gilbert, IEEE Trans. Inf. Theory 53 (2007) 4655–4666.
[33] S.J. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, IEEE J. Sel. Top. Sig. Proc. 1 (2007) 606–617.
[34] M.A.T. Figuiredo, R.D. Nowak, S.J. Wright, IEEE J. Sel. Top. Sig. Proc. 1 (2007) 586–597.
[35] F. Jiru, Eur. J. Radiol. 67 (2008) 202–217.
[36] P.S. Addison, The Illustrated Wavelet Transform Handbook, Taylor & Franics, 2002.
[37] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
[38] K. Fukunaga, D.R. Olsen, IEEE Trans. Comput. C-20 (1971) 176–183.
[39] E.R. Malinowski, Anal. Chem. 49 (1977) 612–617.
[40] E. Levina, A.S. Wagman, A.F. Callender, G.S. Mandair, M.D. Morris, J. Chemom. 21 (2007) 24–34.
[41] F. Westad, M. Kermit, Anal. Chim. Acta 490 (2003) 341–354.
[42] P. McCullagh, Tensor Methods in Statistics, Chapman & Hall, London, 1987.