Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 175 (2005) 242–255
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Novel sparse component analysis approach to free radical EPR spectra decomposition

Chunqi Chang [a,*], Jiyun Ren [a], Peter C.W. Fung [a,b,*], Y.S. Hung [a], J.G. Shen [b], Francis H.Y. Chan [a]

[a] Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong
[b] Division of Medical Physics, Department of Medicine, The University of Hong Kong, Pokfulam Road, Hong Kong

Received 7 January 2005; revised 11 April 2005
Available online 26 May 2005

Abstract

Free radicals play important roles in many physiological and pathological pathways in biological systems. These free radicals can be detected and quantified by their EPR spectra. The measured EPR spectra are often mixtures of pure spectra of several different free radicals and other chemicals. Blind source separation can be applied to estimate the pure spectra of interested free radicals. However, since the pure EPR spectra are often not independent of each other, the approach based on independent component analysis (ICA) cannot accurately extract the required spectra. In this paper, a novel sparse component analysis method for blind source separation, which exploits the sparsity of the EPR spectra, is presented to reliably extract the pure source spectra from their mixtures with high accuracy. This method has been applied to the analysis of EPR spectra of superoxide, hydroxyl, and nitric oxide free radicals, for both simulated data and real world ex vivo experiment. Compared to the traditional self-modeling method and our previous ICA-based blind source separation method, the proposed sparse component analysis approach gives much better results and can give perfect separation for mixtures of superoxide spectrum and hydroxyl spectrum in the ideal noise-free case. This method can also be used in other similar applications of quantitative spectroscopy analysis.

Keywords: Free radical; EPR spectroscopy; Quantitative spectroscopy analysis; Blind source separation; Sparse component analysis

1. Introduction

In biological systems, free radicals play important roles in many physiological and pathological pathways. Such free radicals can be routinely measured by the technique of electron paramagnetic resonance (EPR) spectroscopy. Spin trap agents are usually used for detecting those highly reactive organic free radicals, since such agents help to extend the half-life of those free radicals [1]. However, spin trap agents are not so ideally species-specific. For example, the widely used agent 5-(diethoxyphosphoryl)-5-methyl-1-pyrroline-*N*-oxide (DEPMPO) for superoxide $O_2^{-\cdot}$ radical will also simultaneously trap hydroxyl radical ($OH^{\cdot}$) in in vivo systems [2]. Consequently, the measured EPR spectrum using the spin trap agent DEPMPO will be a linear superposition of the spectra of DEPMPO adducts of the above two free radicals if both radicals are present, a situation common in biological systems. The superimposed spectra can be called multi-component mixtures, while each individual spectrum that constitutes the mixtures is called a component spectrum. Such superposition leads to difficulty in quantitative analysis of the EPR spectra, especially when the component spectra are overlapping, as is the case where superoxide and hydroxyl are co-existing.

* Corresponding authors. Fax: +852 2559 8738 (C. Chang), +852 2904 9443 (P.C.W. Fung).
E-mail addresses: cqchang@eee.hku.hk (C. Chang), hrspfcw@hkucc.hku.hk (P.C.W. Fung).

1090-7807/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2005.04.010

There has been much interest in numerically decomposing multi-component mixtures of EPR spectra into pure component spectra. Traditional approach for determining the component spectra tries to match the mixtures manually to the spectra of all known and possible pure components [3]. This approach is very ineffective and depends highly on the contents of the component spectra. Svistunenko et al. [4] proposed a simplified method that utilizes a set of mixtures of varying compositions, but this method still requires a priori information concerning the shapes of the component spectra. Since the component spectra may be very difficult to acquire a priori, generic species-free methods not relying on a priori knowledge of the shapes of the pure component spectra will be of great advantage. Generally, as in [4], multiple mixtures with distinct compositions are required for such methods. A representative approach was proposed in [5] where a self-modeling procedure and a procedure utilizing the symmetric property of the spectra were applied with preprocessing using the principle component analysis (PCA). PCA has been widely used in analytical spectroscopy, and it can decompose the mixtures into pure components when the peaks in each component are well separated from those of every others [6]. However, the component spectra must be linear combinations of the principle components if they have overlapping peaks, and in this case post-processing is required to get the pure spectra, as used in [5]. The self-modeling technique used in [5] was originally proposed in [7] and recently reviewed in [8]. Another post-processing approach used in [6] assumes that each component spectrum has only one peak.

The above species-free methods require that the pure spectra are either single peak or nearly symmetric or not contaminated by heavy noise. However, organic free radicals always have more complicated EPR spectra, and their concentrations in biological systems are always very low so that the spectra are quite noisy. In fact, the above methods do not work well in our experiments for EPR spectra analysis of superoxide and hydroxyl radicals.

The problem that the species-free methods aim to solve is exactly the blind source separation (BSS) problem familiar to the signal processing community. In [9] we proposed an independent component analysis (ICA)-based BSS approach and obtained promising results. A recent review for BSS and ICA can be found in [10]. Assuming the sources, i.e., the component spectra in this study, to be statistically non-Gaussian and independent, the BSS problem can be solved by ICA, which is based on higher order statistics. Representative ICA methods are JADE [11], InfoMax [12], and FastICA [13]. Alternatively, second order statistics-based methods can also be applied to solve the BSS problem if the sources are statistically uncorrelated over a set of time lags [14,15]. BSS has also been applied for MR spectra analysis, using either ICA [16–20] or second order statistics-based approach [21]. However, since in fact the component EPR spectra are neither completely independent nor completely uncorrelated, the decomposition result of the BSS approach to mixed EPR spectra analysis is not perfect even when the measured spectra are very clean, as shown in our previous studies in [9].

In this study, we propose a novel BSS method to decompose perfectly multi-component EPR spectra mixtures into pure component spectra that employs the sparsity of the component spectra. We will demonstrate its application to the experimental EPR spectral analysis of spin trapping EPR of superoxide and hydroxyl radicals in biological systems. Section 2 presents the theory of the novel BSS method. Details of the experiments are included in Section 3. In Section 4, we test the method on simulated free radical EPR spectra first and then on the experimental data. Discussions and conclusions are presented in Section 5.

2. Theory

2.1. Blind source separation for EPR spectra analysis

As formulated in [9], blind source separation can be used to extract EPR spectra from their mixtures. The blind source separation problem is to extract a number of K unknown source signals $x_1(n), x_2(n), \ldots, x_K(n)$, $n = 1, 2, \ldots, N$, which in our problem are the K unknown pure EPR spectra each sampled to N data points, from a number of M ($M \geqslant K$) known linear mixtures, which are the measured EPR spectra modeled as

$$\mathbf{y}(n) \triangleq \begin{bmatrix} y_1(n) \\ y_2(n) \\ \vdots \\ y_M(n) \end{bmatrix} = \mathbf{A}\mathbf{x}(n) \triangleq \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{MK} \end{bmatrix} \begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_K(n) \end{bmatrix}, \tag{1}$$

where the unknown M by K matrix $\mathbf{A}$ is called the mixing matrix. Our task is to find out a separating matrix $\mathbf{B}$ so that $\mathbf{z}(n) = \mathbf{B}\mathbf{y}(n)$ is an estimate of the source vector $\mathbf{x}(n)$, with preserved waveforms but possibly undetermined scales and orders. With the assumption of independence or uncorrelation (at a set of different lags) among the source signals, there are many well-known methods to solve this blind source separation problem, as described in Section 1. However, such assumption cannot be satisfied well for practical cases of EPR spectra analysis, and consequently the results are not good enough. With the observation that EPR spectra are generally sparse, here we develop a sparse component analysis approach to the blind source separation problem for EPR spectra analysis.

2.2. Sparse signal and the definition of sparsity

A signal is called sparse if it has some peaks and relatively flat area in between the peaks. For an arbitrary signal or data sequence $\mathbf{x} = [x(1)\ x(2) \cdots x(N)]$, its L_0 norm $\|\mathbf{x}\|_0$, which is equal to the number of nonzero samples in $\mathbf{x}$, can be a good indicator of the degree of sparsity. However, the function $\|\mathbf{x}\|_0$ is very sensitive to noise: even a tiny bit of noise would make all the samples nonzero, while noise cannot be avoided in practice. Therefore, we consider proposing a robust definition of sparsity for noisy signals. Since a very sparse signal has its value and hence its energy concentrated in some narrow segments of the signal, we define its *sparsity* as

$$S(\mathbf{x}) = \frac{\sqrt{1/N\sum_{n=1}^{N} x^2(n)}}{1/N\sum_{n=1}^{N}|x(n)|} = \sqrt{N}\frac{\sqrt{\sum_{n=1}^{N} x^2(n)}}{\sum_{n=1}^{N}|x(n)|} = \sqrt{N}\frac{\|\mathbf{x}\|_2}{\|\mathbf{x}\|_1}, \tag{2}$$

where $\|\mathbf{x}\|_2$ and $\|\mathbf{x}\|_1$ represent L_2 norm and L_1 norm of $\mathbf{x}$, respectively.

The sparsity defined by Eq. (2) is a ratio of L_2 norm to L_1 norm scaled by $\sqrt{N}$. Robustness is brought through introducing the L_2 norm in the definition. The sparsity will also tend to the ratio of the second order moment to the first order absolute moment with N tending to infinity when the signal is modeled as a zero mean stationary stochastic process. Such property makes the sparsity function a statistic of a signal. For a stationary signal, its statistic does not change much with the length of the signal, as long as the signal is long enough to provide a stable statistic.

Theorem 1. *The sparsity function has the following properties*:

(1) $S(\mathbf{x}) \geqslant 1$. *Equality holds if and only if* $x(1) = x(2) = \cdots = x(N)$.
(2) $S(\mathbf{x}) \leqslant \sqrt{N}$. *Equality holds if and only if there is only one nonzero term in* $[x(1), \ldots, x(N)]$.
(3) $S(c\mathbf{x}) = S(\mathbf{x})$, *where* c *is a constant*.
(4) $S(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2) \leqslant \max(S(\mathbf{x}_1), S(\mathbf{x}_2))$ *if* $\operatorname{sgn}(\alpha\beta x_1(n) x_2(n)) \geqslant 0$ *for all* n.

The proof is given in Appendix A.1. This theorem shows that the sparsity function is scale-invariant and is constrained between 1 and $\sqrt{N}$, with 1 for the case $x(1) = x(2) = \cdots = x(N)$ and $\sqrt{N}$ for the case where there is only one nonzero term in $x(1), \ldots, x(N)$. All these properties make $S(\mathbf{x})$ a good measure of sparseness of the signal $\mathbf{x}$. This definition regards the signal with prominent peaks as a sparse signal, the sharper the peak, the greater the sparsity $S(\mathbf{x})$. It can be shown that the sparsity of a statistically stationary signal with normal distribution is $\sqrt{\pi/2} = 1.2533$, while the sparsity of a signal with uniform distribution is $2/\sqrt{3} = 1.1547$. This result agrees with the fact that a normal-distribution signal is sparser than a uniform-distribution signal. As shown in Fig. 1, the sparsity of more spiky signals is even larger. The condition of $\operatorname{sgn}(\alpha\beta x_1(n) x_2(n)) \geqslant 0$ can be approximately satisfied for sparse signals, therefore Property 4 shows that the sparsity of a linear combination of two sparse signals is generally less than the sparsity of the original sparser signal, as illustrated in Fig. 1. Conceptually this property can be used to extract sparse source signals from their linear mixtures, as developed in the following Section 2.3.

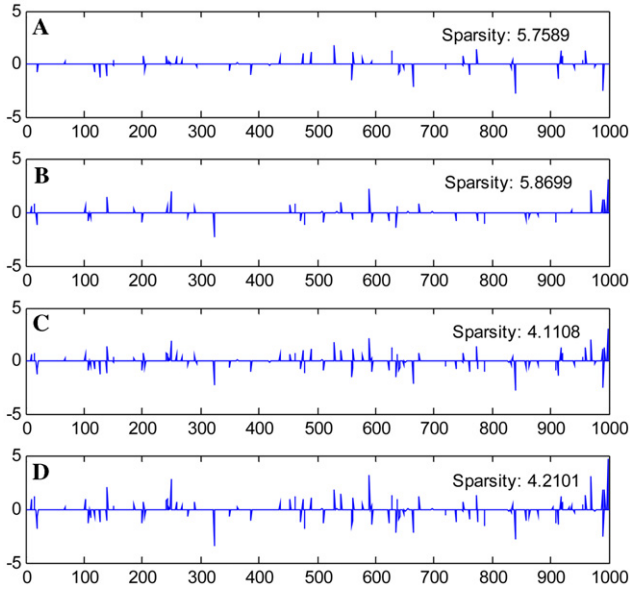

Fig. 1. (A and B) Ideal examples of sparse signal with identity on the x-axis and magnitude on the y-axis (arbitrarily scaled); (C and D) mixtures of the sparse signals in (A and B) with signal amplitude ratios of A/B equal to 1/1 in (C) and 1/1.5 in (D), respectively. The mixtures are less sparse than the source signals, as shown by their sparsity values.

2.3. Sparse component analysis approach to blind source separation

As described in Section 2.1, each source EPR spectrum can be estimated from a linear combination of the mixtures as

$$z(n) = \sum_{m=1}^{M} c_m y_m(n) = \sum_{k=1}^{K} \alpha_k x_k(n), \tag{3}$$

where $y_m(n)$ and $x_k(n)$ are the mixtures and sources defined in Eq. (1). $c_1, \ldots, c_M$ are the parameters need to be determined so that there is only one nonzero among $\alpha_1, \ldots, \alpha_K$, and denote such nonzero to be α_i without loss of generality. Therefore, we have $z(n) = \alpha_i x_i(n)$ with such a determination.

The following theorem shows that the sparsity function has some desired properties that can be used to extract simultaneously all the source signals from their linear mixtures provided that the source signals are

sparse and have some degree of non-overlapping among each other.

Theorem 2. *Let* $\Omega = \{n | x_1(n) = 0, \ x_2(n) \neq 0\}$ *be the support where* $x_2(n)$ *is not overlapping with* $x_1(n)$. *Then,* $J(t) = S(\mathbf{x}_1 + t\mathbf{x}_2)$ *has a local maximum at* $t = 0$ *if any of the following conditions is satisfied*:

(1) $x_1(n)x_2(n) = 0 \ \forall n$;
(2) $E\{x_1(n)x_2(n)\} = 0 \ \forall n$, N *tends to infinity, and* $\Omega \neq \emptyset$;
(3) $|e_\rho| < \frac{\|\mathbf{x}_1\|_2}{\sqrt{N}} S(\mathbf{x}_1) \sum_{n\in\Omega} |x_2(n)|$, *where* $e_\rho = \sum_{n=1}^{N} x_1(n) x_2(n) - S(\mathbf{x}_1) \sum_{n=1}^{N} x_1'(n) x_2(n)$ *and* $x_1'(n) = \frac{\|\mathbf{x}_1\|_2}{\sqrt{N}} \operatorname{sgn}(x_1(n))$.

The proof is given in Appendix A.2. Conceptually, we can regard the linear combination of the sources other than the desired source $\mathbf{x}_i$ as a single virtual source $\tilde{\mathbf{x}}$ such that $\sum_{k\neq i} t_k \mathbf{x}_k = t\tilde{\mathbf{x}}$, then the general case with multiple sources is a direct consequence of this theorem for the simple case with only two sources. This theorem can be used to extract the source signals from the mixtures by finding out the local maxima of the sparsity of the signal $\mathbf{z}$ defined in Eq. (3), $S(\mathbf{z};\mathbf{c})$, with respect to $\mathbf{c} = (c_1 \cdots c_M)$. From Eq. (3) we have

$$S(\mathbf{z};\mathbf{c}) = S\left(\sum_{m=1}^{M} c_m \mathbf{y}_m\right) = S\left(\alpha_i \left(\mathbf{x}_i + \sum_{k\neq i} \frac{\alpha_k}{\alpha_i} \mathbf{x}_k\right)\right) = S\left(\mathbf{x}_i + \sum_{k\neq i} \frac{\alpha_k}{\alpha_i} \mathbf{x}_k\right). \tag{4}$$

According to this theorem, the source signals can be obtained from the local maxima of the function $J(\mathbf{c}) = S(\mathbf{z};\mathbf{c})$. In this paper most examples will be presented for the simple case with two sources.

Comparing to the blind source separation methods based on independent component analysis (ICA), which makes the separated signals as independent of each other as possible, this approach tries to make the separated signals as sparse as possible, thus it can be named as sparse component analysis. In the literature there are several methods for independent component analysis that make use of the sparse property of source signals [22–25]. However, in our approach the sparsity is solely responsible for the blind signal separation, and we do not require that the source signals to be independent of each other.

2.4. Data processing procedure for EPR spectral analysis using sparse component analysis

To prepare the EPR spectrum mixtures for further spectra separation using the proposed sparse component analysis approach, the following preprocessing steps should be performed first:

(1) To make the processing simpler and better conditioned, every data value of the spectrum is subtracted from its mean value and normalized to its standard deviation so that it has zero mean and unit energy.
(2) To remove the peak shifts which often occur in EPR signal monitoring, peak alignment is achieved by evaluating the correlation coefficient [9,26]. The target spectrum is the spectrum with a relatively accurate g value according to literatures [27,28].

A source separation procedure can be obtained by maximizing the sparsity function in Eq. (4) with respect to the parameter vector $\mathbf{c} = (c_1 \cdots c_M)$. Each local maximum gives an estimate of the pure EPR source spectrum. The procedure is detailed as follows:

(1) Select a number of M detected mixture signals denote by $y_1(n), \ y_2(n), \ldots, y_M(n)$, with each of them composed of the same K components but in distinct component ratio (according to simulation control or experimental control).
(2) Since the scale of the source signal cannot be determined, we construct the estimated source as a linear combination of the mixtures with c_1 fixed to unity:
$$z(n) = y_1(n) + \sum_{m=2}^{M} c_m y_m(n). \tag{5}$$
(3) Compute the sparsity of $\mathbf{z}$ using the objective function:
$$J(\mathbf{c}) = S(\mathbf{z};\mathbf{c}) = \sqrt{N} \frac{\|\mathbf{z}\|_2}{\|\mathbf{z}\|_1}. \tag{6}$$
Notice that with $y_1(n), y_2(n), \ldots, y_M(n)$ provided, this is a function with the unknown variables $c_2, \ldots, c_M$.
(4) Determine $\mathbf{c} = \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K$ so that $J(\mathbf{c})$ achieves local maxima with respect to $\mathbf{c}$.
(5) Reconstruct the source EPR spectra using Eq. (5) with the determined $\mathbf{c} = \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K$, where $\mathbf{c}_k = (1 \quad c_{k,2} \cdots c_{k,M})$, thus we get $z_k(n) = y_1(n) + \sum_{m=2}^{M} c_{k,m} y_m(n)$ for $k = 1, \ldots, K$.

For $M = 2$ and $K = 2$ where we have only two mixtures and two sources, Eq. (6) becomes a function of a single variable, thus its local maxima can be easily found by a simple search along a single dimension. Therefore in the following section most results will be given for this simple case for demonstration purpose. For the case of $M = 3$, Eq. (6) becomes a surface in the 3D space, and such case is still tractable. However, when $M > 3$, Eq. (6) becomes a hyper surface in a multi-dimensional space. In this case, a numerical optimization algorithm is required to find the local maxima of $J(\mathbf{c})$.

3. Experimental

3.1. Simulation of some EPR spectra

Some EPR spectra as described in [29] were simulated using the PEST EPR simulation software WinSim [30].

(1) Spectrum for DMPO/·Ot-Bu: *t*-butoxyl (47%; $a^{N} = 14.9$ G, and $a^{H}\beta = 16.15$ G), with parameters: line width, 0.82 G; line shape, 19% Lorentzian, 81% Gaussian (Fig. 1B in [29]).
(2) Spectrum for DMPO/·OCH_3: methyl alkoxyl (36%; $\alpha^{N} = 14.45$ G, $\alpha^{H}\beta = 10.75$ G and $\alpha^{H}\gamma = 1.35$ G), with parameters: line width, 0.64 G; line shape, 3% Lorentzian, 97% Gaussian (Fig. 1C in [29]).
(3) Spectrum for DMPO/·CH_3: methyl radical (17%; $a^{N} = 16.25$ G, and $a^{H}\beta = 23.35$ G), with parameters: line width, 0.87 G; line shape, 65% Lorentzian, 35% Gaussian (Fig. 1D in [29]).
(4) Spectrum for MNP/·CHRR′ (where *R* and *R*′ denote two different substituent alkyl groups): ·CHRR′($a^{N} = 15.72$ G and $a^{H}\beta = 1.68$ G (1H)), with parameters: line width, 0.82 G; line shape, 3% Lorentzian, 97% Gaussian (Fig. 4G in [29]).

Other two virtual EPR source spectra are simulated with the first derivative of the Lorentzian absorption lines. These six EPR source spectra are shown in Fig. 2.

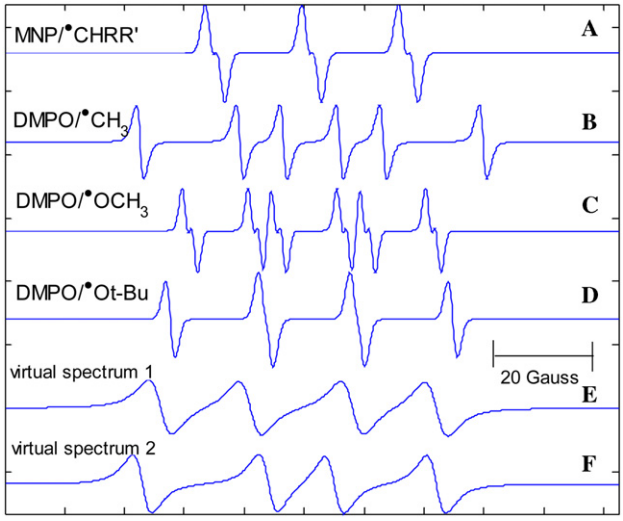

Fig. 2. (A–D) Computer simulated spectra refer to [29]. The hyperfine values used for each species are described in Experimental. (A) MNP/·CHRR′, simulation parameters: line width, 0.82 G; line shape, 3% Lorentzian, 97% Gaussian. (B) DMPO/·CH_3: line width, 0.87 G; line shape, 65% Lorentzian, 35% Gaussian. (C) DMPO/·OCH_3: line width, 0.64 G; line shape, 3% Lorentzian, 97% Gaussian. (D) DMPO/·Ot-Bu: line width, 0.82 G; line shape, 19% Lorentzian, 81% Gaussian. (E and F) Simulated virtual EPR spectra as linear superposition of the first derivatives of four Lorentzian lines.

3.2. Generation, spin trapping, and EPR spectroscopy of superoxide and hydroxyl radicals from chemical systems

The chemical system for $O_2^{-\cdot}$ production is the xanthine/xanthine oxidase system [31], comprised with 0.32 mM xanthine, 9×10^{-3} U/ml of xanthine oxidase, and 20 mM of DEPMPO in 1× PBS solution at pH 7.4. DEPMPO is short for spin trapping agent 5-diethoxyphosphoryl-5-methyl-1-pyrroline *N*-oxide, purchased from the Oxis International. The reaction system for $OH^{\cdot}$ production is the Fenton reaction system [31] consisting of 0.18 mM hydrogen peroxide, 0.09 mM $FeCl_2$, and 20 mM DEPMPO. The spectra of DEPMPO spin adduct were recorded with a Bruker EMX EPR spectrometer. Some representative spectroscopic parameters were: center field 3484 G, microwave frequency 9.76 GHz, microwave power 20 mW, modulation frequency 100 KHz, modulation amplitude 2 G, and time constant 10 ms.

3.3. EPR measurement of nitric oxide radical in rat kidney

All animal experiments were performed in accordance with the guidelines set by the University of Hong Kong Committee for Animal Experimentation. Adult male Sprague–Dawley rats, weighing 220–240 g, were used. The organic nitric oxide level in rat kidneys was detected by the spin trapping EPR method. To detect NO signal in the rats, the spin trap agents diethyldithiocarbamate (DETC, 500 mg/kg i.p.) (Aldrich), ferrous sulphate (50 mg/kg s.c.), and sodium citrate (250 mg/kg s.c.) were administrated 30 min before the sacrifice of the rats. To make NO level different in the rats, L-NAME (30 mg/kg) was injected by i.p. 30 min prior to the ischemia in a few rats. After anesthesia with ketamine (35 mg/kg i.m.) and xylazine (18 mg/kg i.m.), with a midline incision, the abdomen was opened, and the left nephrotomy was performed. The kidney tissue was then immediately cut into small cylinders and put into an EPR tube stored in liquid nitrogen for EPR detection. Measurements were performed with an ESP 300E spectrometer (Bruker) operating at X-band, 77 K, microwave frequency about 9.75 GHz, magnetic field range from 3000 to 3400 G, microwave power 1.0 mW, and modulation amplitude of 5.19 G.

4. Results

4.1. Simulations with software simulated spectra

The satisfaction of the general condition $|e_\rho| < \frac{\|\mathbf{x}_1\|_2}{\sqrt{N}} S(\mathbf{x}_1) \sum_{n\in\Omega} |x_2(n)|$ in Theorem 2 depends on the waveforms of the two signals $\mathbf{x}_1$ and $\mathbf{x}_2$, especially on the correlation and overlapping between the two

signals, and the sparsity of $\mathbf{x}_1$. To evaluate the performance of our sparse component analysis approach, a series of simulations are presented in this section. Pairs of the simulated spectra in Section 3.1 are used to generate linear mixtures in Section 4.1.1,4.1.2,4.1.3 with the model

$$\begin{aligned} y_1(n) &= x_1(n) + x_2(n), \\ y_2(n) &= x_1(n) + 1.75x_2(n), \end{aligned} \tag{7}$$

while in Section 4.1.4 three sources are used with the model

$$\begin{aligned} y_1(n) &= x_1(n) + x_2(n) + x_3(n), \\ y_2(n) &= x_1(n) + 1.25x_2(n) + 1.5x_3(n), \\ y_3(n) &= x_1(n) + 1.6x_2(n) + 0.8x_3(n), \end{aligned} \tag{8}$$

where $x_1(n)$, $x_2(n)$, and $x_3(n)$ are the simulated EPR source spectra, while $y_1(n)$, $y_2(n)$, and $y_3(n)$ are the simulated mixtures. Eq. (5) is used to reconstruct source spectra from their mixtures using the sparse component analysis approach.

4.1.1. Simulations with small correlation between source spectra

The simulated spectrum for DMPO/·OCH_3 is used as $x_1(n)$, while the simulated spectrum for DMPO/·CH_3 is used as $x_2(n)$. The correlation coefficient between these two spectra is -0.0472. Before feeding to Eq. (5), each mixture is normalized to have zero mean and unit variance. The sparsity function $J(c) = S(\mathbf{z}; c)$ is shown in Fig. 3A. The function gives two local maxima which yield the correct source spectra. This result agrees with the expected, since the correlation coefficient between the source spectra is close to zero, and thus the condition (2) in Theorem 2 is approximately satisfied. White Gaussian noise at various levels is added to the mixtures, and we find that the sources can be correctly extracted with SNR $\geqslant$ 30 dB. However, when the noise level is even larger, our method fails to give correct results, as shown in Fig. 3A. These simulations show that our method is sensitive to noise, though it can give correct estimation in high SNR situations. The same conclusion also applies to the independent component analysis approach, as shown in [9].

4.1.2. Simulations with larger correlation between source spectra

The simulated spectrum for DMPO/·Ot-Bu is used as $x_1(n)$, while the simulated spectrum for DMPO/·OCH_3 is used as $x_2(n)$. The correlation coefficient between these two spectra is -0.2845. The sparsity function $J(c) = S(\mathbf{z}; c)$ is shown in Fig. 3B. Though the correlation coefficient between the source spectra is not too small so that condition (2) in Theorem 2 cannot be satisfied, condition (3) in Theorem can still be satisfied.

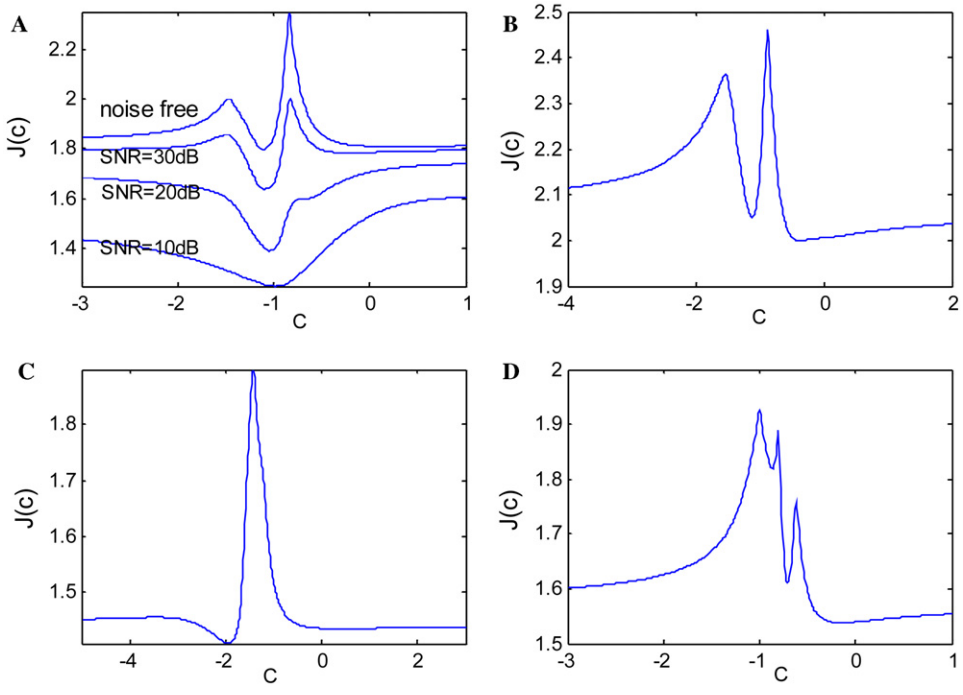

Fig. 3. Sparsity of $z(n) = y_1(n) + cy_2(n)$ for the simulated examples in Section 4.1. (A) The two sources are DMPO/·OCH_3 and DMPO/·CH_3 with correlation coefficient as -0.0472, close to zero. Noise at different levels is added to the mixtures. Results show that the sparsity decreases when noise increases, and fails to give the two exact local maxima when SNR decreases to 20 dB. (B) Two sources are DMPO/·Ot-Bu and DMPO/·OCH_3, with correlation coefficient as -0.2845. The two local maxima are clearly revealed. (C) The sources are the two virtual spectra shown in Fig. 2E and F, with correlation coefficient as 0.3085. Only one false local maximum is present due to heavy overlapping between the two sources. (D) Two mixtures and three sources. The plot gives three local maxima, indicating the existence of three sources.

This is because that each of the two spectra contains prominent peaks non-overlapping with the other, as shown in Fig. 2. Therefore, as shown in Fig. 3B, there are two local maxima in the plot, each corresponding to the estimate of a true EPR source spectrum.

4.1.3. Simulations for an unfavorable case

The two virtual EPR source spectra, simulated with the first derivative of the Lorentzian lines, are used as the two source spectra. The correlation coefficient between these two spectra is 0.3085. The sparsity function $J(c) = S(\mathbf{z}; c)$ is shown in Fig. 3C. The function has only one local maximum, and this local maximum does not correspond to any EPR source spectrum. Our method completely fails in this case. As we can see from Fig. 2, the two source spectra are completely overlapping. It is not unexpected for this failure since none of the three conditions in Theorem 2 can be satisfied, although the correlation coefficient is close to the case in Section 4.1.2.

4.1.4. Simulations with three source spectra

The simulated spectrum for DMPO/·Ot-Bu is used as $x_1(n)$, the simulated spectrum for DMPO/·CH_3 is used as $x_2(n)$, and the simulated spectrum for MNP/·CHRR′ is used as $x_3(n)$. The surface of the sparsity function $J(c_2, c_3) = S(\mathbf{z}; c_2, c_3)$ is shown in Fig. 4. Three local maxima can be clearly observed, with each yields a correct EPR source spectrum. When only two mixtures, $y_1(n)$ and $y_2(n)$, are used in Eq. (5) to extract the source spectra, the sparsity function $J(c) = S(\mathbf{z}; c)$ is shown in Fig. 3D. There are three local maxima found in the plot. Each maximum corresponds to the ratio of the portions of one particular source spectrum in the two mixtures $y_1(n)$ and $y_2(n)$. This means that the mixing model can be estimated. However, the true source spectra still cannot be extracted from the two mixtures without further known information, since the blind source problem is under determined if the number of sources is greater than the number of mixtures.

4.2. Simulations with chemically obtained spectra of superoxide and hydroxyl radical adduct

The spectra of superoxide-DEPMPO and hydroxyl-DEPMPO measured from chemical system were used as 'source spectra' or reference spectra denoted as $x_1(n)$ and $x_2(n)$, as shown in Fig. 5. Overlapping complex spectra were simulated by mixing the source spectra with different mixing coefficients. One group of the simulated mixture spectra for analysis are denoted as $y_1(n) = x_1(n) + x_2(n)$ and $y_2(n) = x_1(n) + 1.75x_2(n)$.

4.2.1. The noise free case and the problem of spurious solution

The noise free mixtures are presented in Fig. 6. For a range of values of c, the sparsity of $z(n) = y_1(n) + cy_2(n)$, $J(c) = S(\mathbf{z}; c)$, is shown in Fig. 7. There are three local maxima according to $c = -0.975$, -0.700, -0.570. Thus we can reconstruct three sources, as shown in Fig. 8. However, compared to the true sources in Fig. 8, only

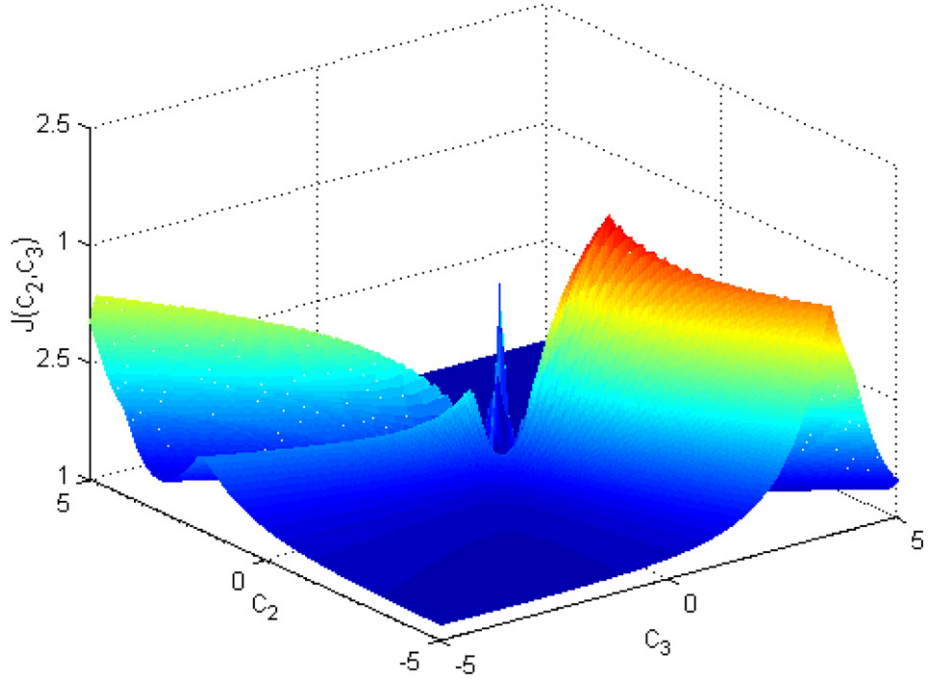

Fig. 4. 2D sparsity function for $z(n) = y_1(n) + c_2y_2(n) + c_3y_3(n)$ for the case with three mixtures and three sources, as described in Section 4.1.4. Three local maxima are demonstrated.

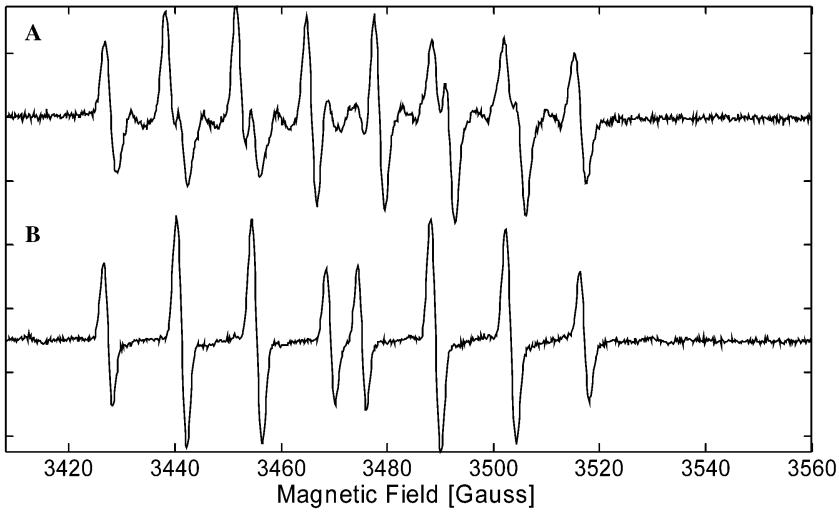

Fig. 5. EPR spectra of two typical free radicals, superoxide radical (A) obtained from xanthine/xanthine oxidase chemical system and hydroxyl (B) measured from Fenton reaction system trapped by DEPMPO. Spectra were recorded on a Bruker EMX EPR spectrometer with a microwave frequency of 9.76 GHz, microwave power of 20 mW and a modulation amplitude of 2.0 G. Other spectroscopic parameters were: center field 3484 G, modulation frequency 100 kHz and time constant 10 ms.

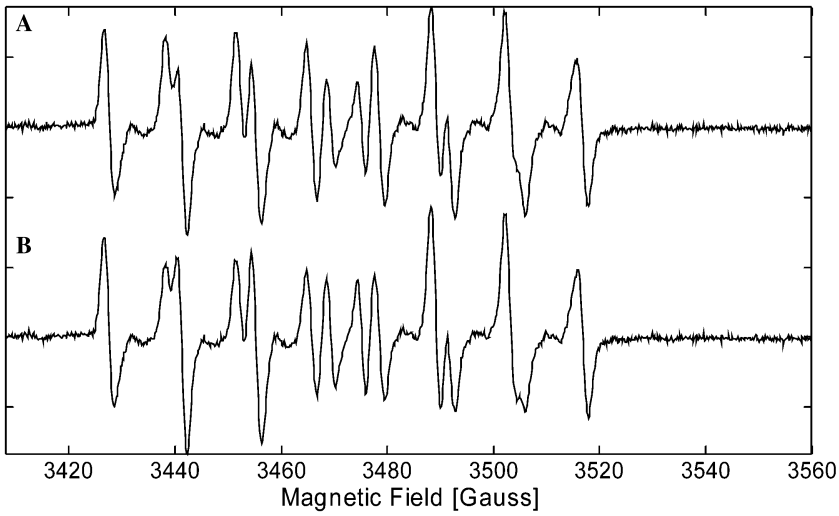

Fig. 6. Simulated mixture spectra from the source spectra of superoxide and hydroxyl radical in Fig. 5. Their proportions are 1.75/1 (A) and 1/1 (B).

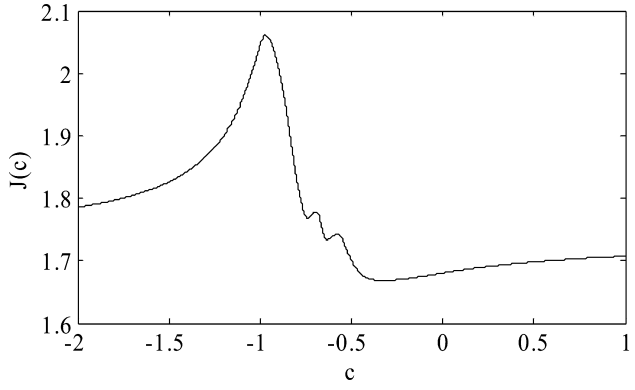

Fig. 7. Sparsity of $z(n) = y_1(n) + cy_2(n)$, with local maxima at $c = -0.975, -0.700, -0.570$. The maximum at $c = -0.700$ is a spurious solution.

the two sources corresponding to $c = -0.570$ and $c = -0.975$ are the estimate of the true sources, while the other one corresponding to $c = -0.700$ is a spurious solution being a superposition of the two true spectra.

As shown in Section 4.2, for a system with greater number of sources than mixtures, the function $J(c) = S(\mathbf{z}; c)$ may have more local maxima than the number of mixtures since the problem is undetermined. However, for a well determined blind source separation problem with equal or greater number of mixtures than sources, the reason why spurious signal may appear in the estimates is not clearly understood, and it may be complicatedly related to the waveforms of the source spectra. Nevertheless, according to Theorem 2, if the condition is satisfied, the true source spectra should be

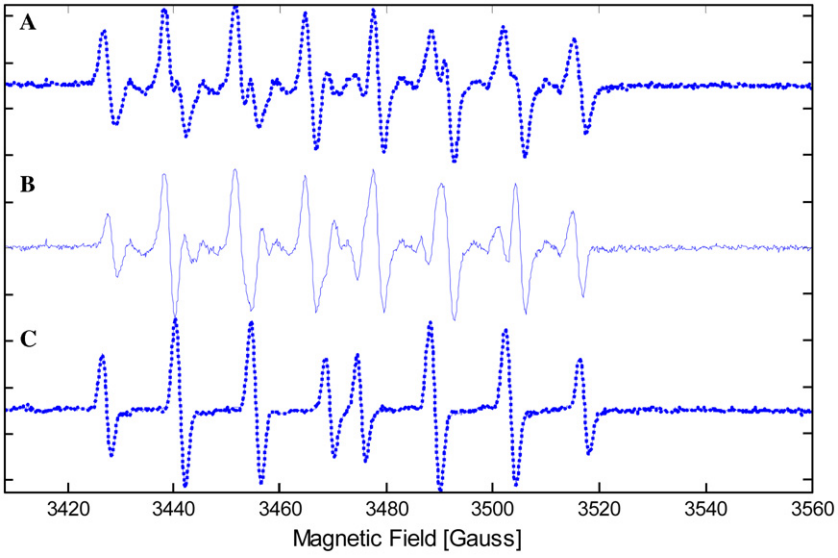

Fig. 8. (Solid lines) Separated spectra from mixtures shown in Fig. 6 using sparse component analysis approach, with $c = -0.975$ (A), $c = -0.700$ (B), and $c = -0.570$ (C). (A) is in accordance with the reference spectrum of superoxide (dotted line), (C) is in accordance with the spectrum of hydroxyl radical (dotted line), while (B) is a spurious estimation.

included in the estimates. Then our task is to select the true source spectra from the estimates. There is no general way to find out the true solutions, and any a prior information may help. One such a prior information is that the source spectra are somewhat independent of each other, as assumed by the independent component analysis. With this a prior information, we can compare the correlation coefficients of each pair of reconstructed sources, and the least correlated pair of reconstructed sources can be selected as the true estimates.

For this model system, the two sources corresponding to $c = -0.570$ and $c = -0.975$ are most uncorrelated with correlation coefficient as 0.3188, compared to the other two pairs of sources with correlation coefficients as 0.4266 and 0.7213, respectively, and therefore such two sources can be selected as the true estimates. As shown in Fig. 8, these two reconstructed source spectra resemble almost exactly the corresponding reference spectra. This is a big improvement when compared to our previous study, though that result is very good already [9].

4.2.2. Simulation with noise added (SNR = 10 dB)

Figs. 9A and B show the mixture spectra with noise added to SNR = 10 dB. As shown in Section 4.1.1, our method is sensitive to noise. Correct results cannot be obtained when the SNR is as low as 10 dB. In practice, we can get multiple measurements for the mixtures and then apply principle component analysis (PCA) to get two principle components which are linear mixtures of the source spectra. The effective SNR will be increased after PCA. In this simulation, multiple measurements are simulated to get a total of 20 measurements. Then the corresponding two principle components are used as two virtual mixtures for sparse component analysis. The sparsity function $J(c) = S(\mathbf{z}; c)$ is shown in Fig. 10. It is not surprising that there are three local maxima in the plot. The methodology for handling spurious solutions described in Section 4.2.1 is also applied here, and the two estimates of the true source spectra are determined, as shown in Figs. 9C and D. Notice that the peaks of the reconstructed spectra are quite discernible, and closely resemble the corresponding reference spectra of superoxide–DEPMPO and hydroxyl–DEPMPO.

4.3. Demonstration with experimental spectra: overlapping spectra of nitric oxide signal from ex vivo rat kidneys

Figs. 11A and B demonstrate two EPR spectra detected from the kidneys of a rat that has received an intraperitoneal injection of DETC 30 min before sacrifice, while the rat in Fig. 8A was with an L-NAME pretreatment which could inhibit NO release. Because the NO concentration was low, much of the spin trap DETC was bound to the copper and a few other sources such as the reduced iron–sulfur proteins, and formed a complex background signal superimposing on the NO EPR signal. Therefore, the EPR signals of $DETC_2$–Fe^{2+}–NO present only a doublet line in Figs. 11A and B. The sparse component analysis approach to blind source separation is applied to estimate the pure EPR spectra from the measured mixtures. As shown in Fig. 12, the sparsity function $J(c) = S(\mathbf{z}; c)$ has only two maxima, with the corresponding estimates shown in Figs. 11C and D. Notice that the distinct triplet hyperfine structure of $DETC_2$–Fe^{2+}–NO is revealed. The appearance of the

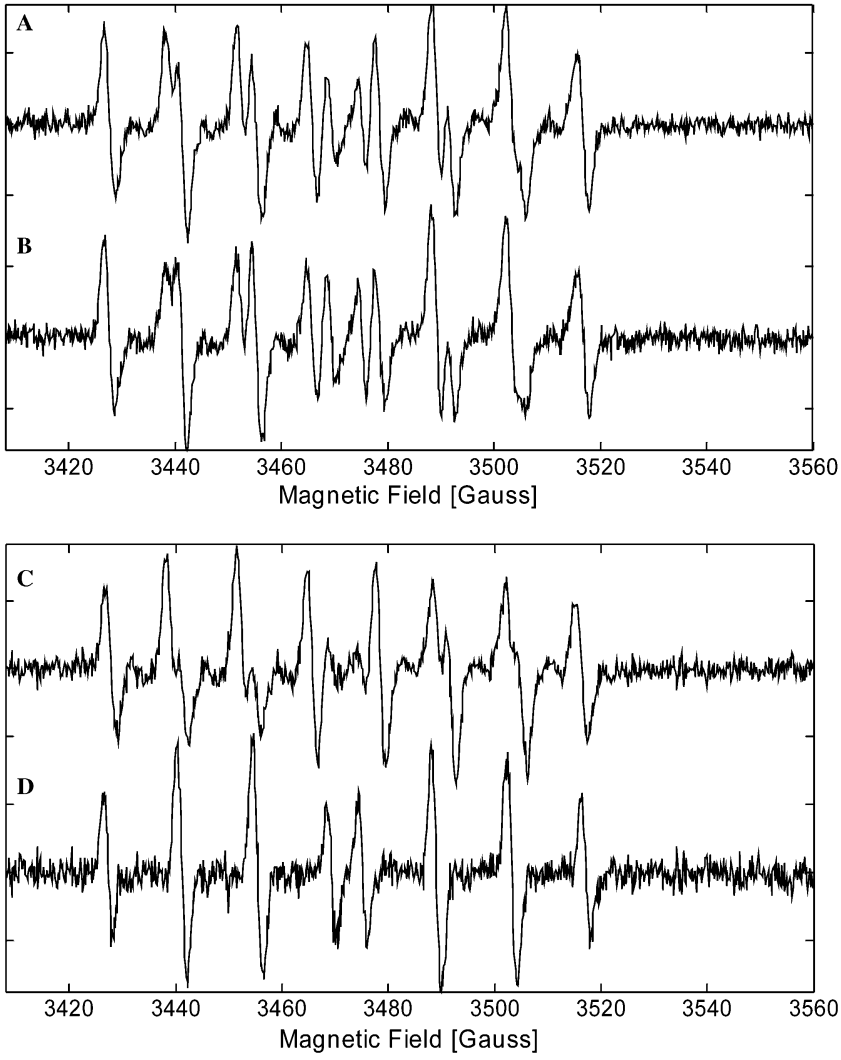

Fig. 9. (A and B) Simulated mixture spectra with random noise added for a SNR of 10 dB; (C and D) Separated signals from mixtures A and B, using sparse component analysis approach.

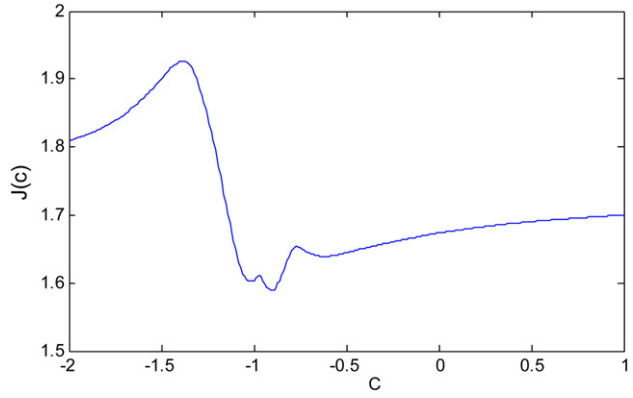

Fig. 10. Sparsity of $z(n) = y_1(n) + cy_2(n)$, where $y_1(n)$ and $y_2(n)$ are the noisy signals shown in Figs. 9A and B. Three local maxima are revealed again as in Fig. 7 with the middle one being spurious.

spectrum in Fig. 11C is mainly due to the contribution of Cu^{2+}–DETC and reduced iron–sulfur proteins, as shown by the evaluation of their g values.

Since there are three pure EPR spectra contained in the mixtures, the sparsity function $J(c_2, c_3) = S(\mathbf{z}; c_2, c_3)$ associated with three mixtures is exploited. It turns out that only two local maxima are present in this function, corresponding to two estimated EPR source spectra.

Both approaches show that only two source spectra can be extracted from the mixtures, with one reconstructed spectrum comprising the spectra of Cu^{2+}–DETC and iron–sulfur proteins. This result agrees with our previous result using a independent component analysis approach [9]. The reason is that the Cu^{2+}–DETC spectrum and the spectrum of iron–sulfur proteins might appear in close concentration proportions in all of the kidney samples, and thus a compound of the two spectra is considered as one virtual source spectrum by the blind source separation problem. This is an inherent property of blind source separation problem.

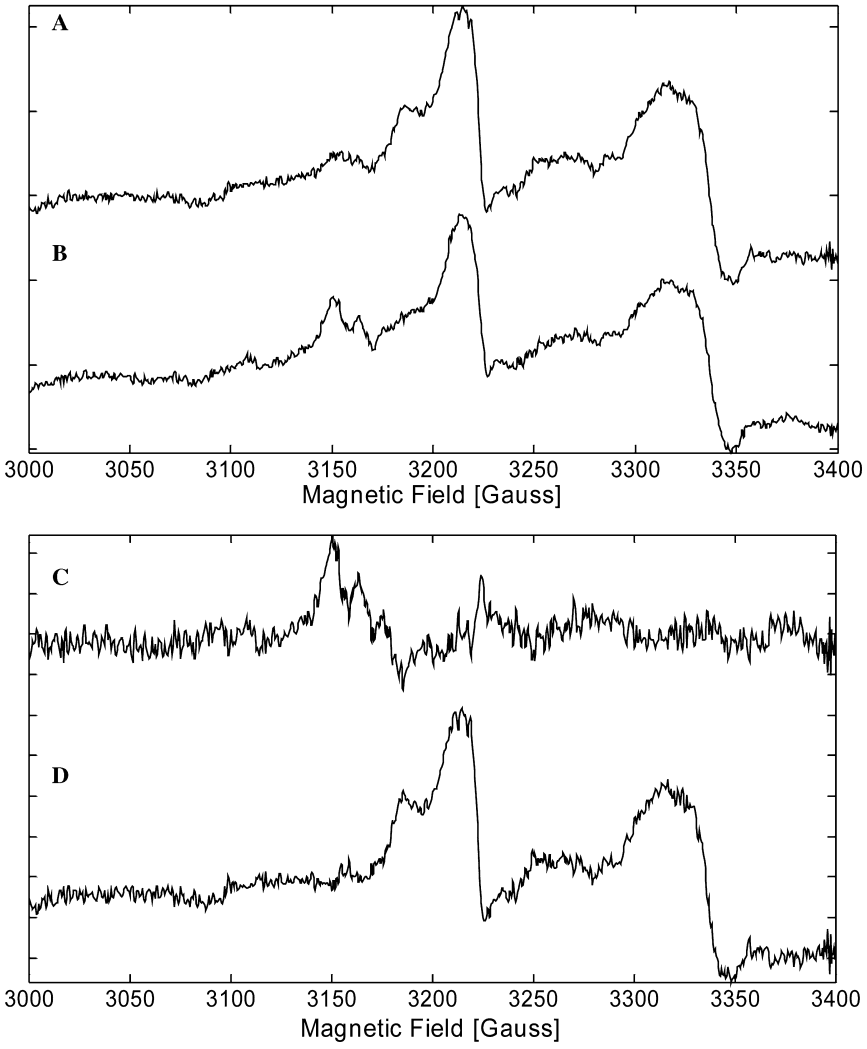

Fig. 11. (A and B) Real world overlapping EPR spectra trapping in rat kidneys labeled with Fe-DETC. NO concentrations were destined to be different in the two measurements. Signals were recorded on frozen tissues at 77 K using a Bruker EPR 300E spectrometer. Representative spectroscopic parameters were: microwave frequency: 9.45 GHz, microwave power 1.0 mW, modulation amplitude 5.19 G. (C and D) Separated signals by sparse component analysis method, in which C reveals typical triplet hyperfine structure of $DETC_2$–Fe^{2+}–NO.

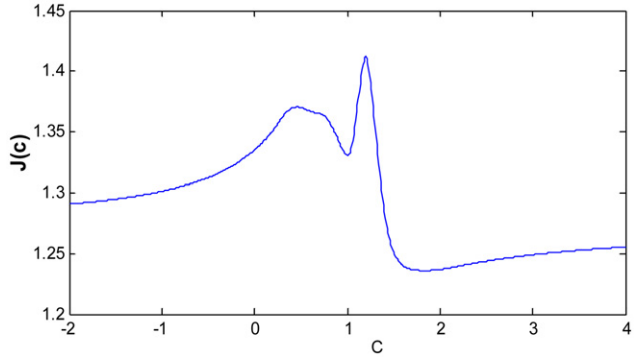

Fig. 12. Sparsity of $z(n) = y_1(n) + cy_2(n)$, where $y_1(n)$ and $y_2(n)$ are the experimental spectra shown in Figs. 11A and B. Two local maxima are revealed here.

5. Discussions and conclusions

For comparison we tested the methods of FastICA [13], PCA combined with self-modeling method [7], and PCA combined with a method employing the spectral symmetric feature [5], on the digitally mixed spectra of superoxide and hydroxyl radical as shown in Fig. 6. Codes of FastICA were obtained from public web sites [32] and were used with default setting of all parameters except for the choice of parameter '*g*' as 'tanh.' Codes of the later two methods were in MATLAB according to Steinbock et al. [5]. We find that the source spectra are retrieved to some extent from their mixtures, but there are mismatching lines between the retrieved spectra and the reference spectra, and such mismatching is most

Table 1
Correlation coefficients between separated signals and reference signals by different methods

Group	Sparse component analysis approach	FastICA	PCA combined with self-modeling method	PCA combined with symmetric method
$O_2^{-\cdot}$	0.9994	0.9612	0.9066	0.8819
$OH^{\cdot}$	1.0000	0.9975	0.9335	0.9999

significant when using self-modeling technique. The experiments demonstrate a high quality of source separation by our sparse component analysis method compared to other known techniques.

The degree of similarity between the retrieved spectra and the reference spectra for each method was measured by correlation coefficient as shown in Table 1. The correlation coefficient between the source signal $x(n)$ and the retrieved signal $z(n)$ is defined as

$$\text{Correlation coefficient} = \frac{E\{(z - E\{z\})(x - E\{x\})\}}{\sqrt{E\{(z - E\{z\})^2 (x - E\{x\})^2\}}}, \tag{9}$$

where E is the statistical expectation function. When using our novel sparse component analysis method, the average correlation coefficient is 1.0000 for superoxide-DEPMPO spectrum, and 0.9994 for hydroxyl-DEPMPO spectrum. The consistence between reference and separated spectra is the highest by our sparse component analysis method among all the methods, with FastICA ranking second. This comparison illustrates that the introduction of sparse component analysis method achieves a big improvement in solving BSS problem for the case of relative sparse signals such as the DEPMPO adduct spectra of superoxide and hydroxyl. Comparing to traditional self-modeling methods, blind source separation based on ICA (for example FastICA) gives much better separation results for mixtures of free radical EPR spectra, but the retrieved spectra may still have significant distortion even in the ideal noise-free case. Such distortion occurs because the sources are not as perfectly independent as required by the ICA approach. In fact, the correlation coefficient between superoxide-DEPMPO spectrum and hydroxyl-DEPMPO spectrum is as high as 0.3454, therefore the two spectra are far from independent. In contrast, though the correlation between source spectra violates the first two conditions in Theorem 2, condition (3) can still be satisfied since there is a high degree of non-overlapping between the two EPR spectra. Therefore, the true source spectra can still be retrieved perfectly by our sparse component analysis approach in the noise-free case.

The ideal requirements for blind source separation using sparse component analysis are that the sources are relatively sparse and have no overlapping nonzero data samples. This is the condition (1) in Theorem 2. The later requirement also assumes inherently that the sources are uncorrelated. In practical applications such requirements are most probably neither achievable nor necessary. Since according to the conditions (2) and (3) in Theorem 2, the degree of overlapping is the most important factor which determines whether the sparse component analysis approach for blind source separation can be applied effectively. The greater the sparsity of the source signals and the less the correlation among the sources, the greater the probability that the nonzero data samples of the sources are not overlapped. As shown in Fig. 5, the EPR spectra of superoxide-DEPMPO and hydroxyl-DEPMPO are much sparser than random signals with either uniform or normal distribution. Therefore, despite the strong correlation between these two EPR spectra with many overlapping peak lines, there are still some other peaks not overlapping, and such non-overlapping peaks make the sparse component analysis approach applicable. In some practical cases, only one source signal is sparse, while the others are not. Nevertheless, the sparse source can be retrieved by the sparse component analysis approach, and since one source is obtained, one can continue to use other signal processing methods to further determine other sources. Since the sparse component analysis approach favors sparse signals, a proper way to improve its efficacy is to preprocess the signals to make them sparser by making use of some appropriate mathematical transforms, as adopted in [22–24].

Acknowledgments

This work is supported in part by Hong Kong RGC grants under HKU7180/03E and N_HKU703/03, and the University of Hong Kong seed grant for medical physics. The authors thank two anonymous reviewers for helpful suggestions.

Appendix A

A.1. Proof of Theorem 1

(1) The Cebysev's inequality [33] states that if $\mathbf{a} = (a_1, \ldots, a_N)$ and $\mathbf{b} = (b_1, \ldots, b_N)$ are two non-deceasing (or non-increasing) sequences then

$\sum_{n=1}^{N} a_n b_n \geqslant \frac{1}{N}\sum_{i=1}^{N} a_n \sum_{i=1}^{N} b_n$.Substitute a_n and b_n with $|x(n)|$, we get $\|\mathbf{x}\|_2^2 \geqslant \frac{1}{N}\|\mathbf{x}\|_1^2$, and thus $S(\mathbf{x}) \geqslant 1$. If $\mathbf{x} = (c, \ldots, c)$, $S(\mathbf{x}) = \sqrt{N}\frac{c\sqrt{N}}{cN} = 1$.

(2) Since $\|\mathbf{x}\|_1^2 = (\sum_{n=1}^{N}|x(n)|)^2 = \sum_{n=1}^{N} x^2(n) + \sum_{n=1}^{N}\sum_{m\neq n}|x(n)||x(m)| \geqslant \sum_{n=1}^{N} x^2(n) = \|\mathbf{x}\|_2^2$, it is straightforward that $S(\mathbf{x}) \leqslant \sqrt{N}$.If there is only one nonzero in $[x(1), \cdots, x(N)]$, assume it to be c, then $S(\mathbf{x}) = \sqrt{N}\frac{\|\mathbf{x}\|_2}{\|\mathbf{x}\|_1} = \sqrt{N}\frac{c}{c} = \sqrt{N}$.

(3) $S(c\mathbf{x}) = \sqrt{N}\frac{\|c\mathbf{x}\|_2}{\|c\mathbf{x}\|_1} = \sqrt{N}\frac{c\|\mathbf{x}\|_2}{c\|\mathbf{x}\|_1} = S(\mathbf{x})$.

(4) Assume $S(\mathbf{x}_1) \geqslant S(\mathbf{x}_2)$, let $c = |\frac{\beta}{\alpha}|$. Due to Theorem 1.3, without loss of generality, assume $\sum_{n=1}^{N}|x_1(n)| = \sum_{n=1}^{N}|x_2(n)| = 1$. If $\operatorname{sgn}(\alpha\beta x_1(n) x_2(n)) \geqslant 0$, then

$$\begin{aligned}
S(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2) &= S\left(\mathbf{x}_1 + \frac{\beta}{\alpha}\mathbf{x}_2\right)\\
&= N\frac{\sqrt{\sum_{n=1}^{N}(x_1^2(n) + c^2x_2^2(n) + 2c|x_1(n)x_2(n)|)}}{\sum_{n=1}^{N}(|x_1(n)| + c|x_2(n)|)}\\
&= \frac{S(\mathbf{x}_1)}{1+c}\sqrt{1 + c^2\left(\frac{S(\mathbf{x}_2)}{S(\mathbf{x}_1)}\right)^2 + 2c\frac{\sum_{n=1}^{N}|x_1(n)x_2(n)|}{\sum_{n=1}^{N}x_1^2(n)}}\\
&\leqslant \frac{S(\mathbf{x}_1)}{1+c}\sqrt{1 + c^2 + c\frac{\sum_{n=1}^{N}(x_1^2(n) + x_2^2(n))}{\sum_{n=1}^{N}x_1^2(n)}}\\
&\leqslant \frac{S(\mathbf{x}_1)}{1+c}\sqrt{1 + c^2 + 2c} = S(\mathbf{x}_1)\\
&= \max(S(\mathbf{x}_1), S(\mathbf{x}_2)).
\end{aligned}$$

A.2. Proof of Theorem 2

$J_0(t) = \frac{1}{N}J^2(t) = \frac{\sum_{n=1}^{N}[x_1(n)+tx_2(n)]^2}{[\sum_{n=1}^{N}|x_1(n)+tx_2(n)|]^2} = \frac{\|\mathbf{x}_1+t\mathbf{x}_2\|_2^2}{\|\mathbf{x}_1+t\mathbf{x}_2\|_1^2}$ will have the same maxima as $J(t)$, thus in the following we analyze the property of $J_0(t)$:

$$\begin{aligned}
\frac{\partial J_0}{\partial t} &= \frac{1}{\|\mathbf{x}_1 + t\mathbf{x}_2\|_1^4}\left(\|\mathbf{x}_1 + t\mathbf{x}_2\|_1^2\frac{\partial}{\partial t}\|\mathbf{x}_1 + t\mathbf{x}_2\|_2^2\right.\\
&\quad \left. - \|\mathbf{x}_1 + t\mathbf{x}_2\|_2^2\frac{\partial}{\partial t}\|\mathbf{x}_1 + t\mathbf{x}_2\|_1^2\right),
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial t}\|\mathbf{x}_1 + t\mathbf{x}_2\|_2^2 &= 2\sum_{n=1}^{N}[x_1(n) + tx_2(n)]x_2(n)\\
&= 2\sum_{n=1}^{N}x_1(n)x_2(n) + 2t\sum_{n=1}^{N}x_2^2(n),
\end{aligned}$$

$$\frac{\partial}{\partial t}\|\mathbf{x}_1 + t\mathbf{x}_2\|_1^2 = 2\|\mathbf{x}_1 + t\mathbf{x}_2\|_1\sum_{n=1}^{N}\operatorname{sgn}(x_1(n) + tx_2(n))x_2(n).$$

Let

$$\begin{aligned}
g(t) &= \frac{1}{2}\|\mathbf{x}_1 + t\mathbf{x}_2\|_1^3\frac{\partial J_0}{\partial t} = \frac{1}{2}\|\mathbf{x}_1 + t\mathbf{x}_2\|_1\frac{\partial}{\partial t}\|\mathbf{x}_1 + t\mathbf{x}_2\|_2^2\\
&\quad - \frac{1}{2}\frac{\|\mathbf{x}_1 + t\mathbf{x}_2\|_2^2}{\|\mathbf{x}_1 + t\mathbf{x}_2\|_1}\frac{\partial}{\partial t}\|\mathbf{x}_1 + t\mathbf{x}_2\|_1^2\\
&= \|\mathbf{x}_1 + t\mathbf{x}_2\|_1\left[\sum_{n=1}^{N}x_1(n)x_2(n) + t\sum_{n=1}^{N}x_1^2(n)\right]\\
&\quad - \|\mathbf{x}_1 + t\mathbf{x}_2\|_2^2\sum_{n=1}^{N}\operatorname{sgn}(x_1(n) + tx_2(n))x_2(n).
\end{aligned}$$

(1) If $x_1(n)x_2(n) = 0\ \forall n$, then

$$\begin{aligned}
g(t\to 0^-) &\to -\|\mathbf{x}_1\|_2^2\sum_{n=1}^{N}\operatorname{sgn}(t)\operatorname{sgn}(x_2(n))x_2(n)\\
&= \|\mathbf{x}_1\|_2^2\|\mathbf{x}_2\|_1 > 0,
\end{aligned}$$

$$g(t\to 0^+) \to -\|\mathbf{x}_1\|_2^2\|\mathbf{x}_2\|_1 < 0.$$

Therefore $J(t)$ has a local maximum at $t = 0$.

(2) If $E\{x_1(n)x_2(n)\} = 0$, when N tends to infinity

$$\begin{aligned}
g(t\to 0) &\to t\|\mathbf{x}_1\|_1\|\mathbf{x}_2\|_2^2 - \|\mathbf{x}_1\|_2^2\sum_{n=1}^{N}\operatorname{sgn}(x_1(n)\\
&\quad + tx_2(n))x_2(n).
\end{aligned}$$

Since $E\{x_1(n)x_2(n)\} = 0$, $x_1(n)$ and $x_2(n)$ are independent, $x_1(n)$ and $\operatorname{sgn}(x_2(n))$ are also independent, therefore $E\{x_1(n)\operatorname{sgn}(x_2(n))\} = 0$, then we have

$$\begin{aligned}
g(t\to 0^-) &\to -\|\mathbf{x}_1\|_2^2\left[\sum_{n=1}^{N}\operatorname{sgn}(x_1(n))x_2(n)\right.\\
&\quad \left. + \sum_{n\in\Omega}\operatorname{sgn}(tx_2(n))x_2(n)\right] = \|\mathbf{x}_1\|_2^2\sum_{n\in\Omega}|x_2(n)| > 0
\end{aligned}$$

if $\Omega \neq \emptyset$.

$$\begin{aligned}
g(t\to 0^+) &\to -\|\mathbf{x}_1\|_2^2\left[\sum_{n=1}^{N}\operatorname{sgn}(x_1(n))x_2(n)\right.\\
&\quad \left. + \sum_{n\in\Omega}\operatorname{sgn}(tx_2(n))x_2(n)\right] = -\|\mathbf{x}_1\|_2^2\sum_{n\in\Omega}|x_2(n)| < 0
\end{aligned}$$

if $\Omega \neq \emptyset$.

Therefore $J(t)$ has a local maximum at $t = 0$.

(3) If $|e_\rho| < \frac{\|\mathbf{x}_1\|_2}{\sqrt{N}}S(\mathbf{x}_1)\sum_{n\in\Omega}|x_2(n)|$, that is $\|\mathbf{x}_1\|_2^2\sum_{n\in\Omega}|x_2(n)| > |e_\rho|\|\mathbf{x}_1\|_1$, since

$$\begin{aligned}
g(t\to 0) &\to \|\mathbf{x}_1\|_1\left[\sum_{n=1}^{N}x_1(n)x_2(n)\right.\\
&\quad \left. - S(\mathbf{x}_1)\sum_{n=1}^{N}x_1'(n)x_2(n)\right] - \|\mathbf{x}_1\|_2^2\sum_{n\in\Omega}\operatorname{sgn}(tx_2(n))x_2(n)\\
&= e_\rho\|\mathbf{x}_1\|_1 - \|\mathbf{x}_1\|_2^2\sum_{n\in\Omega}\operatorname{sgn}(tx_2(n))x_2(n),
\end{aligned}$$

$$g(t \to 0^-) \to e_\rho \|\mathbf{x}_1\|_1 + \|\mathbf{x}_1\|_2^2 \sum_{n\in\Omega} |x_2(n)| > 0,$$

$$g(t \to 0^+) \to e_\rho \|\mathbf{x}_1\|_1 - \|\mathbf{x}_1\|_2^2 \sum_{n\in\Omega} |x_2(n)| < 0.$$

Therefore $J(t)$ has a local maximum at $t = 0$.

References

[1] L.J. Berliner, V. Khramtsov, H. Fujii, T.L. Clanton, Unique in vivo applications of spin traps, Free Radic. Biol. Med. 30 (2001) 489–499.

[2] K.J. Liu, M. Miyake, T. Panz, H. Swartz, Evaluation of DEPMPO as a spin trapping agent in biological systems, Free Radic. Biol. Med. 26 (1999) 714–721.

[3] B. Kirste, Methods for automated analysis and simulation of electron paramagnetic resonance spectra, Anal. Chim. Acta 265 (1992) 191–200.

[4] D.A. Svistunenko, M.A. Sharpe, P. Nicholls, M.T. Wilson, C.E. Cooper, A new method for quantitation of spin concentration by EPR spectroscopy: application to methemoglobin and metmyoglobin, J. Magn. Reson. 142 (2000) 266–275.

[5] O. Steinbock, B. Neumann, B. Cage, J. Saltiel, S.C. Muller, N.S. Dalal, A demonstration of principal component analysis for EPR spectroscopy: identifying pure component spectra from complex spectra, Anal. Chem. 69 (1997) 3708–3713.

[6] R. Stoyanova, A.C. Kuesel, T.R. Brown, Application of principal-component analysis for NMR spectral quantitation, J. Magn. Reson. Ser. A 115 (1995) 265–269.

[7] W.H. Lawton, E.A. Sylvestre, Self modeling curve resolution, Technometrics 13 (1971) 617–633.

[8] J.-H. Jiang, Y. Ozaki, Self-modeling curve resolution (SMCR): principles, techniques, and applications, Appl. Spectrosc. Rev. 37 (2002) 321–345.

[9] J.Y. Ren, C.Q. Chang, P.C. Fung, J.G. Shen, F.H. Chan, Free radical EPR spectroscopy analysis using blind source separation, J. Magn. Reson. 166 (2004) 82–91.

[10] V.D. Sanchez A, Frontiers of research in BSS/ICA, Neurocomputing 49 (2002) 7–23.

[11] J.F. Cardoso, Higher-order contrasts for independent component analysis, Neural Comput. 11 (1999) 157–192.

[12] A.J. Bell, T.J. Sejnowski, An information–maximization approach to blind separation and blind deconvolution, Neural Comput. 7 (1995) 1129–1159.

[13] A. Hyvarinen, Fast and robust fixed-point algorithms for independent component analysis, IEEE Trans. Neural Networks 10 (1999) 626–634.

[14] L. Tong, R.-W. Liu, V.C. Soon, Y.-F. Huang, Indeterminacy and identifiability of blind identification, IEEE Trans. Circuits Syst. 38 (1991) 499–509.

[15] C.Q. Chang, Z. Ding, S.F. Yau, F.H.Y. Chan, A matrix-pencil approach to blind separation of colored nonstationary signals, IEEE Trans. Signal Process. 48 (2000) 900–907.

[16] D. Nuzillard, S. Bourg, J.-M. Nuzillard, Model-free analysis of mixtures by NMR using blind source separation, J. Magn. Reson. 133 (1998) 358–363.

[17] Y. Lee, Y. Huang, W. El-Deredy, P. Lisboa, C. Arus, P. Harris, Robust methodology for the discrimination of brain tumours from in vivo magnetic resonance spectra, IEE Proc.—Sci. Meas. Technol. 147 (2000) 309–314.

[18] J. Pulkkinen, A.-M. Hakkinen, N. Lundbom, R. Kauppinen, Y. Hiltunen, Application of independent component analysis to the proton spectroscopic imaging of human brain tumours, in: Proceedings of the 10th Annual Meeting of the ISMRM, Honolulu, USA, 2002.

[19] C. Ladroue, F.A. Howe, J.R. Griffiths, A.R. Tate, Independent component analysis for automated decomposition of in vivo magnetic resonance spectra, Magn. Reson. Med. 50 (2003) 697–703.

[20] F. Calamante, M. Morup, L.K. Hansen, Defining a local arterial input function for perfusion MRI using independent component analysis, Magn. Reson. Med. 52 (2004) 789.

[21] K. Stadlthanner, A.M. Tome, F.J. Theis, W. Gronwald, H.-R. Kalbitzer, E.W. Lang, Blind source separation of water artefacts in NMR spectra using a matrix pencil, in: Proceedings of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), Nara, Japan, 2003, pp. 167–172.

[22] M. Zibulevsky, B.A. Pearlmutter, Blind source separation by sparse decomposition in a signal dictionary, Neural Comput. 13 (2001) 863–882.

[23] P. Bofill, M. Zibulevsky, Underdetermined blind source separation using sparse representations, Signal Process. 81 (2001) 2353–2362.

[24] M.S. Lewicki, T.J. Sejnowski, Learning overcomplete representations, Neural Comput. 12 (2000) 337–365.

[25] A. Prieto, B. Prieto, C.G. Puntonet, A. Canas, P. MartinSmith, Geometric separation of linear mixtures of sources: application to speech signals, in: J.F. Cardoso, C. Jutten, P. Loubaton (Eds.), Independent Component Analysis and Signal Separation (Proc. ICA'99), 1999, pp. 295–300.

[26] V. Pravdova, B. Walczak, D.L. Massart, A comparison of two algorithms for warping of analytical signals, Anal. Chim. Acta 456 (2002) 77–92.

[27] A. Keszler, B. Kalyanaraman, N. Hogg, Comparative investigation of superoxide trapping by cyclic nitrone spin traps: the use of singular value decomposition and multiple linear regression analysis, Free Radic. Biol. Med. 35 (2003) 1149–1157.

[28] W. Zhu, P.C. Fung, The roles played by crucial free radicals like lipid free radicals, nitric oxide, and enzymes NOS and NADPH in CCl(4)-induced acute liver injury of mice, Free Radic. Biol. Med. 29 (2000) 870–880.

[29] Y.R. Chen, R.P. Mason, Mechanism in the reaction of cytochrome *c* oxidase with organic hydroperoxides: an ESR spin-trapping investigation, Biochem. J. 365 (2002) 461–469.

[30] NIH, PEST EPR simulation software WinSim, http://epr.niehs.nih.gov/pest.html.

[31] C. Frejaville, H. Karoui, B. Tuccio, F.L. Moigne, M. Culcasi, S. Pietri, 5-(Diethoxyphosphoryl)-5-methyl-1-pyrroline *N*-oxide: a new efficient phosphorylated nitrone for the in vitro and in vivo spin trapping of oxygen-centered radicals, J. Med. Chem. 38 (1995) 258–265.

[32] J. Hurri, H. Gävert, J. Särelä, A. Hyvärinen, The FastICA software package. http://www.cis.hut.fi/projects/ica/fastica.

[33] D.S. Mitrinovic, J.E. Pecaric, A.M. Fink, Classical and New Inequalities in Analysis, first ed., Kluwer Academic, Dordrecht, 1993.