

Blind separation of electron paramagnetic resonance signals using diversity minimization [☆]

Xiansheng Guo ^{a,b}, Chunqi Chang ^a, Edmund Y. Lam ^{a,*}

^a Department of Electrical and Electronic Engineering, University of Hong Kong, Pokfulam Road, Hong Kong

^b Department of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China

ARTICLE INFO

Article history:

Received 31 July 2009

Revised 23 December 2009

Available online 11 February 2010

Keywords:

Blind source separation

Non-negative sources

Diversity measures

Electron Paramagnetic Resonance

spectroscopy

ABSTRACT

This paper presents a method for the blind separation of Electron Paramagnetic Resonance (EPR) spectroscopy signals that can aid in the detection of free radicals in living organisms. Observed EPR signals are often mixtures of source signals that are approximately “sparse”, with a small number of narrow segments of the signal much larger than the remaining parts. We develop a method to separate the sources through minimizing a p -norm-like diversity measure under some mild assumptions which are generally valid for EPR signals. Simulations demonstrate that the proposed method performs well on EPR signal separation, with better robustness to noise compared to other techniques.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Electron Paramagnetic Resonance (EPR) spectroscopy is an effective method for quantifying free radicals, i.e., chemical species that have unpaired electron(s), in living organisms. Since such molecules are paramagnetic, EPR, similar to the more widely used nuclear magnetic resonance (NMR), can be used to detect the response of the unpaired electrons under a changing magnetic field. This EPR signal can then be used to quantify the amount of free radicals in human and animal tissues. More recently, the technique of EPR spectroscopy has been further developed to be a medical imaging technique called electron paramagnetic resonance imaging (EPRI) [1,2], as analog to the magnetic resonance imaging (MRI) that evolved from nuclear magnetic resonance spectroscopy.

Free radicals play important roles in living organisms. For example, the free radical nitric oxide (NO) is an important neural transmitter. More often than not, free radicals have adverse effects on living tissues. The unpaired electrons have a strong tendency to attract other electrons to pair them, and such a property makes free radicals chemically very active, leading to oxidation reactions with other molecules in the tissue. However, this should be avoided since it is a major cause of aging. Therefore, the quantification of free radicals in living tissues is important to biomedical engineering [3].

1.1. EPR signal analysis

One major difficulty in quantitative analysis of free radicals is that in living tissues there are always several different free radicals that cannot be easily measured separately. For a real EPR measurement, it is often a mixture of the signals generated from these multiple free radical species. Therefore, there is a need to develop an effective computational algorithm to estimate the individual EPR spectra of each free radical from their mixture, which is very often the only available measurement. Traditional approach matches the mixture manually to the spectra of all known and possible pure components, which is very ineffective and depends highly on the contents of the component spectra. It has been observed that multiple measurements, which contain different proportions of mixing components, can bring additional information and thus has been utilized in [4] to get improved result. However, it still requires *a priori* information concerning the shapes of the component spectra. Given multiple mixtures with distinct compositions such as the setting in [4], it is possible to separate them without prior knowledge on the shapes of the pure components using, for example, the self-modeling technique proposed in [5,6], which is basically a principal component analysis (PCA) with post-processing based on a single peak or symmetric assumption.

Furthermore, real organic free radicals always have more complicated EPR spectra, and their concentrations in biological systems are so low that the spectra are deemed quite noisy. Research has found that the self-modeling technique may not work well for separating free radicals with complicated spectra, as evidenced in the study of [7] for separating the spectra of superoxide and hydroxyl.

[☆] This work was in part supported by the University of Hong Kong under Project 10208648 at the University of Hong Kong.

* Corresponding author.

E-mail address: elam@eee.hku.hk (E.Y. Lam).

URL: <http://www.eee.hku.hk/~elam> (E.Y. Lam).

On the other hand, the same study argues that determining component EPR spectra from measurement mixtures can be considered a blind source separation (BSS) problem, and it applies some popular BSS algorithms to EPR spectra separation for the first time with reasonably good results [7]. Yet, because conventional BSS methods' assumption of independence of source signals are not strictly valid for EPR signals, the source spectra cannot be perfectly determined with established techniques. In [8], a new BSS method is developed that uses the fact that EPR spectra are generally sparse. It is demonstrated that in many cases the pure source spectra can be determined by this novel sparse component analysis method. However, a drawback is that it is not very scalable, and therefore with a large number of free radicals the computational time can be very high. Therefore, in this paper we aim to develop a new BSS method to determine the pure component spectra from EPR mixtures that is computationally scalable and does not require independence of the sources. For this purpose, note that EPR spectra are naturally sparse and non-negative, which we will elaborate further in Section 2.

1.2. Non-negative blind source separation techniques

Blind separation of non-negative sources arises in many applications of signal processing, ranging from biomedical imaging (such as electroencephalogram (EEG) [9] and magnetoencephalogram (MEG) [10]), hyperspectral imaging, to analytic chemistry [11,12]. The word “blind” refers to the fact that we do not know the mixing matrix, just as in blind deconvolution that we do not know the impulse response of a linear time-invariant system [13]. Typical methods include the non-negative independent component analysis (nICA) [14] and non-negative matrix factorization (NMF) [15]. The nICA method assumes that the source signals are non-negative and mutually independent, while NMF decomposes the observation matrix into a product of two non-negative matrices with one serving as the estimate of the sources and the other as the mixing matrix [16]. However, neither of them is suitable for EPR signal separation since nICA requires independence of source spectra while the decomposition of NMF is not unique.

With essentially the same assumption of sparse and local dominance as in [8], and non-negativity, the two algorithms named NMF with sparse constraints (NMFSC) [17] and convex analysis of mixtures of non-negative sources (CAMNS) [18] have been developed recently. By introducing the concept of an affine hull, the true source vectors can be found from its extreme points in CAMNS. Notice that NMFSC exploits only the sparseness and CAMNS only the local dominance of the source signals. In this paper we aim to develop a new method that exploits both aspects that are appropriate for EPR signals. The method basically is to minimize a diversity measure, ℓ_p norm (where $0 < p \leq 1$) of the source spectra.

1.3. Organization

This paper is organized as follows. In Section 2, the problem formulation is presented, followed by the derivation of the diversity minimization approach to BSS of non-negative sparse sources that are locally dominant. Simulation results for both randomly generated non-negative sources and real EPR signals are then presented in Section 3. Finally, some concluding remarks are given in Section 4.

2. Problem formulation

Assume that there are N EPR signal sources, denoted s_1, s_2, \dots, s_N , which are mixed together to form M mixtures, denoted y_1, y_2, \dots, y_M . If we let

$$\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_N \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix},$$

Then the two are related by the equation

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \mathbf{v}, \quad (1)$$

where \mathbf{A} is an $M \times N$ mixing matrix. The vector \mathbf{v} represents the additive random noise. Since \mathbf{A} is typically unknown, finding the EPR sources signals \mathbf{s} from the observed mixture \mathbf{y} can be considered a BSS problem described in the introduction of this paper.

Eq. (1) is in fact only an instantaneous mixture model. In reality, we need a generalized version, where each source and observed mixture is a length L signal. In that case, the length- M vector \mathbf{y} becomes a matrix \mathbf{Y} of size $M \times L$; the length- N vector \mathbf{s} becomes a matrix \mathbf{S} of size $N \times L$; and the length- M vector \mathbf{v} becomes a matrix \mathbf{V} of size $M \times L$. The mixing model is then

$$\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathbf{V}. \quad (2)$$

The small letters corresponding to these matrices will be used to denote the individual entries. For example, s_{nl} would be the n th row and l th column of the matrix \mathbf{S} , which is the value of the source s_n at time l .

In solving the BSS for the EPR signal, it is safe to make the following four assumptions:

1. There are more mixtures than sources, i.e., $M \geq N$. In other words, we have an overdetermined system of equations. Also, \mathbf{A} is of full column rank.
2. The mixing matrix has unit row sum, i.e., $\sum_{n=1}^N a_{mn} = 1$. Therefore, the average intensity of the sources is preserved in the mixtures.
3. All sources are non-negative, i.e., $s_{nl} \geq 0$.
4. The source signals to be separated satisfy local dominance [18]. In other words, for each source we can always find at least one time point with the property that only this source is large while the other sources are approximately zero. We can compute the local dominance κ_n for source s_n by

$$\kappa_n = \max_l \left(\frac{s_{nl}}{\sum_{n=1, n \neq n'}^N s_{nl}} \right). \quad (3)$$

This value should be large for all sources in our EPR signals.

Consequent to assumption (4) above, EPR source signals are often considered to be approximately sparse. We can also say that EPR signals have high contrast. In other words, many entries of s_{nl} are close to zero, and the energy of the source concentrates on a few nonzero entries. On the other hand, the data are not strictly sparse because there are often residual signals at other sources even when they are dominated by another one. As we will see more clearly in the experimental section, our method is capable of handling this “approximately sparse” BSS.

2.1. Diversity measure

Assume for the time being that we have a strictly sparse set of EPR signals. We can then retrieve the EPR sources from the mixtures by solving the following optimization problem:

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2 + \lambda E^{(0)}(\mathbf{S}) \\ & \text{subject to} && \mathbf{S} \succeq 0 \\ & && \sum_{n=1}^N a_{mn} = 1, \end{aligned} \quad (4)$$

where λ is a regularization parameter balancing data fidelity and sensitivity to noise [19], and $\|\cdot\|_F$ is the Frobenius norm [20]. The inequality \succeq denotes component-wise greater than or equal to. The function $E^{(0)}(S)$ is used to denote an “ ℓ_0 norm” of the matrix, which counts the number of nonzero terms in S .

It is known that finding a global solution to an ℓ_0 norm minimization requires a combinatorial search, which is computationally unattractive [21]. Instead, we can use an ℓ_1 norm minimization by substituting $E^{(0)}(S)$ by $E^{(1)}(S)$ in Eq. (4) and leaving all other terms intact, where

$$E^{(1)}(S) = \sum_{n=1}^N \sum_{l=1}^L |s_{nl}|.$$

The efficiency of ℓ_1 norm minimization is explored in [22], and recent research has looked into finding conditions where the substitution gives the same solutions [23].

There are however two reasons why one should also explore the ℓ_p norm, with $0 < p < 1$. First, it is differentiable everywhere while ℓ_1 norm is not differentiable at the origin. Second, it can give a more sparse solution, as p gets smaller [24]. Thus, ℓ_p norm promises to be a practical compromise between sparseness of the solution and solvability in a reasonable amount of time. Note also that the ℓ_p norm is also called a diversity measure. In this nomenclature, “diversity” refers to the lack of sparsity, or anti-sparsity. When sources are sparse, minimizing the diversity (anti-sparsity) is equivalent to maximizing the concentration (sparsity), which is the basis of compressive sensing [25–27].

Here, we define the diversity measure of S as

$$E^{(p)}(S) = \sum_{n=1}^N \sum_{l=1}^L |s_{nl}|^p. \quad (5)$$

As p approaches zero, the diversity measure provides a count of the number of nonzero entries in S . Using diversity measure minimization leads to the following optimization problem

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \|Y - AS\|_F^2 + \lambda E^{(p)}(S) \\ & \text{subject to} && S \succeq 0 \\ & && \sum_{n=1}^N a_{mn} = 1, \end{aligned} \quad (6)$$

which we will solve in the next section.

2.2. Solving the non-negative BSS

The product AS can be alternatively written as follows. Let us denote the entire length- L signal from source s_n as \mathbf{s}_n^T ; mathematically, this equals the n th row of S put in a column format. In addition, we denote the n th column of the matrix A as \mathbf{a}_n . With these, AS can be written as [20]

$$AS = \sum_{n=1}^N \mathbf{a}_n \mathbf{s}_n^T. \quad (7)$$

Based on this decomposition, we propose an alternating minimization algorithm to estimate the unknowns iteratively, one term at a time. We define the q th residual as

$$R_q = Y - \sum_{n=1, n \neq q}^N \mathbf{a}_n \mathbf{s}_n^T, \quad (8)$$

which is the part of the data Y not explained by the component $\mathbf{a}_n \mathbf{s}_n$. Thus, estimating \mathbf{s}_n assuming that A and \mathbf{s}_q (for $q \neq n$) are fixed leads to the component-wise optimization problem

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \|R_n - \mathbf{a}_n \mathbf{s}_n^T\|_F^2 + \lambda E^{(p)}(\mathbf{s}_n) \\ & \text{subject to} && \mathbf{s}_n \succeq 0 \\ & && \sum_{n=1}^N a_{mn} = 1. \end{aligned} \quad (9)$$

To solve Eq. (9), a fixed-point algorithm using alternating minimization can be derived. We set the derivative of its objective function with respect to \mathbf{s}_n to be zero, i.e.,

$$\frac{\partial}{\partial \mathbf{s}_n} \left\{ \frac{1}{2} \|R_n - \mathbf{a}_n \mathbf{s}_n^T\|_F^2 + \lambda E^{(p)}(\mathbf{s}_n) \right\} = 0. \quad (10)$$

This gives the Karush–Kuhn–Tucker (KKT) condition

$$-R_n^T \mathbf{a}_n + \|\mathbf{a}_n\|_2^2 \mathbf{s}_n + \lambda p D(\mathbf{s}_n)^{p-1} \text{sgn}(\mathbf{s}_n) = 0 \quad (11)$$

$$\mathbf{s}_n \succeq 0 \quad (12)$$

$$\sum_{n=1}^N a_{mn} = 1, \quad (13)$$

where $\text{sgn}(\cdot)$ is the signum function, and $\|\cdot\|_2$ is the ℓ_2 norm. The matrix $D(\mathbf{s}_n)$ is a diagonal matrix of size $L \times L$ where the diagonal entries are given by \mathbf{s}_n .

From Eqs. (11) and (12), we can compute \mathbf{s}_n using the following iterative procedure:

$$\mathbf{s}_n^{(k+1)} = \frac{R_n^T \mathbf{a}_n - \lambda p D(\mathbf{s}_n)^{p-1} \text{sgn}(\mathbf{s}_n)}{\|\mathbf{a}_n^{(k)}\|_2^2}, \quad (14)$$

where the superscript $k = 1, 2, \dots, K$ denotes the iteration step. After obtaining an estimate of \mathbf{s}_n , we fix $\{\mathbf{a}_q\}_{q \neq n}$ and S , and update the column \mathbf{a}_n from a least-squares estimate

$$\mathbf{a}_n^{(k+1)} = \frac{R_n}{\|\mathbf{s}_n^{(k)}\|_2^2} \mathbf{s}_n^T, \quad (15)$$

which is further scaled to meet the following constraint that

$$\sum_{n=1}^N a_{mn}^{(k+1)} = 1. \quad (16)$$

The Eqs. (14)–(16) are the three main iterative steps in our algorithm. We summarize it in Table 1.

2.3. Parameter selection in our approach

The first parameter to be determined in our algorithm is the maximum iteration number K . Empirically, picking it between 50 and 100 leads to good convergence. Beyond this point, we have not observed any substantial improvement to the source separa-

Table 1
The alternating minimization method based on diversity measure minimization.

Step (1)	Initialization
	<ul style="list-style-type: none"> • K • p • A • $S \leftarrow A^+ Y$
Step (2)	For $k = 1, 2, \dots, K$ For $n = 1, 2, \dots, N$ <ul style="list-style-type: none"> • Compute the residual R_n using Eq. (8) • Estimate \mathbf{s}_n using the iterative Eq. (14) • Estimate \mathbf{a}_n using the iterative Eq. (15) • Enforce the constraint given in Eq. (16) • If $s_n^{(k)} < 0$, then set $s_n^{(k)} = 10^{-9}$; else, go to Step 3
Step (3)	$k \leftarrow k + 1$
Step (4)	If $k > K$, return; else, go to Step 2

Table 2
The parameters in the three periodic Gaussian source signals.

Source signal	b_n	τ_n (s)	σ_n (ms)	T_n (s)
Source 1	1.0	0.1	12.50	1.00
Source 2	0.5	0.2	6.25	0.69
Source 3	0.7	0.2	2.50	0.80

tion results. An alternative way is to terminate our iteration by setting a threshold of estimation error [22].

Another parameter to be chosen is p in the ℓ_p norm. As mentioned earlier, a smaller p leads to a more sparse solution. However, the tradeoff is that the algorithm will have a higher likelihood of getting trapped in a local minimum. In practice, values of p between 0.8 and 1 have been found to represent a good

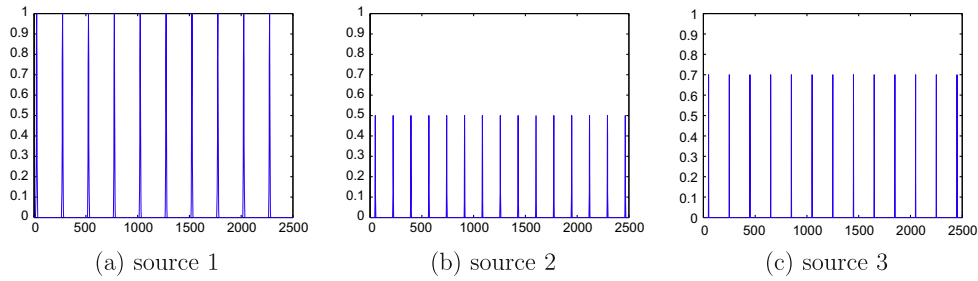


Fig. 1. The three source signals, each being a periodic truncated Gaussian pulse.

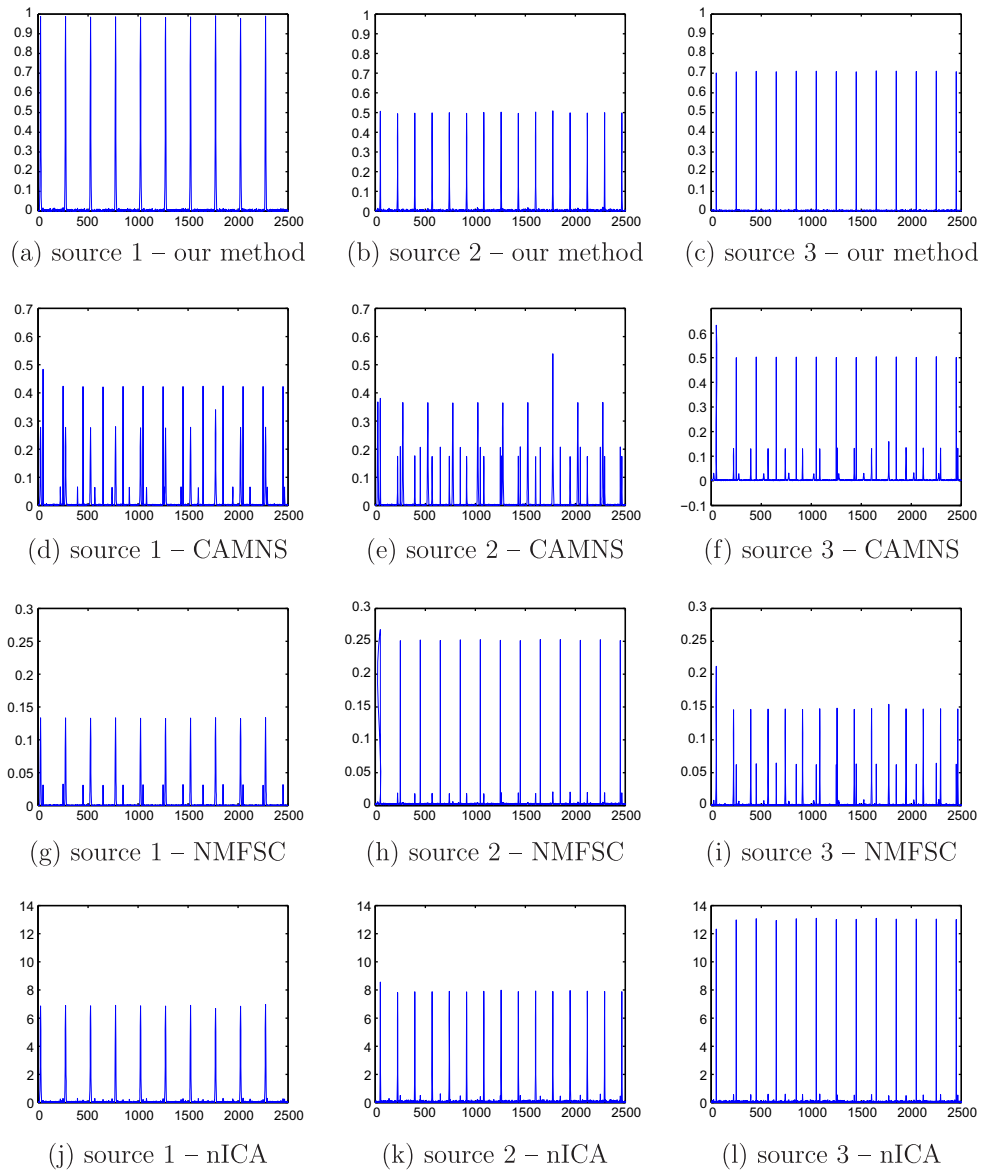


Fig. 2. The source signal separation results by the four different methods for comparison.

compromise between the speed of convergence and the quality of the resulting sparse solution.

The mixing matrix A can be initialized with randomly generated zero-mean and unit-variance Gaussian random variables. The initial source estimation can be given by $S = A^\dagger Y$, where A^\dagger is the Moore–Penrose pseudoinverse of A .

As for the regularization parameter λ , we can determine it using the traditional methods such as the modified L-curve [28]. This balances the need for a sparse solution and one that gives a small error with the observation.

3. Simulation results

In simulation with synthetic data, we can measure the accuracy of our algorithm after computing the estimate of the sources by comparing with the “real” S . Let \bar{S} denote the estimate of the sources at the K th iteration, i.e., $\bar{S} = S^{(K)}$. We define the correlation error (CE) of the source signals as

$$CE_{ij}(S, \bar{S}) = 1 - \frac{\mathbf{s}_i^T \bar{\mathbf{s}}_j}{\|\mathbf{s}_i\| \|\bar{\mathbf{s}}_j\|}, \quad (17)$$

for $i, j = 1, 2, \dots, N$. The cumulative correlation error (CCE) of the source signals is then

$$CCE_{ij}(S, \bar{S}) = \sum_{i=1}^N \min_j CE_{ij}(S, \bar{S}) = \sum_{i=1}^N \min_j \left(1 - \frac{\mathbf{s}_i^T \bar{\mathbf{s}}_j}{\|\mathbf{s}_i\| \|\bar{\mathbf{s}}_j\|} \right). \quad (18)$$

This problem can be effectively solved by the Kuhn–Munkres algorithm [29].

To gauge the performance of our algorithm, we compare results with three existing non-negative BSS algorithms, i.e., CAMNS [18],

Table 3
The CCEs of the four non-negative BSS methods.

Scheme	Experiment 1	Experiment 2	
	$M = 4, N = 3$	Noiseless	20 dB
Our method	0.0020	0.0013	0.081
CAMNS	0.0112	0.0064	0.139
NMFSC	0.0101	0.0121	0.115
nICA	0.0032	0.0520	0.175

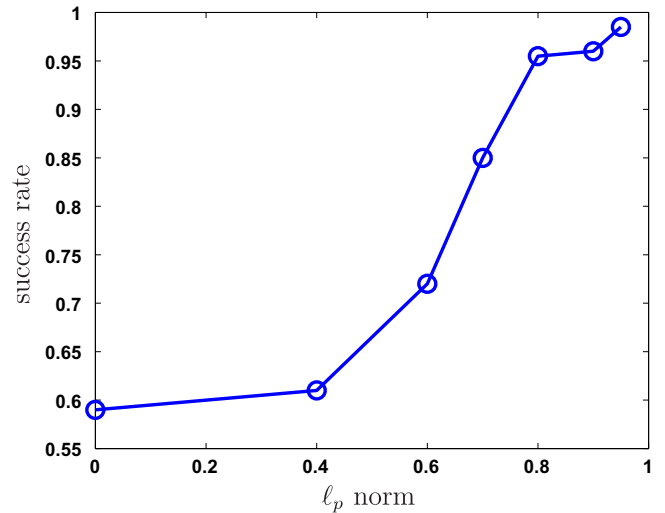


Fig. 4. The success rate of our method for different values of p .

NMFSC [17], and nICA [14]. The regularization parameter is determined by the modified L-curve [28]. We also use $K = 100$ and $p = 0.8$.

3.1. Example of mixture of three non-negative periodic signals ($M = 4, N = 3$)

In this test we use three synthetic sources that are designed to be uncorrelated. The sources are labeled \mathbf{s}_1 , \mathbf{s}_2 and \mathbf{s}_3 . Each is a periodic signal, where each period is a truncated Gaussian. In other words, for source n , let b_n be the peak, σ_n be the standard deviation of the Gaussian pulse, and τ_n be the time shift of the first period. The value of s_{nl} , within the first period, is then given by

$$s_{nl} = b_n e^{-\frac{(l-\tau_n)^2}{2\sigma_n^2}}. \quad (19)$$

This is repeated with a period of T_n . With signals that are 10 s long at a sampling frequency of 250 Hz, the parameters used for the three source signals are listed in Table 2 and the shapes of them are depicted in Fig. 1.

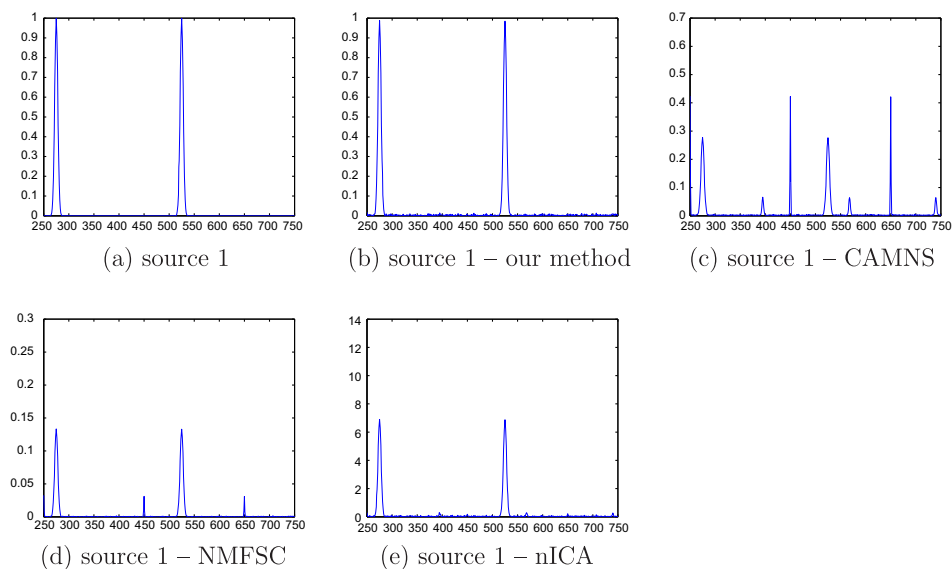


Fig. 3. The zoomed-in version of the first source signal separation results by the four different methods.

The three source signals are combined using the mixing matrix A , where

$$A = \begin{bmatrix} 0.5170 & 0.0866 & 0.3964 \\ 0.0033 & 0.4203 & 0.5764 \\ 0.8807 & 0.0301 & 0.0893 \\ 0.3953 & 0.2321 & 0.3746 \end{bmatrix}, \quad (20)$$

with a signal-to-noise ratio (SNR) at 40 dB. The source separation results using our method are given in Fig. 2a–c. These are compared with those from CAMNS in (d)–(f), NMFSC in (g)–(i), and the nICA in (j)–(l). We can see that our method yields the best separation results among the four competing schemes visually. Furthermore, in Table 3, the various methods are quantitatively compared, using the CCE defined in Eq. (18). It substantiates our claim that our method achieves the best signal separation results for the simulated EPR signals. Additionally, in order to show the

separated results clearly, we present a zoomed-in version of the four methods in Fig. 3a–e.

Our analyses of why our method is superior to the competing ones in this example is as follows. The maximum correlation coefficient among the three signals is 0.073, so they are essentially uncorrelated. As can be seen from Figs. 2 and 3, the nICA method and our method give the most precise source separation results, while the separation results of NMFSC and CAMNS are slightly worse than the two former methods. As mentioned in [17], the performance of NMFSC strongly depends on the prior knowledge of the sparsity of the sources to be separated, which is often unknown in general. Although CAMNS has good performance in separating general sources that satisfy the local dominance requirement, for the sources that are very sparse, i.e., when most parts of the sources are zero, its performance is poor because the convex hull analysis is not satisfied. Meanwhile, nICA is a classical method requiring assumption about independence

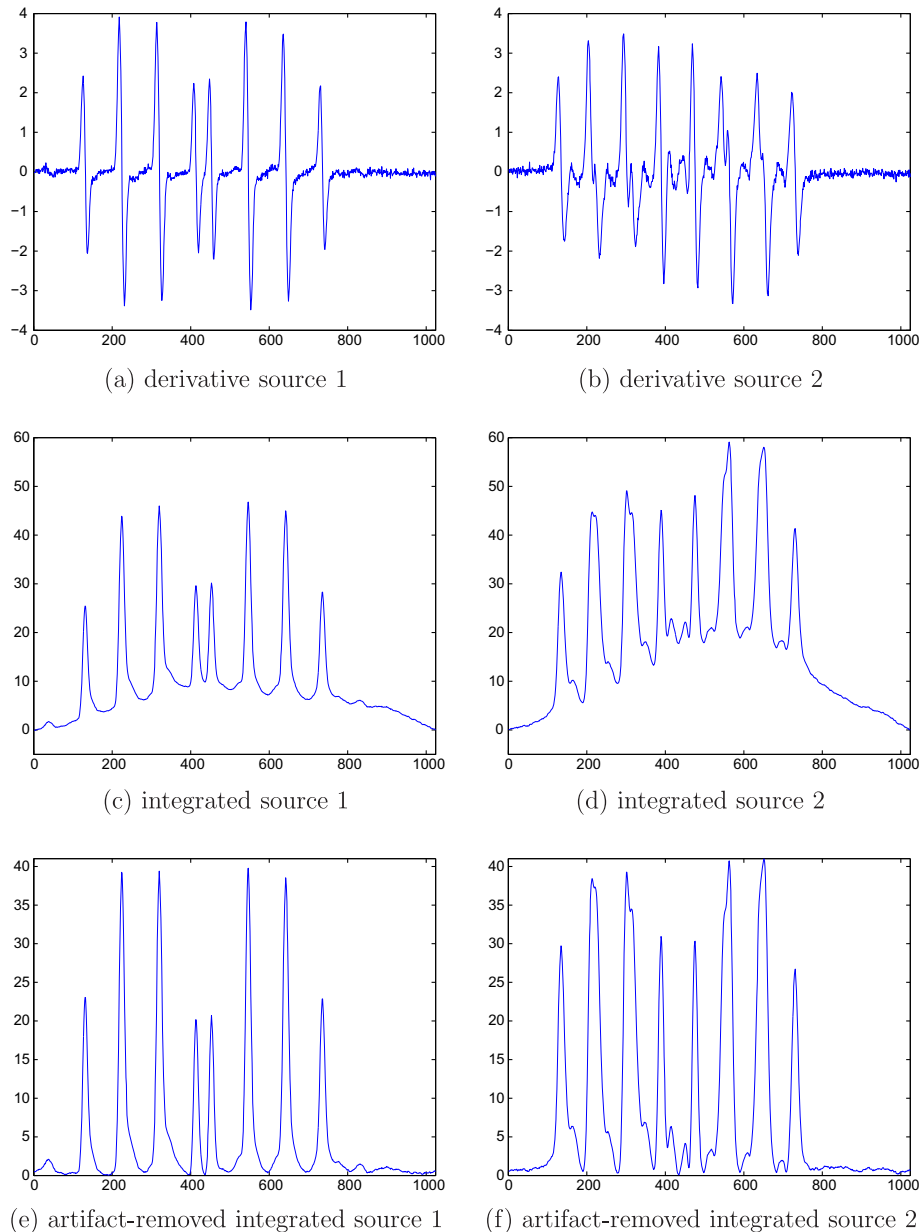


Fig. 5. The two derivative, integrated and artifact-removed EPR signals.

among the sources. For coherent source signals, the performance of nICA will degenerate substantially even for an overdetermined system. This will be further confirmed for EPR signals in the next simulation. In addition, compared with our method, CAMNS and NMFSC take a longer time to reach a stationary point. The former requires 1.704 s, the latter 7.463 s, compared with nICA which uses 0.898 s and our method, which uses 0.020 s.

We can also define the success rate of non-negative BSS as the percentage of trials in which the CCEs are less than a threshold, say, 0.05. Fig. 4 depicts the success rates of our method under different values of p with 100 independent trials. It can be seen that the results are better when p is close to 1. On the other hand, a lower value of p leads to faster convergence, but the algorithm will have a higher likelihood of getting trapped in a lo-

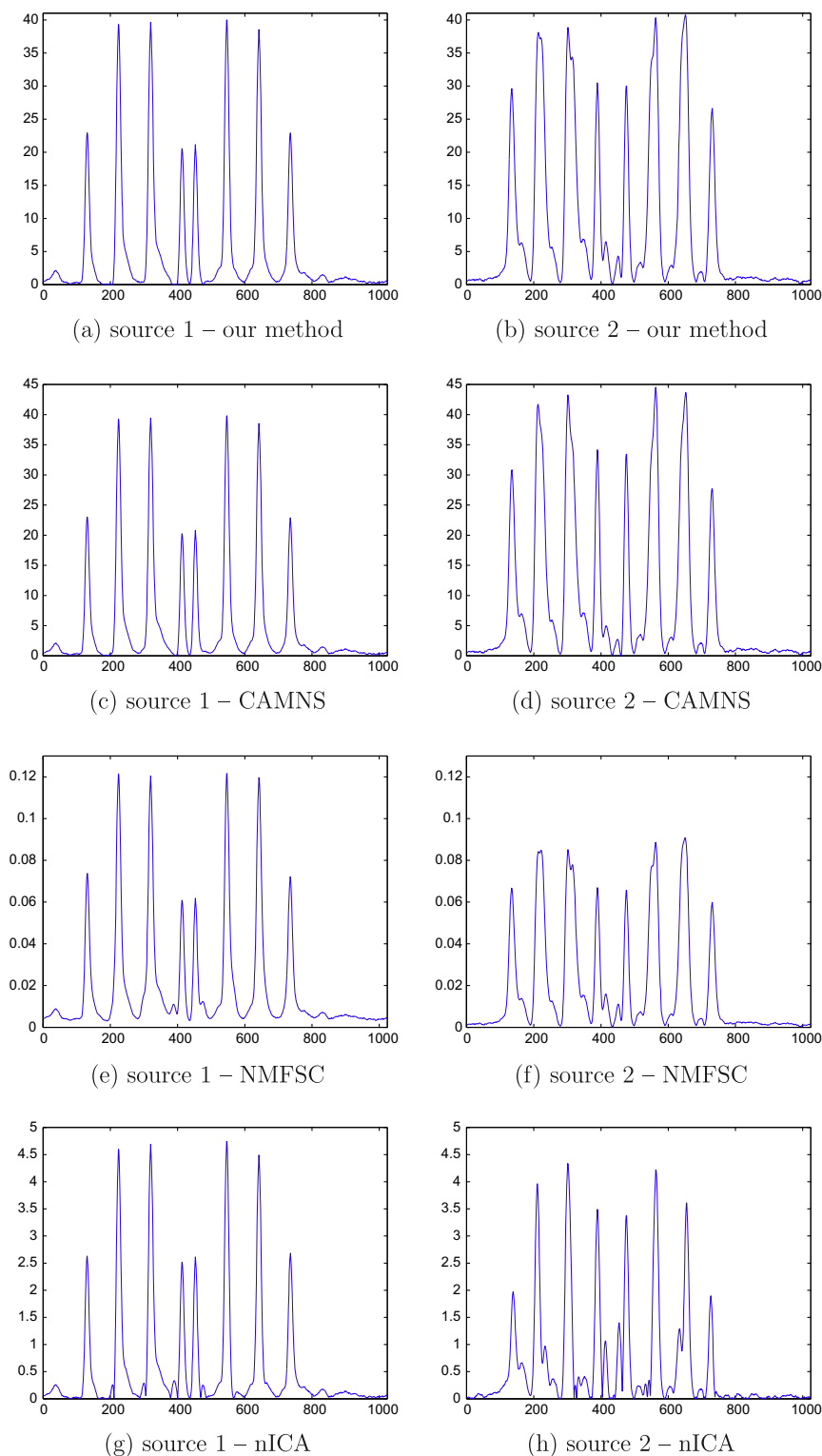


Fig. 6. The EPR source signal separation results under a noiseless scenario.

cal minimum. Overall, we pick p between 0.8 and 1 for a good tradeoff between convergence and success rate.

3.2. Example of mixture of two non-negative EPR spectra data ($M = 2, N = 2$)

Here, we consider the mixing of two continuous non-negative EPR spectra signals carefully captured experimentally in a laboratory so that each is as pure as possible. The two source EPR sig-

nals used here are the same as reported in [7]. The EPR spectra were measured with the help of the spin trap agent 5-(diethoxyphosphoryl)-5-methyl-1-pyrroline-N-oxide (DEPMPO). The spectra of superoxide-DEPMPO and hydroxyl-DEPMPO measured from chemical system are used as our source signals in this study. The measured signals are first-derivative lines of EPR spectra. All the signal intensities were digitalized to 1024 equidistant points for further processing, so the length of each signal is 1024.

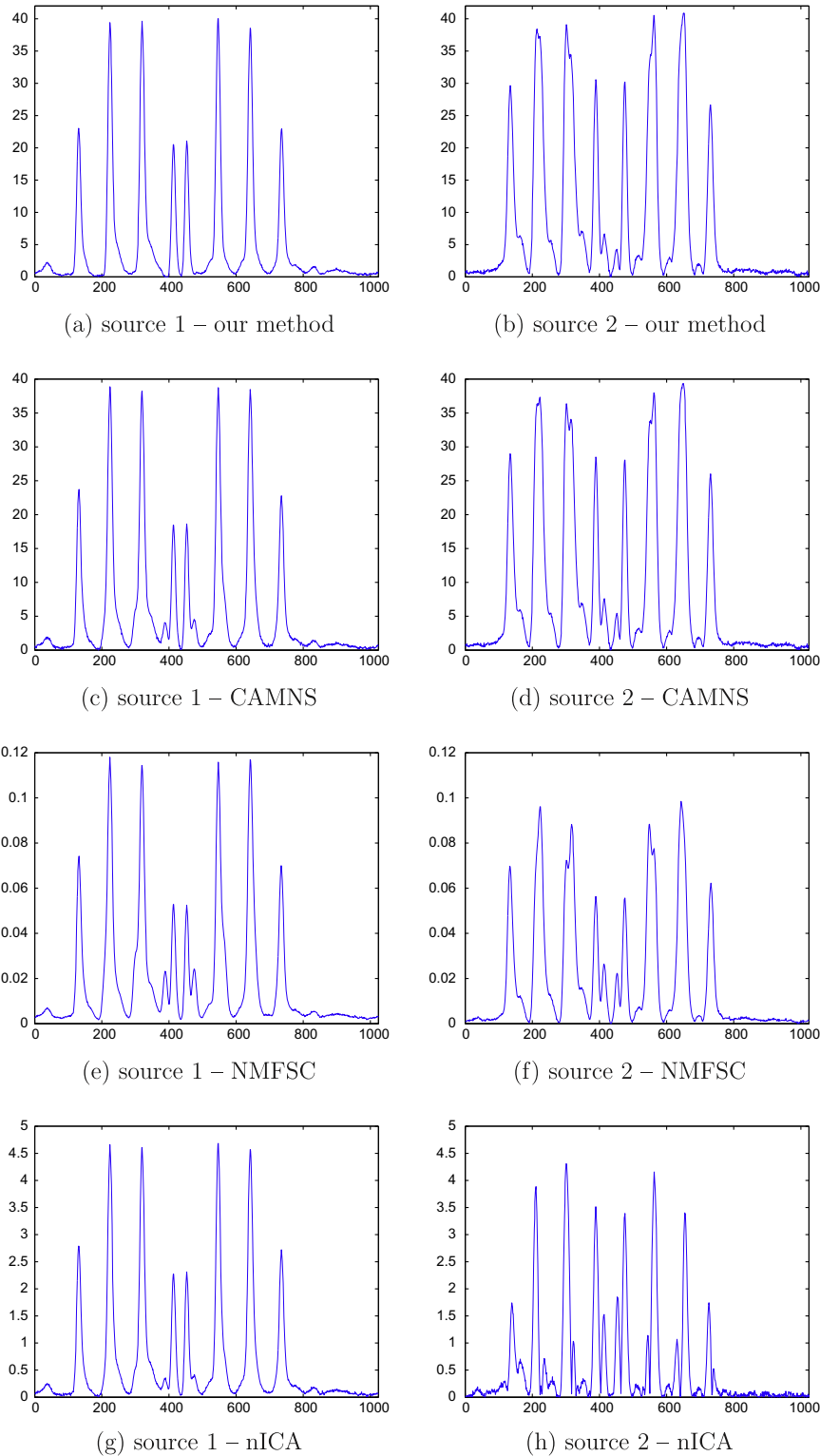


Fig. 7. The EPR source signal separation results at 20 dB.

To recover the original EPR spectra, we integrate their measured first derivatives. The spectrum should be non-negative if the response absorption is linear. In the presence of noise, the spectrum will inevitably have small negative values, particularly in the baseline region, so we set those small negative values to zero before simulating the mixed spectra. Due to some artifacts in the derivation process during measurement, it is possible that the inte-

grated spectra show an artifact of broad background under the sharply resolved peaks. This background artifact can be removed by a spline curve fitting. The measured first derivative of the two EPR spectra, their integrations, and the artifact-removed integrated spectra are shown in Fig. 5. It can be seen that the integrated spectra are nearly non-negative almost everywhere, with exceptions in the baseline region where some values are small and negative.

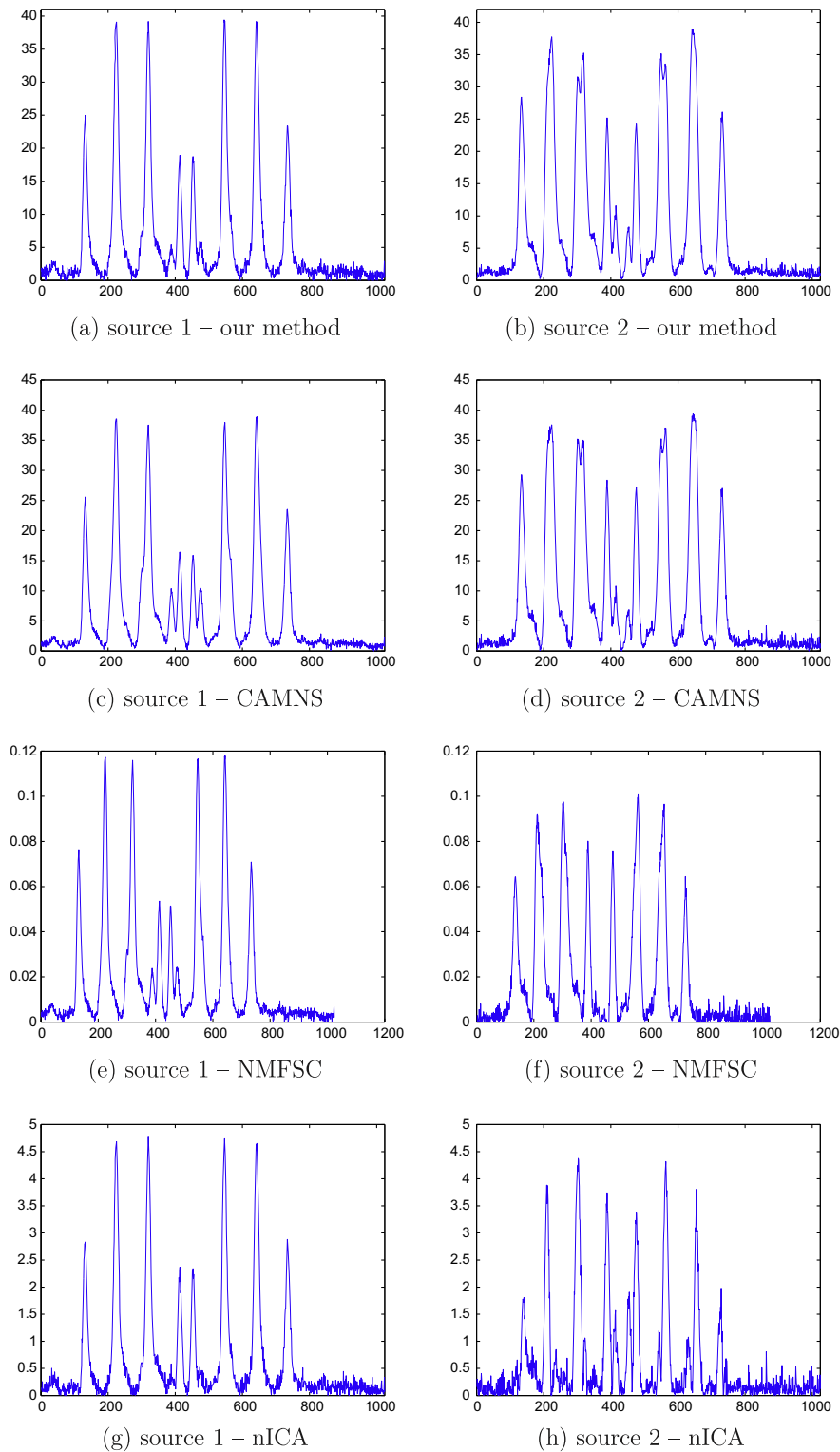


Fig. 8. The EPR source signal separation results at 0 dB.

After replacing these small negative values by zero, the source signals satisfy the non-negativity requirement of our algorithm. In fact, even if these small negative values are not replaced by zero, our algorithm can still provide a very good result, showing good robustness to slight violations to the non-negativity requirement. We state that although the two source spectra are highly overlapped, the requirement of local dominance can still be approximately satisfied, since there are still distinct peaks for each source spectra. At the distinct peaks of a source spectrum, it has a dominant value over another one. To simulate the EPR spectra of mixtures of free radicals superoxide and hydroxyl, the two artifact-corrected integrated source EPR spectra are combined using the mixing matrix

$$A = \begin{bmatrix} 0.4509 & 0.5491 \\ 0.7197 & 0.2803 \end{bmatrix}. \quad (21)$$

Fig. 6a and b give the source separation results of our method under a noiseless scenario. These are compared with CAMNS in (c) and (d), NMFSC in (e) and (f), and nICA in (g) and (h). Fig. 7 give the corresponding results for a SNR of 20 dB. Their CCEs are tabulated in Table 3.

The correlation coefficient between the two EPR signals is 0.7247, suggesting that they are not independent and, in fact, highly coherent. In the noiseless scenario depicted in Fig. 6, the waveforms of the EPR spectra are not definitive in distinguishing the four methods. However, the CCEs tabulated in Table 3 shows that CAMNS and our method have better performance than NMFSC and nICA. On the other hand, at 20 dB, the performance of CAMNS is worse than the other two methods. In fact, the separated signals contain negative values for some parts of the signals in both noiseless and noisy cases, which violates the non-negativity constraint. Note also that the performance of CAMNS degenerates as SNR decreases because the convex hull analysis is increasingly not satisfied. Meanwhile, the performance of nICA is the worst among the four methods because the assumption of independence does not hold even for the noiseless scenario, and it degrades heavily as the noise level increases. In comparison, NMFSC and our method employ regularization and therefore are more robust against noise.

We also plot the average CCEs of the four methods versus SNR from 0 to 50 dB, with 100 independent trials, as shown in Fig. 9. Our method has the smallest SNR for at various noise levels. Additionally, note that the EPR source signals we considered have a

very small number of zero points, i.e., most parts of the signals have nonzero amplitude, showing that even though the four methods are all designed for strictly sparse signal separations, ours seems to outperform the competing schemes in separating EPR signals that are only approximately sparse. Further research can probe into the mathematical reasons behind such superior performance.

As seen from Figs. 6 and 7, for the real EPR spectra of free radicals which have significant spectral overlap, our algorithm performs perfectly when there is no noise and very well when there is a high SNR of 20 dB. For a real *in vivo* EPR measurement of free radicals, the SNR may be much lower than 20 dB. However, even for a signal-to-noise ratio of 0 dB, meaning the noise energy is equal to the signal energy, our algorithm can still give a reasonably good separation with an average CCE equal to 0.2, as shown in Fig. 9. A typical separation result for the case of 0 dB is shown in Fig. 8.

4. Conclusion

This paper addresses the problem of blind separation of EPR data based on diversity measure minimization. A fixed-point alternating minimization algorithm is derived to separate the EPR source signals from linear instantaneous mixtures. We demonstrate that the proposed method can work well for signals that are not exactly sparse, but only have small amplitudes in most of the data points, which is typical in practice.

References

- [1] Y. Deng, G. He, S. Petryakov, P. Kuppusamy, J.L. Zweier, Fast EPR imaging at 300 MHz using spinning magnetic field gradients, *J. Magn. Reson.* 168 (2) (2004) 220–227, doi:10.1016/j.jmr.2004.02.012.
- [2] F. Hyodo, S. Matsumoto, N. Devasahayam, C. Dharmaraj, S. Subramanian, J.B. Mitchell, M.C. Krishna, Pulsed EPR imaging of nitroxides in mice, *J. Magn. Reson.* 197 (2) (2009) 181–185, doi:10.1016/j.jmr.2008.12.018.
- [3] G.M. Rosen, B.E. Britigan, H.J. Halpern, S. Pou, *Free radicals: biology and detection by spin trapping*, Oxford University Press, 1999.
- [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 (2) (2000) 266–275, doi:10.1006/jmre.1999.1935.
- [5] W. Lawton, E. Sylvestre, Self modeling curve resolution, *Technometrics* 13 (3) (1971) 617–633.
- [6] O. Steinbock, B. Neumann, B. Cage, J. Saltiel, S. Muller, N. Dalal, A demonstration of principal component analysis for EPR spectroscopy: identifying pure component spectra from complex spectra, *Anal. Chem.* 69 (18) (1997) 3708–3713.
- [7] J.Y. Ren, C.Q. Chang, P.C.W. Fung, J.G. Shen, F.H.Y. Chan, Free radical EPR spectroscopy analysis using blind source separation, *J. Magn. Reson.* 166 (1) (2004) 82–91.
- [8] C. Chang, J. Ren, P.C. Fung, Y. Hung, J. Shen, F.H. Chan, Novel sparse component analysis approach to free radical EPR spectra decomposition, *J. Magn. Reson.* 175 (2005) 242–255.
- [9] C. Hesse, C. James, On semi-blind source separation using spatial constraints with applications in EEG analysis, *IEEE Trans. Biomed. Eng.* 53 (12) (2006) 2525–2534.
- [10] M. Guimaraes, D.K. Wong, E. Uy, L. Grosenick, P. Suppes, Single-trial classification of MEG recordings, *IEEE Trans. Biomed. Eng.* 54 (3) (2007) 436–443.
- [11] A. Cichochi, S. Amari, *Adaptive Blind Signal and Image Processing*, John Wiley and Sons, New York, 2002.
- [12] D. Nuzillard, J.M. Nuzillard, Application of blind source separation to 1-D and 2-D nuclear magnetic resonance spectroscopy, *IEEE Trans. Signal Process. Lett.* 5 (8) (1998) 209–211.
- [13] Z. Xu, E.Y. Lam, Maximum a posteriori blind image deconvolution with Huber–Markov random-field regularization, *Opt. Lett.* 34 (9) (2009) 1453–1455.
- [14] M.D. Plumbley, Algorithms for non-negative independent component analysis, *IEEE Trans. Neural Network* 14 (3) (2003) 534–543.
- [15] D.D. Lee, H.S. Seung, Learning the parts of objects by non-negative matrix factorization, *Nature* 6755 (401) (1999) 788–791.
- [16] E.Y. Lam, Non-negative matrix factorization for images with Laplacian noise, in: *IEEE Asia Pacific Conference on Circuits and Systems*, 2008, pp. 798–801.
- [17] P.O. Hoyer, Non-negative matrix factorization with sparseness constraints, *J. Mach. Learn. Res* 5 (2004) 1457–1469.

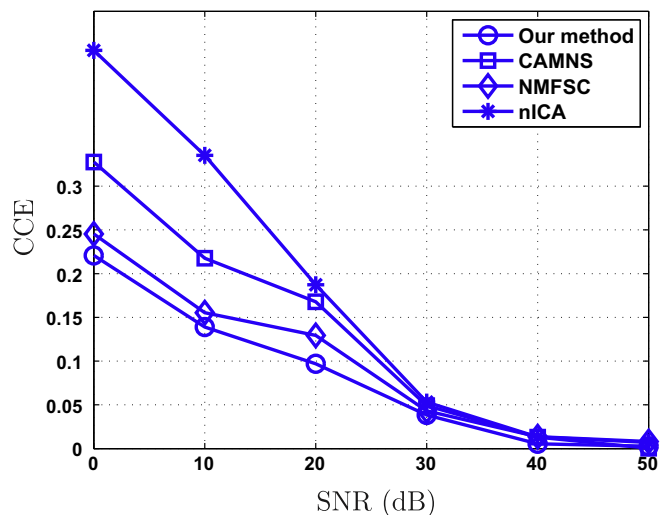


Fig. 9. The average CCEs of four different methods versus SNR with 100 independent trials.

- [18] T.-H. Cha, W.-K. Ma, C.-Y. Chi, Y. Wang, A convex analysis framework for blind separation of non-negative sources, *IEEE Trans. Signal Process.* 56 (10) (2008) 5120–5134.
- [19] Y. Shen, E.Y. Lam, N. Wong, Binary image restoration by positive semidefinite programming, *Opt. Lett.* 32 (2) (2007) 121–123.
- [20] G.H. Golub, C.F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Maryland, 1996.
- [21] Y. Shen, E.Y. Lam, N. Wong, A signomial programming approach for binary image restoration by penalized least squares, *IEEE Trans. Circuits Syst. II* 55 (1) (2008) 41–45.
- [22] J. Bobina, J.-L. Starck, J. Fadili, Y. Moudden, Sparsity and morphological diversity in blind source separation, *IEEE Trans. Signal Process.* 16 (11) (2007) 2662–2674.
- [23] J.A. Tropp, *Topics in sparse approximation*, Ph.D. thesis, The University of Texas at Austin, Texas, August 2004.
- [24] B.D. Rao, K. Kreutz-Delgado, An affine scaling methodology for basis selection, *IEEE Trans. Signal Process.* 47 (1) (1999) 187–200.
- [25] D.L. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory* 52 (4) (2006) 1289–1306.
- [26] V.K. Goyal, A.K. Fletcher, S. Rangan, Compressive sampling and lossy compression, *IEEE Signal Process. Mag.* (2008) 48–56.
- [27] H. Rauhut, K. Schnass, P. Vandergheynst, Compressed sensing and redundant dictionaries, *IEEE Trans. Inform. Theory* 54 (5) (2008) 2210–2219.
- [28] B.D. Rao, K. Egan, S.F. Scotter, J. Palmer, K. Kreutz-Delgado, Subset selection in noise based on diversity measure minimization, *IEEE Trans. Signal Process.* 51 (3) (2003) 760–770.
- [29] K.-M. algorithm, Online. Available from: <http://si.utia.cas.cz/downloadPT.htm%3e>.