A Novel Underdetermined Source Recovery Algorithm based on k-Sparse Component Analysis

**Ehsan Eqlimi · Bahador Makkiabadi ·
Nasser Samadzadehaghdam · Hassan
Khajehpour · Saeid Sanei**

Abstract Sparse component analysis (SCA) is a popular method for addressing underdetermined blind source separation (UBSS) in array signal processing applications. We are motivated by problems that arise in the applications where the sources are densely sparse (i.e. only a limited number of sources are inactive at each time instant). The separation performance of current underdetermined source recovery (USR) solutions, including the relaxation and greedy families, reduces with decreasing the mixing system dimension and increasing the sparsity level (k). In this paper, we present a k-SCA based algorithm that is suitable for USR in low dimensional mixing systems. Assuming the sources are at most $(m-1)$-sparse where m is the number of mixtures, the proposed method is capable of recovering the sources from the mixtures given the mixing matrix using a subspace detection framework. Simulation results show that the proposed algorithm achieves better separation performance in k-SCA conditions compared with state of the art USR algorithms such as basis pursuit (BP), minimizing L1-norm (ML1), smoothed L0 (SL0), focal underdetermined system solve (FOCUSS) and orthogonal matching pursuit (OMP).

E.Eqlimi, B.Makkiabadi, N.Samadzadehaghdam and H.Khajehpour
1- Department of Medical Physics and Biomedical Engineering, School of Medicine,
Tehran University of Medical Sciences (TUMS), Tehran, Iran
2-Research Center for Biomedical Technology and Robotics (RCBTR),
Institute of Advanced Medical Technologies (IAMT), TUMS, Tehran, Iran
E-mail: Eghlimi@razi.tums.ac.ir and B-Makkiabadi@sina.tums.ac.ir

S. Sanei
Department of Computer Science, University of Surrey, Surrey GU2 7XH, UK

1 Introduction

Blind source separation (BSS) problem refers to recovering the source signals from their mixtures (so-called observations) with little information about the mixing system and the source signals [1]. However, to solve the BSS problem, we need to exploit some a priori knowledge, e.g. the independence of source signals in independent component analysis (ICA) or the sparsity of the source signals in sparse component analysis (SCA) [2–5].

An instantaneous linear mixing system in the absence of noise can be formulated as follows:

$$\mathbf{X} = \mathbf{AS}, \tag{1}$$

where $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_T] \in \mathbb{R}^{m \times T}$ is the mixture matrix, $\mathbf{A} = [\mathbf{a}_1, ..., \mathbf{a}_n] \in \mathbb{R}^{m \times n}$ is the mixing matrix, and $\mathbf{S} = [\mathbf{s}_1, ..., \mathbf{s}_T] \in \mathbb{R}^{n \times T}$ is the source matrix (m, n and T are the number of sensors, sources and signal smaples, respectively). If there are more sensors than sources i.e. $m > n$, BSS is an overdetermined problem and could be solved using ICA where the sources are statistically independent and non-Gaussian [6, 7]. If there are more sources than sensors i.e. $m < n$, the source separation problem is an underdetermined BSS (UBSS) problem [4,8]. In an underdetermined case, the mixing matrix is non-invertible and therefore the source recovery is a challenging problem even if the mixing matrix is known or could be identified.

A two step technique is employed to address the UBSS problem: underdetermined blind identification (UBI) and underdetermined source recovery (USR) [9]. The mixing matrix ($\mathbf{A}$) is identified using UBI whereas USR recovers the source signals in equation (1).

SCA [2–5, 10, 11] is a general approach to address UBSS problem which benefits the sparseness of the source signals. In contrast with ICA algorithm, the SCA could separate the source signals even when the source signals are dependent or have Gaussian distributions. In SCA approach, a few sources take the significant values in each instant i.e. the columns of matrix $\mathbf{S}$ are sparse vectors.

In the underdetermined cases, the SCA-based UBSS has applications in many signal processing problems such as seperation of speech signals [12], extraction of single trial event related potential (ERP) signals [13] and electroencephalography (EEG) source imaging where it is aimed to reconstruct the brain sources with few sensors [14, 15].

Generally, there are two groups of methods for solving UBSS problem based on SCA. A number of methods have been proposed employing the sparsity of the sources in the temporal domain [4] and some other approaches operate in time-frequency (TF) domain [16–18].

There are several TF-SCA approachs which assume all or most of the sources are W-disjoint, i.e. there is at most one active source at each time instant. In this case, the scatter plot of mixed signals shows clear geometrical line orientations related to the columns of mixing matrix while W-disjointendness is a restrictive requirement in the most real applications.

The degenerate unmixing estimation technique (DUET) for two sensors [19] and multiple sensor DUET (MENUET) for multiple sensors [20] are well-known TF-SCA methods which rely on source W-disjointedness. DUET allows one to estimate the mixing parameters by clustering ratios of the time-frequency representations of the mixtures. Then, the mixing parameters are employed to separate the time-frequency representation of one mixture to recover the original sources. MENUET generalized DUET to multiple sensors using arbitrary geometric arrangement using K-means clustering.

Time-frequency ratio of mixtures (TIFROM) [21, 22] method is another TF-SCA which first finds single-source TF zones, i.e. a set of adjacent TF windows, where a single source is active. Then, in each of the mentioned zones, it estimates a column of the mixing matrix. When all the columns of the mixing matrix have been estimated, the last step consists in recovering the sources. TIFROM method relaxes the restrictive W-disjoint assumption but still needs to have adjacent TF windows with one active source. It is notable that the source separation performance of the TF-SCA UBSS algorithms depends highly on the correctly detection of single source points (SSPs).

Georgiev et $al.$ [4] proposed k-sparse component analysis (k-SCA) approach to solve the UBSS problem assuming at most $m-1$ active sources in each instant. The letter of k denotes the sparsity level and refers to the cardinality of column source vector at each time instant or any transformed domain bin. Indeed, they proved that the UBSS problems are solvable as long as the number of active sources in each time instant is less than the number of mixtures ($k \leq m-1$). During these years, some k-SCA based UBSS methods have been proposed based on the k-SCA assumptions in the literature. Most of the developed k-SCA algorithms for solving UBSS problem focus on identifying the mixing matrix (UBI) [23–30] and utilise conventional source recovery algorithms i.e. greedy methods, e.g., orthogonal matching pursuit (OMP) [31–33], relaxation methods, e.g., basis pursuit (BP) [11,34–36] and smoothed L0 (SL0) algorithm [37,38].

Generally, BP methods expolit the ℓ_1-norm minimization algorithm assuming the source signals (i.e. the rows of matrix $\mathbf{S}$) are independent and identically distributed (i.i.d) [34, 39] and their source recovery performance decreases with higher k. Actually, the most conventional USR algorithms like BP algorithm achieve a good separation performance for very high dimensional mixing system ($m \times n$) and they fail when the mixing system dimension is low.

With regards to above discussion, we are motivated to develop a new USR algorithm based on k-SCA assumptions [4] (discussed in section 3) where $k \leq m-1$ and hence is able to solve USR step of UBSS problem without using any conventional USR algorithm. The proposed method is not based on any BP algorithms and its approach is more similar to the greedy based source recovery

approaches. Indeed, based on the k-SCA assumptions, we propose a novel and effective method to solve the USR problem assuming the mixing matrix ($\mathbf{A}$) is known or could be identified using methods such as those proposed in [23–30]. It detects the generating subspaces of each mixed data based on eigenvalue decomposition of concatenated mixed data and the columns of the mixing matrix. Also, it is able to solve the USR problem even for the cases where $k = m - 1$. The proposed algorithm achieves better performance compared to the standard source recovery methods such as BP, OMP and SL0 algorithms for low dimensional mixing systems.

The remainder of the paper is organized as follows: The next section briefly explains the standard approaches for the sparse source recovery. Section 3 reviews the k-SCA assumptions and related theorem. Section 4 describes our proposed solution to solve the USR problem based on k-SCA assumptions. In Section 5 the numerical experiments to compare our proposed algorithm with standard USR methods are presented. We consider four scenarios to examine the effects of mixture noise, source noise, sparsity level, mixing matrix estimation error, and dimension of mixing system on source separation signal to noise Ratio (SNR). In addition, the computational cost is also discussed and reported for the algorithms in subsection 5.6. Finally, concluding remarks are discussed in Section 6.

2 The review of standard algorithms for sparse source recovery

We have compared the results of the proposed USR algorithm with those of standard algorithms. Therefore, in this section, the standard algorithms for sparse source recovery will be reviewed briefly.

Consider $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a full rank matrix with $m < n$. Therefore, the system $\mathbf{x} = \mathbf{As}$ is an underdetermined linear system and obviously there are infinitely many possible solutions $\mathbf{s}$. In order to obtain a single and well- defined solution, we need an additional criteria. If we are interested in seeking the sparsest solution, the criteria is sparsity of $\mathbf{s}$ and thus the problem is as follows:

$$(P_0): \quad \min_{\mathbf{s}} \ ||\mathbf{s}||_0 \quad \text{subject to} \quad \mathbf{x} = \mathbf{As}, \tag{2}$$

where $||\mathbf{s}||_0$ denotes the ℓ_0-norm of $\mathbf{s}$, counting the number of nonzero entries in it. In general, there is no straightforward method to solve (P_0) since the ℓ_0-norm is a highly non-smooth function. Therefore, we need some specific conditions to solve it. Mostly, three approaches has been taken in the literature to solve the sparse source recovery problem. One approach that is often called greedy family focuses on the support and then the nonzero values of $\mathbf{s}$ could be obtained using the simple least squares method. The greedy algorithms are mostly discrete due to the discrete nature of support.

On the contrary, a different approach that often called relaxation family adopt a continuous optimization to solve (P_0). Generally, the relaxation approach smooths the ℓ_0-norm in various form mostly using the convex relaxation technique [40].

Furthermore, an alternative view to solve (P_0) is smoothed ℓ_0-norm (SL0) algorithm which try to directly minimize the ℓ_0-norm [38]. Unlike the BP methods, which replace ℓ_0-norm by other convex functions, SL0 is a non-convex method and employs graduated-non convex (GNC) procedure to escape the local minima.

In the following, each approach will be explained individually.

2.1 Greedy approach

Mallat and Zhang [31] have introduced a greedy and myopic method for sparse decomposition which is usually referred to as matching pursuit (MP). They shown when the dictionary (mixing matrix) is orthogonal and the observed signal is composed of $k \ll n$ atoms, the algorithm recovers the sparse decomposition exactly after n steps [41]. In fact, MP is based on a suboptimal forward search through the mixing matrix. Succeeding greedy algorithms such as OMP [42] and compressive sampling matched pursuit (CoSaMP) [33] are based on MP concept. OMP method chooses the column of maximal correlation with residue, in fact, its also the one having the steepest decline in residue, which implies OMP is greedy.

Let $\mathbf{x} = \mathbf{A}\mathbf{s}$ and $\mathbf{s}$ is a k-sparse signal. Since the number of nonzero coefficients of a k-sparse signal is equal to k, if the places of these coefficients are determined, the non-zero values could be found by inverting or pseudo-inverting of the corresponding $m \times k$ sub-matrices in $\mathbf{A}$. In fact, finding the non-zero places is the main complexity. In OMP algorithm, the non-zero places are found one by one. We could use the similarity between $\mathbf{x}$ and $\mathbf{A}$ and actually the most likely column is belonged to the support. To find the other members of the support, the found column effect is deflated. This process will be repeated to find the other members of support.

2.2 Relaxation approach

On the other hand, relaxation methods smooth the ℓ_0-norm and use continuous optimization techniques with the hope to simplify the problem while preserving its essence. Some example for smoothed version of ℓ_0-norm are $\rho_\alpha(\mathbf{s}) = 1 - exp(-\mathbf{s}^2/\alpha)$, $\rho_\alpha(\mathbf{s}) = \frac{\mathbf{s}^2}{\alpha + \mathbf{s}^2}$ and $\rho_\alpha(\mathbf{s}) = |\mathbf{s}|^\alpha$. As α tends to zero, these functions get closer to ℓ_0-norm.

An popular and interesting algorithm of this family is the focal underdetermined system solver (FOCUSS) [43]. FOCUSS employs a tricky optimization method called iterative recursive least-squares (IRLS). The core idea is the representation of smoothed ℓ_0-norm ($\rho_\alpha(\mathbf{s})$) as a weighted ℓ_2-norm as follows:

$$\sum_{q=1}^{n} \rho_\alpha(s_q) \rightarrow \sum_{q=1}^{n} \frac{\rho_\alpha(s_q)}{s_q^2} s_q^2 = \sum_{q=1}^{n} w_q s_q^2. \tag{3}$$

IRLS iterates between a solution of the ℓ_2-norm and update of the weights ($\mathbf{w}$).

An alternative possible function to smooth the ℓ_0-norm, which has drawn much attention is ℓ_1-norm , i.e. $|\mathbf{s}|$. The core idea is to replace ℓ_1-norm by an absolute value of $\mathbf{s}$, which indeed leads to an ℓ_1-norm penalty is as follows:

$$(P_1): \quad \min_{\mathbf{s}} \ ||\mathbf{s}||_1 \quad \text{subject to} \quad \mathbf{x} = \mathbf{As}, \tag{4}$$

This convex relaxation technique is called BP [44] and could admit a linear programming form. In general, BP method could provide good estimation of sources in comparison with OMP and FOCUSS methods but it is still slow.

2.3 Smoothed ℓ_0 approach

Smoothed ℓ_0 algorithm [38] approximates the ℓ_0-norm of vector $\mathbf{s}$ by a smooth function as follows:

$$f_\sigma(\mathbf{s}) = exp(-\mathbf{s}^2/\sigma) \tag{5}$$

The σ parameter controls the quality of approximation such that the smaller σ results the closer behaviour of $f_\sigma(\mathbf{s})$ to ℓ_0-norm. The aim of this algorithm is to maximize $f_\sigma(\mathbf{s})$ for a very small σ, which gives the minimum ℓ_0-norm solution. However, $f_\sigma(\mathbf{s})$ for very small σ is highly non-smooth with a lot of local minima. Smoothed ℓ_0 algorithm employs a graduated non-convex procedure with the hope to escape from these local minima. SL0 shows the lower complexity in compression with BP and OMP methods.

3 k-Sparse Component Analysis

Most of the sparseness based USR methods, which be reviewd in previous section, achieve good results when $k \ll m$ and the dimension of mixing system ($m \times n$) is high. To overcome this limitation, Georgive $et\ al.$ [3] proved their theorem, refereed to as k-SCA. In contrast to the existing SCA algorithms which generally rely on having $k \ll m$, they proved that the UBSS problem can be tackled using the k-SCA assumptions as long as $k \leq m - 1$.

Based on equation (1), the adopted version of instantaneous mixing system in a vector-wise scheme, can be formulated as follows:

$$\mathbf{x}_t = \sum_{q=1}^{n} \mathbf{a}_q s_t(q), \tag{6}$$

which $\mathbf{x}_t \in \mathbb{R}^{m \times 1}$ includes the m mixture values in instant t, $s_t(q) \in \mathbb{R}^{1 \times 1}$ is the qth source value in instant t and $\mathbf{a}_q \in \mathbb{R}^{m \times 1}$ is the qth column of mixing matrix.

According to new formulation, the mixture signals in each time instant (shown by t) is built by linear combination of k weighted columns of $\mathbf{A}$ where k is the number of active sources in each time instant.

The k-SCA assumptions are listed as follows [4]:

1. Each square $m \times m$ submatrix of $\mathbf{A}$ is nonsingular.
2. Source matrix $\mathbf{S}$ has at most $m - 1$ active source in each column.
3. Source matrix is able to excite all the possible subspaces sufficiently i.e., for any index subset of $n - (m - 1)$ elements $\mathcal{I} = \{i_1, ..., i_{n-m}, i_{n-m+1}\} \subset \{1, ..., n\}$ there exist at least m columns of $\mathbf{S}$ matrix such that each of them has inactive elements in places with indices in $\mathcal{I}$ and each $m - 1$ of them are linearly independent.

Based on the theorem of uniqueness of sparse represenation [4], for a linear system of the form $\mathbf{As} = \mathbf{x}$ and $\mathcal{H}$, the set of all $\mathbf{x} \in \mathbb{R}^m$, such that $\mathbf{s}$ vector is found with at least $n - (m - 1)$ zero elements, if the k-SCA assumptions are satisfied, then there is a subset $\mathcal{H}_0 \subset \mathcal{H}$, such that for every $\mathbf{x} \in \mathcal{H} \backslash \mathcal{H}_0$ ($\{\mathbf{x} \in \mathcal{H} \mid x \notin \mathcal{H}_0\}$) this system has no other solution with this property.

Let $\mathcal{I}$ be the set of all subsets of $\{1, .., n\}$ containing $k \leq m - 1$ elements. It is clear that $\mathcal{I}$ consists of C_k^n (combination k from n) elements. Each element of $\mathcal{I}$ shows the indices of the columns of $\mathbf{A}$ that are used to generate the observed data points. Therefore, each data point lies in a union of k-dimensional subspaces. The number of all possible subspaces will be C_k^n. The basis vectors of the subspaces are k columns of $\mathbf{A}$ with the indices in $\mathcal{I}$.

Based on Georgive *et al*'s proof [4], as long as $k \leq m - 1$, where k is the number of active sources at each column of $\mathbf{S}$, we can identify $\mathbf{A}$ and estimate $\mathbf{S}$. Essentially, Georgive's proof guarantees both stages of UBSS (UBI and USR), however he and his co-authors did not provide any detailed or robust algorithm to justify the stages for UBI and USR. Albeit, few methods, including two approaches proposed by our group recently, extended this method to solve the UBI problem [23–30].

4 Proposed USR algorithm based on k-SCA assumptions

As mentioned before, so far, no direct method has been proposed to tackle the USR problem which considers the k-SCA assumptions and the least possible relaxation to estimate the sparse sources. Therefore, a dedicated method for solving the above-mentioned USR problem is necessary especially for applications with more dense sparse sources. In this regard, we have tried to introduce a fundamental approach when k is known or detected using the proposed methods in [27, 30] for each time instant.

Lets define the disjoint and f-nondisjoint subspaces concepts.

Definition 1 (disjoint subspaces) If there is no common mixing vector between the subspaces then they are called disjoint subspaces

Definition 2 (f-nondisjoint subspaces) If there are f common mixing vector between the subspaces then they are called f-nondisjoint subspaces. Obviously, the number of f-nondisjoint subspaces is $nd_f = C_{m-1-f}^{n-f}$ where $f = 1, ..., m - 1$.

For example, lets subspace $\mathcal{P}_1$ is spanned by 1st, 2nd and 3rd columns of mixing matrix and subspace $\mathcal{P}_2$ is spanned by 1st, 3rd and 4th columns of mixing matrix. In this case, $\mathcal{P}_1$ and $\mathcal{P}_2$ are 2-nondisjoint subspaces ($f = 2$) beacuse there are two common mixing vectors between them (1st and 3rd columns of A).

For $k = 1$ the subspaces are disjoint because there are not common mixing vector to span the subspaces (i.e. $k - 1 = 0$) but if $k > 1$ there are some nondisjoint subspaces. Therefore, in general, the subspaces could be nondisjoint.

If we identify an orthogonal complement space which is orthogonal to the observed data point and k columns of $\mathbf{A}$ then, the observed data point lies on that subspace spanned by those k columns of $\mathbf{A}$, i.e., the observed data point will be the linear combination of those k vectors.

In the proposed algorithm, in order to find the closest subspace to each observed vector, we first find the set of all subsets of $\{1, .., n\}$ containing $k \leq m - 1$ elements called $\mathcal{I}$. $\mathcal{I}^i$ is the ith member of $\mathcal{I}$ representing the indices of k columns of $\mathbf{A}$ where $i = 1 : C_k^n$. For example, assume $m = 4$, $n = 5$ and $k = m - 1 = 3$. In this case the number of all subspaces are $d = C_k^n = C_3^5 = 10$ where each $nd_1 = C_{k-1}^{n-1} = C_{3-1}^{5-1} = 6$ is a 1-nondisjoint subspace, each $nd_2 = C_{k-2}^{n-2} = C_{3-2}^{5-2} = 3$ is a 2-nondisjoint subspace and each $nd_3 = C_{k-3}^{n-2} = C_{3-3}^{5-3} = 1$ is a 3-nondisjoint subspace.

Assuming the columns of the mixing matrix are normalized to ℓ_2-norm, for a given data point $\mathbf{x}_t$ and all members of $\mathcal{I}$, we calculate the eigenvalues and eigenvectors of covariance matrix of $\mathbf{F}^i = [\mathbf{x}_t \ \mathbf{A}(:, \mathcal{I}_i)]$ (i.e. $E\{\mathbf{F}^i \mathbf{F}^{i^T}\}$) called augmented mixing matrix (AMM) using eigenvector decomposition (EVD) method where $i = 1 : C_k^n$. $\mathbf{F}^i$ or ith AMM represents a matrix containing the observed data in each instant and the selected columns of mixing matrix which correspond to $\mathcal{I}_i$.

The eigenvector corresponding to minimum eigenvalue is most orthogonal to $\mathbf{F}^i$. Indeed, we need to find the minimum eigenvalue of each $\mathbf{F}^i$ and there are C_k^n minimum eigenvalues which are inserted in *min eigenvalues vector* ($\mathbf{v}$). We employed an iterative and numerical method [45] to compute the smallest eigenvalue.

The number of k-nondisjoint subspaces is used to compute the orthogonal vectors to the covariance of AMM. For $k = m - 1$, the number of k-nondisjoint subspaces is $nd_k = C_{m-1-m-1}^{n-k} = C_0^{n-k} = 1$. Consequently, when $k = m - 1$ we find the element of $\mathcal{I}$ corresponding to minimum value of $\mathbf{v}$ because $nd_k = 1$. But if $k < m - 1$ the number of k-nondisjoint interferer subspaces increases i.e. $nd_k = C_{m-1-k}^{n-k} > 1$. In this case we find nd_k member of $\mathcal{I}$ corresponding to nd_k first lower values of $\mathbf{v}$. Then, we find k basis vectors that have the maximum redundancy between nd_k members of $\mathcal{I}$. The estimated indices are inserted in a new vector $\mathbf{h}$. These k basis vectors span those subspaces that $\mathbf{x}_t$ lies in it. Thus, we find the generating mixing vectors of each data point ($\mathbf{W}$). Finally, each source point could be recovered by pseudo-inverse minimum mean squared error (MMSE) [46, 47] as follows:

$$\hat{\mathbf{s}}_t = \mathbf{W}^{t^T}.(\mathbf{W}^t.\mathbf{W}^{t^T} + p_N.\mathbf{I}_m)^{-1}.\mathbf{x}_t, \tag{7}$$
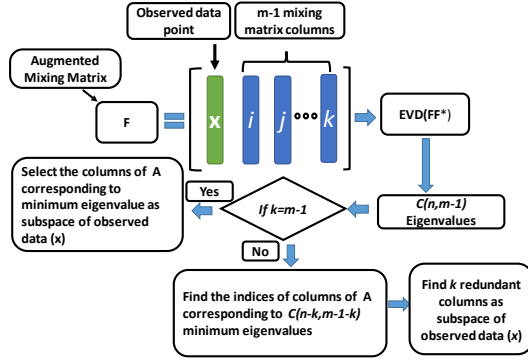
Fig. 1: Block diagram to find the subspaces of each data point

where $t = 1 : T$ (T = the number of data points or instant), $\mathbf{W}^t = \mathbf{A}(:, \mathbf{h}^t)$, $\hat{\mathbf{s}}_t = \hat{\mathbf{S}}(\mathbf{h}^t, t)$, $\mathbf{x}_t = \mathbf{X}(:, t)$, $\mathbf{I}_m$ is the $m \times m$ identity matrix and p_N is the power of the additive noise which is estimated using the iterative algorithm in [48].

Based on the above description, we propose Algorithm 1 to recover the source matrix based on k-SCA assumptions. The schematic block diagram to find the subspaces of each data point ($\mathbf{x}_t$) has been illustrated in Fig. 1. Since the degree of membership of each data point to the detected subspaces is measured using the EVD method, our proposed algorithm is called eigenvalue membership (EigMem) USR.

Remark: It is worthy to note that for high SNR mixing scenarios, the proposed algorithm is capable of source recovery without estimating $k \leq m-1$. In these cases, because of having negligible effect of noise on the recovered sources, the condition on step 7 of Algorithm 1 can be removed, and considered as *true* for all $k \leq m - 1$.

Algorithm 1 The Proposed USR Algorithm based on k-SCA assumptions

---

Input: $\mathbf{X} \in \mathbb{R}^{m \times T}$: the mixed signals, $\mathbf{A} \in \mathbb{R}^{m \times n}$: the mixing matrix, k: the number of dominant sources at each time instant.
Output: $\hat{\mathbf{S}} \in \mathbb{R}^{n \times T}$: the estimated source matrix
1: for $i = 1 : C_{m-1}^n$
2: $\mathcal{I}(i, :) \leftarrow [i_1, ..., i_k] \subset [1, ..., n]$; **end for**.
3: for $t = 1 : T$ **do**,
4: for $j = 1 : C_{m-1}^n$ **do**,
5: $\mathbf{F} = [\mathbf{x}_t, \mathbf{A}(:, \mathcal{I}(j, :)]$,
6: $\mathbf{v}(j, :) \leftarrow$ minimum eigenvalue of $\mathbf{FF}^T$; **end for**.
7: if $k = m - 1$ then $\mathbf{h}^t \leftarrow$ index from $\mathcal{I}$ corresponding to minimum eigenvalue of $\mathbf{v}$,
8: else,
9: $\mathbf{h}^t \leftarrow C_{m-1-k}^{n-k}$ indices from $\mathcal{I}$ corresponding to C_{m-1-k}^{n-k} first lower eigenvalues of $\mathbf{v}$ and find the k frequent indices; **end if**.
10: $\mathbf{W}^t = \mathbf{A}(:, \mathbf{h}^t)$;
11: Employ equation 8 to recover the source data point i.e. $\mathbf{S}(:, t)$; **end for**.

---

5 Experiments and Results

In this section, the performance of the proposed approach is experimentally verified and compared with SL0[1] [38], BP and BP denoise (BPDN) using SPGL1 toolbox[2] [49,50], $ML1$[3] [8],FOCUSS[4] [43,51] and OMP [42][5].

The evaluation process of the separation performances are quantified in terms of separation SNR as follows:

$$SNR_q = 10 \log_{10}(\frac{\sum_{t=1}^{T} \mathbf{s}_t^2(q)}{\sum_{t=1}^{T} \left(\mathbf{s}_t(q) - \hat{\mathbf{s}}_t(q)\right)^2}), \tag{8}$$

where $\mathbf{s}_t(q)$ and $\hat{\mathbf{s}}_t(q)$ are qth ($q = 1 : n$) original and estimated source values in instant t, respectively and SNR_q denotes the separation SNR (the measure of separation performance) for qth source . The y-axis of all evaluation figures (2-7) is the average of separation SNRs over n channels, i.e.:

$$SNR = \frac{\sum_{q=1}^{n} SNR_q}{n} \tag{9}$$

The effects of the additive mixture noise (SNR^X), variance of source noise (Var^S), normalized mean square error (NMSE) of the mixing matrix estimation ($NMSE^A$), sparsity level (k) and dimension ($m \times n$) on the performance are experimentally analysed. In the following, we will clarify the mixture, source noise and the estimation error of mixing matrix .

The mixture noise refers to the noise added to the mixture signals. The amount of additive mixture noise to the rth mixture is quantified by

$$SNR_r^X = 10 \log_{10}(\frac{\sum_{t=1}^{T} x_t^2(r)}{\sum_{t=1}^{T} n_t^2(r)}), \tag{10}$$

where $x_t(r)$ and $n_t(r)$ are the rth ($r = 1 : m$) mixture and noise values in instant t, respectively and SNR_r^X denotes the mixture SNR for rth sensor. In fact, equation (1) is an ideal model and practically we deal with the following model:

$$\tilde{\mathbf{x}}(r) = \mathbf{x}(r) + \mathbf{n}(r), \quad \mathbf{n}(r) \sim \mathcal{N}(0, Var_r^X) \tag{11}$$

where

$$Var_r^X = \frac{1}{T}(\frac{\sum_{t=1}^{T} x_t^2(r)}{10^{(SNR_r^X/10)}}), \tag{12}$$

---

[1] http://ee.sharif.edu/ SLzero/

[2] http://www.cs.ubc.ca/ mpf/spgl1/

[3] https://www.mathworks.com/matlabcentral/fileexchange/48641-sparse-blind-source-separation–sparse-component-analysis-modal-identification

[4] http://dsp.ucsd.edu/ jfmurray/software.htm

[5] https://www.mathworks.com/matlabcentral/fileexchange/32402-cosamp-and-omp-for-sparse-recovery

where T is the number of signal samples, Var_r^X and SNR_r^X are the variance of additive noise and the predefined SNR value for rth sensor.

The sparse sources artifically are generated similar to the employed procedure in [38]. Source noise refers to the noise during source inactive state i.e. how much inactive components are really zero. Indeed, the k active sources in each instant are a zero-mean Gaussian random variable with unit variance and $n - k$ inactive sources in each instant are a zero-mean Gaussian random variable with variance $Var^S \ll 1$.

Let us demonstrate source noise on an example. Consider m mixture signals are generated using equation (1) with k normally distributed sources ($\mathcal{N}(0, 1)$). Ideally, k-sparse vector $\mathbf{s} \in \mathbb{R}^{n \times 1}$ has $n - k$ zero components but practically it's not possible to have exactly zero components. This non-strict sparsity state has been modelled by considering the variance of source noise (Var^S). For example, if the variance of source noise is 0.001 then the inactive sources in each time instant takes the random values with variance 0.001 rather than zero.

Matrix $\mathbf{A}$ is assumed to be known or could be identified using the existing UBI algorithms [23–30]. Note that our algorithm focuses on the source recovery but we add Gaussian noise to $\mathbf{A}$ with a predefined $NMSE^A$ in order to evaluate the sensitivity of algorithms to any error in estimation $\mathbf{A}$:

$$NMSE^A = 10 \log_{10}(\frac{\sum_{ij}(a_{ij} - \hat{a}_{ij})^2}{\sum_{ij} a_{ij}^2}), \qquad (13)$$

where $\hat{a}_{ij}$ and $\hat{a}_{ij}$ are the (i, j)th element of the original ($\mathbf{A}$) and estimated ($\hat{\mathbf{A}}$) mixing matrix.

In the next subsections, we present our experiments and their results. In all experiments, we assume $T = 500$ (the number of signal samples).

5.1 Experiment 1: Effect of mixture noise on separation SNR

In the this experiment, the effect of additive Gaussian noise variance (Var^X) on the separation SNR is considered. We evaluate the separation of the source recovery methods in terms of the 6 levels of SNR^X (10, 20, 30, 40, 50 and 60 dB). To simulate rth ($r = 1 : m$) noisy mixture ($\tilde{\mathbf{x}}(r)$), we add a Gaussian noise signal ($\mathcal{N}(0, Var_r^X)$) to the simulated noiseless mixture ($\mathbf{x}(r)$) according to equations (12) and (13).

Fig. 2 depicts the average separation performance over 100 simulation trials versus SNR^X for our proposed source recovery method (EigMem-SCA) and those of 5 conventional methods. In this simulation, we have assumed $m = 3$, $n = 4$ and $k = 2$ and know the exact $\mathbf{A}$, i.e. the estimation error of the mixing matrix is zero and there is no source noise, i.e. all the inactive sources are exactly zero. It can be seen because of employing a robust subspace matching

process, the proposed algorithm has shown better performance results compared to the conventional methods for all SNR^X. It is important to note that all methods (except BP) has employed the predefined values for noise variance (Var^X) as an input parameter to recover the sources. Although the BPDN method considers the variance of noise, the result is not significantly different from the BP method.
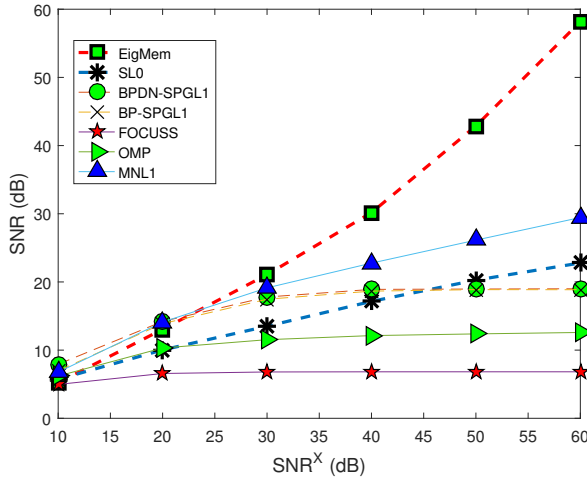

Fig. 2: Effect of mixture noise on separation SNR

5.2 Experiment 2: Effect of source noise on separation SNR

In order to consider more realistic sparse sources, the source noise is modeled using the variance of normal noise over the inactive sources in each time instant. Fig. 3 denotes the average separation SNR over 100 simulation trials in terms of the variance of source noise ($m = 3, n = 4, k = m - 1$, noiseless mixtures and with exact $\mathbf{A}$). It can be seen that the proposed method achieves better performance than the other algorithms due to the membership function assigned to each subspace instead of following an ℓ_1 or ℓ_0 norm minimization framework. Note the standard methods unlike the proposed method yield almost constant results and are not sensitive to the changes of source noise variance. Focusing on k-sparsity rather than ℓ_1- norm (BP, BPDN and ML1), pseudo ℓ_2- norm (FOCUSS) and greedy way (OMP) results the significant difference between our algorithm and others in this scenario. In fact, ℓ_1- norm , pseudo ℓ_2- norm and greedy way have undergone larger changes rather than our method by increasing the source noise specifically in low dimensional problems with $k \lll m$ and they couldn't be a suitable estimation for k-sparsity.
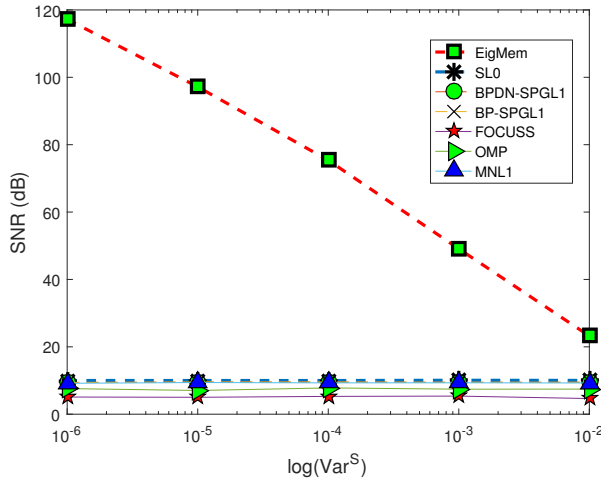
Fig. 3: Effect of source noise on separation SNR

5.3 Experiment 3: Effect of estimation error of the mixing matrix on separation SNR

Fig. 4 shows the average separation SNR of the algorithms versus the estimation error of $\mathbf{A}$ over 100 simulation trials. We have assumed $m = 3$, $n = 4$, $k = 2$ and no mixture and source noise. It can be seen that for $NMSE^A = -60$ to -20 dB, the proposed algorithm achieves better separation SNR than the other algorithms. Albeit the separation SNR of proposed algorithm for $NMSE^A = -10dB$ is not satisfactory due to lack of any remedy process for high mixing estimation error in our proposed algorithm.

5.4 Experiment 4: Effect of sparsity level on separation SNR

Fig. 5 shows the average separation SNR versus the sparsity level (k) over 100 simulation trials for EigMem-SCA algorithm and compares it with those of others. We have assumed $m = 7$, $n = 8$ and $k = 1 : m - 1$ and no mixture, source noise and the mixing estimation error. The results indicate that the proposed method is not sensitive to sparsity level while the separation performance of other algorithms reduces with increasing k. Therefore, the EigMem-SCA algorithm is capable of capturing the problems with densely sparse sources. As mentioned before, the assumptions of standard algorithms are violated in low dimensional problems by approaching k to m. As a result, a significant difference could be seen in Fig. 5 between our algorithm and others, specifically FOCUSS method. It's worthwhile to explain about the huge performance difference between FOCUSS and the proposed algorithm. In FOCUSS method (based on IRLS algorithm) according equation 3, we employ the pseudo ℓ_2-norm which is a trivial modification. It's obvious that when the dimension is
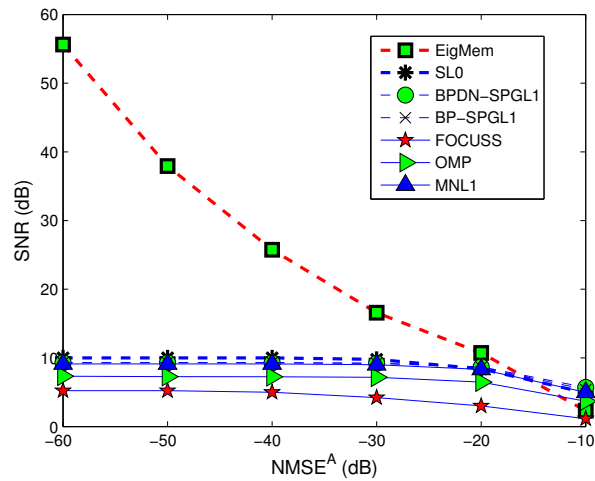
Fig. 4: Effect of the NMSE^A on separation SNR

low and k is close to m, the FOCUSS trick is not able to approximate the sparse solutions and pseudo ℓ_2-norm isn't a good alternative for k-sparsity. FOCUSS method yield the better performance when the $k \ll m$.
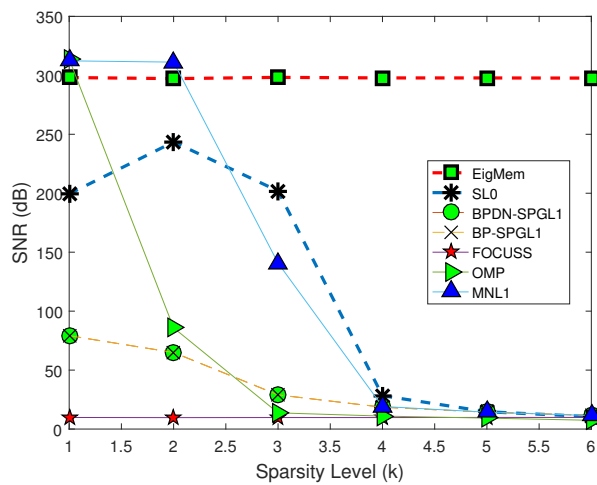

Fig. 5: Effect of sparsity level on separation SNR

5.5 Experiment 5: Effect of dimension on separation SNR

Fig. 6 depicts the average separation performance over 100 simulation trials for different $m \times n$ configurations without any mixture noise, source noise or mixing system error. It is assumed that $k = m - 1$. Observing Fig. 6, in contrast to other algorithms, in the proposed USR algorithm, with increasing the dimension, the separation SNR remains approximately constant.

In a similar way, the experiment 5 has been done considering the mixture noise (SNR^X=30 dB). The results has been denoted in Fig. 7. In this case, the proposed algorithm outperforms others but its performance like other algorithms denotes an decreasing trend by increasing the dimension. For the reason that has been mentioned before in previous subsection, FOCUSS method yields the weak performance in comparison with other methods.
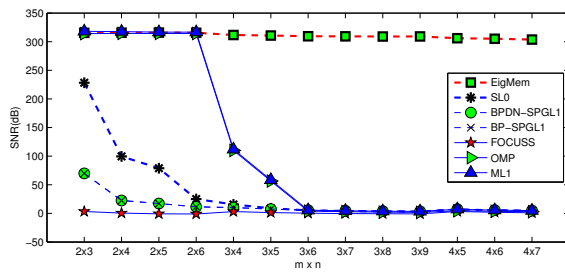

Fig. 6: Effect of $m \times n$ on separation SNR
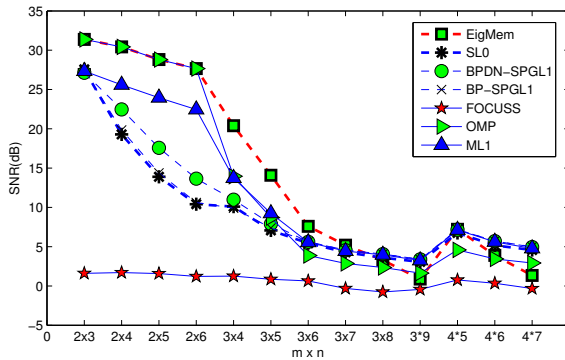(without any mixture noise)

Fig. 7: Effect of $m \times n$ on separation SNR
(with the mixture noise SNR^X=30 dB)

5.6 Computational cost

Our proposed method just like other SCA algorithms suffers from exponential growth in computation cost. Therefore, more work is needed to solve the problem in the large-scale. In fact, EigMem-SCA algorithm has more computational cost than other algorithms. However, we have tried to employ the iterative methods to reduce computational costs. It can also be reduced by using parallel programming, which is one of the future plan. Numerous algorithms are available for computing the eigenvalues and eigenvectors of a matrix. The inverse power method [52] can be used to calculate the smallest eigenvalue. Moreover, Kryolv subspace algorithms [53] could be be employed to find the first k eigenvalues of a general and also large-scale matrix. These algorithms reduce the computational cost from $\mathcal{O}(m^3)$ to $\mathcal{O}(m^2)$ for a $m \times m$ matrix. If we only need the k smallest or largest eigenvalues, the complexity order will be only $\mathcal{O}(k^2 m)$

It's better to explain the computational cost of EigMem-SCA algorithm in more detail. Let be assumed a high SNR scenario with the sparsity level $k = m - 1$. According to Algorithm 1, to find the spanning subspaces in each data sample ($\mathbf{x}_t$), we need to calculate the smallest eigenvalue of $\mathbf{FF}^T \in \mathbb{R}^{m \times m}$ matrix for C_k^n times. Therefore, we need to do a $m \times m$ matrix multiplication and finding the smallest eigenvalue of a $m \times m$ matrix. The matrix multiplication using optimized CW-like algorithm [54] and finding the smallest eigenvalue using Kryolv subspace algorithm require a computational cost of order $\mathcal{O}(m^{2.373})$ and $\mathcal{O}(m)$, respectively. These two operations should be done C_k^n times, which needs a computational cost of order $\mathcal{O}(n^{min(n-k,k)})$. Indeed, finding the spanning subspaces for each data sample requires the computational cost of order $\mathcal{O}(m^{2.373} n^{min(n-k,k)})$, which is more than other standard USR algorithms. For example, OMP algorithm requires only $\mathcal{O}(kmn)$ floating point operations per second (FLOPS).

The high computational cost is extremely restrictive in large-scale problems. Although our experimental scenario in this paper is small-scale problems. Therefore, the average computation time (ACT) could be compared between algorithms. The ACT in seconds over 100 trials for a problem with $T = 500$, $m = 4$, $n = 7$ and $k = 3$ (without any mixture, source noises and mixing matrix error) has been reported in Table 1. In average, EigMem-SCA finds the sparse solution ($k = 3$) of a linear system of 4 equations and 7 unknowns in $0.6280/500 = 1.3ms$. As can be seen from Table 1, FOCUSS is the fastest algorithm and BP-SPGL1 is the slowest algorithm.

6 Conclusion

In this work, we proposed a novel approach for recovering the sources in the underdetermined cases given the mixing matrix based on the k-SCA assumptions. The proposed algorithm considers the possible least relaxation ($k \leqslant m - 1$) to estimate sparse sources and the method is suitable for applications where

Table 1: Average computation time (seconds) over 100 trials
for a problem with $T = 500$, $m = 4$, $n = 7$ and $k = 3$

Algorithm	ACT(seconds)
FOCUSS	0.0459
ML1	0.4580
OMP	0.4809
EigMem-SCA	0.6280
SL0	0.6469
BP-SGL1	82.8522

the sources are densely sparse. The condition $k \leqslant m - 1$ is very promising in comparison to $k \ll m$. The separation performance of most state of the art USR approaches (e.g. the relaxation and greedy families) deteriorate when the sparsity level approaches the number of sensors i.e. $k \to m$ as occurred in low dimensional mixing systems. We tried to modify the k-SCA approaches propsed by Georgiev *et al.* [4]. We showed experimentally that the proposed method achieves better separation performance compared with SL0, BP, BPDN, OPM, FOCUSS and ML1. Moreover, by simulating the low dimensional synthetic mixing system, it was shown that the proposed algorithm (in contrast to the other algorithms) keeps the good separation performance when the sparsity level (k) and also the mixing dimension ($m \times n$) increases. Future works may include making the algorithm robust against noise capable of coping with the low SNR mixing ($SNR \leq 10dB$) scenarios, extending to high dimensional mixing systems. The most widely used analytic tool for recoverability and stability in the sparse recovery methods is restricted isometry property (RIP) leading to the restriction on the mixing matrix ($\mathbf{A}$). Therefore, it is worth researching further in analysing and discussing the extension of proposed algorithm to Non-RIP systems. Interesting avenues of research which we are currently investigating include applying the algorithm to the sparse representation of signals or compressed sensing (CS). Finally, investigating to reduce the computational costs and time (especially using parallel processing) and applying the proposed algorithm to more realistic problems in higher dimensions are the interesting and the practical subjects for the future works.

References

1. Pierre Comon and Christian Jutten. *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.
2. A Cichocki, Y Li, P-G Georgiev, and S Amari. Beyond ICA: Robust sparse signal representations. In *Proceedings of 2004 IEEE International Symposium on Circuits and Systems*, pages 684–687, 2004.
3. P. Georgiev, F. Theis, and A. Cichocki. Sparse component analysis and blind source separation of underdetermined mixtures. *IEEE Transactions on Neural Networks*, 16(4):992–996, July 2005.
4. F Georgiev, Fabian Theis, and Andrzej Cichocki. Blind source separation and sparse component analysis of overcomplete mixtures. In *Acoustics, Speech, and Signal Pro-*

cessing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on, volume 5, pages V–493. IEEE, 2004.

5. Pando Georgiev, Fabian Theis, Andrzej Cichocki, and Hovagim Bakardjian. Sparse component analysis: a new tool for data mining. In *Data mining in biomedicine*, pages 91–116. Springer, 2007.

6. Aapo Hyvärinen, Juha Karhunen, and Erkki Oja. *Independent component analysis*, volume 46. John Wiley & Sons, 2004.

7. Andrzej Cichocki and Shun-ichi Amari. Adaptive blind signal and image processing: learning algorithms and applications. 2002.

8. Pau Bofill and Michael Zibulevsky. Underdetermined blind source separation using sparse representations. *Signal processing*, 81(11):2353–2362, 2001.

9. Stefan Winter, Hiroshi Sawada, Shoko Araki, and Shoji Makino. Overcomplete bss for convolutive mixtures based on hierarchical clustering. In *International Conference on Independent Component Analysis and Signal Separation*, pages 652–660. Springer, 2004.

10. Yuanqing Li, Andrzej Cichocki, and Shun-Ichi Amari. Sparse component analysis for blind source separation with less sensors than sources. In *ICA2003*, pages 89–94. Citeseer, 2003.

11. Michael Zibulevsky and Barak A Pearlmutter. Blind source separation by sparse decomposition in a signal dictionary. *Neural computation*, 13(4):863–882, 2001.

12. Shoko Araki, Shoji Makino, Audrey Blin, Ryo Mukai, and Hiroshi Sawada. Underdetermined blind separation for speech in real environments with sparseness and ICA. In *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*, volume 3, pages iii–881. IEEE, 2004.

13. Mohammad Niknazar, Hanna Becker, Bertrand Rivet, Christian Jutten, and Pierre Comon. Blind source separation of underdetermined mixtures of event-related sources. *Signal Processing*, 101:52–64, 2014.

14. Yuanqing Li, Andrzej Cichocki, and S-I Amari. Blind estimation of channel parameters and source components for EEG signals: a sparse factorization approach. *IEEE Transactions on Neural Networks*, 17(2):419–431, 2006.

15. Yuanqing Li, Zhu Liang Yu, Ning Bi, Yong Xu, Zhenghui Gu, and Shun-ichi Amari. Sparse representation for brain signal processing: a tutorial on methods and applications. *IEEE Signal Processing Magazine*, 31(3):96–106, 2014.

16. Yuanqing Li, Shun-Ichi Amari, Andrzej Cichocki, Daniel WC Ho, and Shengli Xie. Underdetermined blind source separation based on sparse representation. *IEEE Transactions on signal processing*, 54(2):423–437, 2006.

17. Abdeldjalil Aissa-El-Bey, Nguyen Linh-Trung, Karim Abed-Meraim, Adel Belouchrani, and Yves Grenier. Underdetermined blind separation of nondisjoint sources in the time-frequency domain. *IEEE Transactions on Signal Processing*, 55(3):897–907, 2007.

18. Dezhong Peng and Yong Xiang. Underdetermined blind source separation based on relaxed sparsity condition of sources. *IEEE Transactions on Signal Processing*, 57(2):809–814, 2009.

19. Alexander Jourjine, Scott Rickard, and Ozgur Yilmaz. Blind separation of disjoint orthogonal signals: Demixing n sources from 2 mixtures. In *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*, volume 5, pages 2985–2988. IEEE, 2000.

20. Shoko Araki, Hiroshi Sawada, Ryo Mukai, and Shoji Makino. Underdetermined blind sparse source separation for arbitrarily arranged multiple sensors. *Signal Processing*, 87(8):1833–1847, 2007.

21. Frédéric Abrard and Yannick Deville. Blind separation of dependent sources using the" time-frequency ratio of mixtures" approach. In *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, volume 2, pages 81–84. IEEE, 2003.

22. Frédéric Abrard and Yannick Deville. A time–frequency blind signal separation method applicable to underdetermined mixtures of dependent sources. *Signal Processing*, 85(7):1389–1403, 2005.

23. Chu-Jun Yao, Hai-Lin Liu, and Zhi-Tao Cui. Mixing matrix recovery of underdetermined source separation based on sparse representation. In *Computational Intelligence and Security, 2007 International Conference on*, pages 1–5. IEEE, 2007.

24. Jiechang Wen, Hailin Liu, Suxian Zhang, and Mingqing Xiao. A new fuzzy K-EVD orthogonal complement space clustering method. *Neural Computing and Applications*, 24(1):147–154, 2014.

25. Fabian J Theis, Pando G Georgiev, and Andrzej Cichocki. Robust overcomplete matrix recovery for sparse sources using a generalized Hough transform. In *ESANN*, pages 343–348. Citeseer, 2004.

26. Y. Washizawa and A. Cichocki. On-line k-plane clustering learning algorithm for sparse comopnent analysis. In *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, volume 5, pages V–V, May 2006.

27. E. Eqlimi and B. Makkiabadi. Multiple sparse component analysis based on subspace selective search algorithm. In *2015 23rd Iranian Conference on Electrical Engineering*, pages 550–554, May 2015.

28. Ehsan Eqlimi and Bahador Makkiabadi. An efficient K-SCA based unerdetermined channel identification algorithm for online applications. In *Signal Processing Conference (EUSIPCO), 2015 23rd European*, pages 2661–2665. IEEE, 2015.

29. Wen Yang and Hongyi Zhang. Blind source separation based on K-SCA assumption. In *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*, volume 9, pages 116–121, July 2010.

30. Zhaoshui He, Andrzej Cichocki, Yuanqing Li, Shengli Xie, and Saeid Sanei. K-hyperline clustering learning for sparse component analysis. *Signal Processing*, 89(6):1011–1022, 2009.

31. Stéphane G Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on signal processing*, 41(12):3397–3415, 1993.

32. Rémi Gribonval and Pierre Vandergheynst. On the exponential convergence of matching pursuits in quasi-incoherent dictionaries. *IEEE Transactions on Information Theory*, 52(1):255–261, 2006.

33. Deanna Needell and Joel A Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, 2009.

34. Ichigaku Takigawa, Mineichi Kudo, and Jun Toyama. Performance analysis of minimum 1-norm solutions for underdetermined source separation. *IEEE Transactions on Signal Processing*, 52(3):582–591, 2004.

35. David L Donoho, Michael Elad, and Vladimir N Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Transactions on information theory*, 52(1):6–18, 2006.

36. Rémi Gribonval and Morten Nielsen. Sparse representations in unions of bases. *IEEE Transactions on Information Theory*, 49(12):3320–3325, 2003.

37. G Hosein Mohimani, Massoud Babaie-Zadeh, and Christian Jutten. Fast sparse representation based on smoothed 0 norm. In *International Conference on Independent Component Analysis and Signal Separation*, pages 389–396. Springer, 2007.

38. Hosein Mohimani, Massoud Babaie-Zadeh, and Christian Jutten. A fast approach for overcomplete sparse decomposition based on smoothed norm. *IEEE Transactions on Signal Processing*, 57(1):289–301, 2009.

39. Benxu Liu, Vaninirappuputhenpurayil Gopalan Reju, and Andy WH Khong. A linear source recovery method for underdetermined mixtures of uncorrelated AR-model signals without sparseness. *IEEE Transactions on Signal Processing*, 62(19):4947–4958, 2014.

40. Michael Elad. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer Publishing Company, Incorporated, 1st edition, 2010.

41. Farokh Marvasti, Arash Amini, Farzan Haddadi, Mahdi Soltanolkotabi, Babak Hossein Khalaj, Akram Aldroubi, Saeid Sanei, and Janathon Chambers. A unified approach to sparse signal processing. *EURASIP journal on advances in signal processing*, 2012(1):44, 2012.

42. Joel A Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information theory*, 50(10):2231–2242, 2004.

43. Irina F Gorodnitsky and Bhaskar D Rao. Sparse signal reconstruction from limited data using focuss: A re-weighted minimum norm algorithm. *IEEE Transactions on signal processing*, 45(3):600–616, 1997.

44. Scott Shaobing Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM review*, 43(1):129–159, 2001.
45. Axel Målqvist and Daniel Peterseim. Computation of eigenvalues by numerical upscaling. *Numerische Mathematik*, 130(2):337–361, 2015.
46. Babak Hassibi. An efficient square-root algorithm for blast. In *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*, volume 2, pages II737–II740. IEEE, 2000.
47. Dirk Wubben, Ronald Bohnke, Volker Kuhn, and K-D Kammeyer. Mmse extension of v-blast based on sorted qr decomposition. In *Vehicular technology conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, volume 1, pages 508–512. IEEE, 2003.
48. Hadi Zayyani, Massoud Babaie-Zadeh, and Christian Jutten. An iterative bayesian algorithm for sparse component analysis in presence of noise. *IEEE Transactions on Signal Processing*, 57(11):4378–4390, 2009.
49. E. van den Berg and M. P. Friedlander. Probing the Pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, 31(2):890–912, 2008.
50. E. van den Berg and M. P. Friedlander. SPGL1: A solver for large-scale sparse reconstruction, June 2007. http://www.cs.ubc.ca/labs/scl/spgl1.
51. Shane F Cotter, Bhaskar D Rao, Kjersti Engan, and Kenneth Kreutz-Delgado. Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Transactions on Signal Processing*, 53(7):2477–2488, 2005.
52. RV Mises and Hilda Pollaczek-Geiringer. Praktische verfahren der gleichungsauflösung. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 9(1):58–77, 1929.
53. Yousef Saad. Krylov subspace methods for solving large unsymmetric linear systems. *Mathematics of computation*, 37(155):105–126, 1981.
54. François Le Gall. Powers of tensors and fast matrix multiplication. In *Proceedings of the 39th international symposium on symbolic and algebraic computation*, pages 296–303. ACM, 2014.