RESEARCH

Open Access

Decoupled 2D direction-of-arrival estimation based on sparse signal reconstruction

Feng Wang[*], Xiaowei Cui, Mingquan Lu and Zhenming Feng

Abstract

A new two-dimensional direction-of-arrival estimation algorithm called 2D-l_1-singular value decomposition (SVD) and its improved version called enhanced-2D-l_1-SVD are proposed in this paper. They are designed for rectangular arrays and can also be extended to rectangular arrays with faulty or missing elements. The key idea is to represent direction-of-arrival with two decoupled angles and then successively estimate them. Therefore, two-dimensional direction finding can be achieved by applying several times of one-dimensional sparse reconstruction-based direction finding methods instead of directly extending them to two-dimensional situation. Performance analysis and simulation results reveal that the proposed method has a much lower computational complexity and a similar statistical performance compared with the well-known l_1-SVD algorithm, which has several advantages over conventional direction finding techniques due to the application of sparse signal reconstruction. Moreover, 2D-l_1-SVD has better robustness to the assumed number of sources over l_1-SVD.

Keywords: Antenna arrays; Direction-of-arrival estimation; Sparse signal reconstruction

1 Introduction

Direction-of-arrival (DOA) estimation has been playing a significant role in many applications such as radar and wireless communication, and it has been well studied in literature [1,2]. Conventional DOA estimation algorithms can be classified into three broad categories: beamforming [3], subspace-based methods [4,5], and maximum likelihood methods [6]. In the last decade, DOA estimation algorithms based on sparse signal reconstruction (SSR) were proposed. They enforce sparsity on the spatial spectrum and pose source localization as an over-complete basis representation problem. Many of them such as l_1-singular value decomposition (SVD) [7] and JLZA-DOA [8] work directly on the input data while others such as SPICE [9] and SpSF [10] work on the covariance matrix. DOA estimation algorithms based on SSR have several advantages over conventional methods, including increased resolution, better robustness to noise, limitations in data quantity, and correlated sources, as well as not requiring an accurate initialization [7]. Improvements

like [11] which are based on weighted l_1 minimization can get a better performance.

Although one-dimensional (1D) DOA estimation has been widely investigated, two-dimensional (2D) DOA estimation is of greater practical importance. However, most existing 2D DOA estimation methods have met problems of high arithmetic complexity and pair-matching due to a more complicated array manifold [12]. What is more, 2D methods extended from 1D conventional algorithms may still have shortcomings such as requirement of sufficient samples and performance degradation in the presence of correlated sources. In [13], the authors proposed a 2D algorithm whose key idea is to successively apply several times of 1D multiple signal classification (MUSIC) [4] in tree structure. It has a much lower complexity than 2D MUSIC but is limited to uniform rectangular arrays (URAs) and needs extra processes to deal with coherent sources. Sparse methods can also be directly extended to 2D situation, but they will significantly increase the dimension of over-complete basis or the dictionary [14]. A fast orthogonal matching pursuit (OMP) method for 2D angle estimation in multiple-input multiple-output (MIMO) radar has been proposed in [15], which decomposes the 2D dictionary into two sub-dictionaries. However, the bistatic MIMO

*Correspondence: wangfeng12@mails.tsinghua.edu.cn
Department of Electronic Engineering, Tsinghua University, 30 Shuang Qing Lu, 100084 Beijing, China

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 2 of 16

radar considered in [15] consists of a transmit array and a receive array, and the signal at the receiver is analyzed to estimate the transmit angle and the receive angle. So its model differs from 2D DOA estimation. Some 2D DOA estimation algorithms using L-shaped array have also been proposed [16,17]. They may have a low computational complexity and need less antenna elements than those algorithms using rectangular arrays. However, these methods are based on the second-order statistics or the cross-correlation matrix of the received data, so their performances may deteriorate in the presence of correlated sources or insufficient snapshots.

In this paper, a 2D DOA estimation method based on SSR for rectangular arrays called 2D-l_1-SVD is proposed. The key idea is that 2D DOA estimation, which is usually implemented by estimating azimuth and elevation angles jointly, can be accomplished by successively solving several 1D direction finding problems. It is illustrated in this paper that 2D-l_1-SVD reduces the computational load due to the successive parameter estimation and has a similar performance with the direct extension of l_1-SVD [7] to 2D situation. What is more, the other SSR based DOA estimation methods that work directly on the input data, including the weighted version of l_1-SVD [11] or JLZA-DOA [8], can also be extended to 2D situations similar with a much lower complexity than direct extensions. Moreover, it is possible that the array is conformal to follow some prescribed shape or a few elements fail to work in practical applications, and 2D-l_1-SVD also works properly for these nonrectangular applications. Therefore, 2D-l_1-SVD keeps robustness to non-uniform array manifolds. It is illustrated in both theoretical analysis and simulation results that the proposed algorithm performs properly and effectively.

The rest of this paper is organized as follows. Section 2 addresses the problem formulation. Section 3 presents the proposed algorithm. Section 4 investigates the performance of the proposed method and Section 5 demonstrates the simulation results. Section 6 concludes the paper.

In this paper, we use a bold small letter to represent a vector and a bold capital letter to represent a matrix.

2 Problem formulation
2.1 Array configuration
Firstly, consider a rectangular planar array that consists of $M \times N$ omnidirectional and well-calibrated antenna elements. The coordinates (x_{mn}, y_{mn}) of the antenna element at the mth row and nth column ($1 \leq n \leq M, 1 \leq n \leq N$) satisfy:

$$\begin{cases} x_{m1} = \ldots = x_{mn} = \ldots = x_{mN} \\ y_{1n} = \ldots = y_{mn} = \ldots = y_{Mn} \end{cases}, \quad (1)$$

so they are denoted as (x_m, y_n) and the element at the mth row and nth column is indexed as (m, n). The rectangular array may be non-uniform.
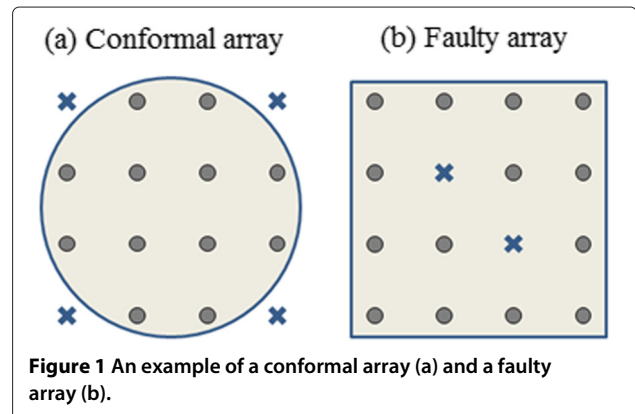
Although rectangular arrays, including uniform rectangular arrays (URA), are quite common in practical applications, it is possible that the array is conformal to follow some prescribed shape or some of the elements fail to work, as illustrated in Figure 1, where the valid elements are denoted as circles in the area rounded up by the deep blue line while the blue 'X' denotes missing or faulty elements (similarly hereinafter). So the array manifold is no longer rectangular, and it turns out to be a sub-array of the rectangular array.
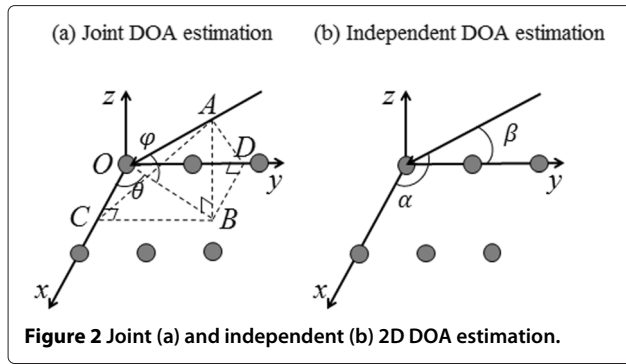
In order to deal with irregular arrays mentioned above, consider the rectangular array that consists of $M \times N$ antenna elements in Equation 1. And it is assumed that there are D invalid (missing or faulty) sensors indexed as (m_d, n_d), $d = 1, \cdots, D$, where m_d and n_d are the row and column number of the dth invalid sensor, correspondingly. The location of all invalid sensors is supposed to be known *a priori* in this paper.

2.2 Representation of 2D DOA
Most existing 2D DOA algorithms try to estimate the azimuth and elevation angle (θ, φ), which is $(\angle BOC, \angle AOB)$ in Figure 2a. However, we consider another definition of 2D DOA [16] in this paper, as given in Figure 2b or $(\angle AOC, \angle AOD)$ in Figure 2a, where the purpose of direction finding is to estimate the two angles (α, β) between the incoming signal and x-axis or y-axis, respectively. It is assumed that the signals come from above the antenna, as shown in Figure 2. Since $OC \perp BC$ and $OC \perp AB$, then we have $OC \perp AC$. As a result, $OC = OA \cdot \cos \alpha = OB \cdot \cos \theta = OA \cdot \cos \varphi \cos \theta$. Similarly, $OD = OA \cdot \cos \beta = OB \cdot \sin \theta = OA \cdot \cos \varphi \sin \theta$. Therefore, there exists a correspondence between (θ, φ) and (α, β):

$$\begin{cases} \cos \alpha = \cos \varphi \cos \theta \\ \cos \beta = \cos \varphi \sin \theta \end{cases}. \quad (2)$$

Figure 1 An example of a conformal array (a) and a faulty array (b).

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 3 of 16

Figure 2 Joint (a) and independent (b) 2D DOA estimation.

In this paper, we focus on the narrowband DOA estimation problem and the noise signals are assumed to be Gaussian additive noises. The incoming signals can be correlated or even coherent with each other. Using the narrowband model, we get the digital vector $y(k) = [y_1(k), \cdots, y_{MN-D}(k)]^T \in \mathbb{C}^{(MN-D)\times 1}$ of complex amplitudes of the sensors at time instant k bellow (the superscript T denotes the transpose operation):

$$y(k) = As(k) + n(k), \qquad (3)$$

where it is assumed that there are P narrowband far-field signals $s(k) = [s_1(k), \cdots, s_P(k)]^T \in \mathbb{C}^{P\times 1}$. $n(k) = [n_1(k), \cdots, n_{MN-D}(k)]^T \in \mathbb{C}^{(MN-D)\times 1}$ is the noise vector and $A = [a(\alpha_1, \beta_1), \cdots, a(\alpha_P, \beta_P)] = [a(\theta_1, \varphi_1), \cdots, a(\theta_P, \varphi_P)] \in \mathbb{C}^{(MN-D)\times P}$ is the array manifold matrix whose columns are comprised of P manifold vectors. According to Equation 2, the element corresponding to a valid sensor at the mth row and nth column in the manifold vector $a(\alpha_p, \beta_p)$ is:

$$
\begin{aligned}
a_{m,n}(\alpha_p, \beta_p) &= a_{m,n}(\theta_p, \varphi_p) \\
&= e^{-j2\pi(x_m \cos\theta_p + y_n \sin\theta_p)\cos\varphi_p/\lambda} \\
&= e^{-j2\pi x_m \cos\alpha/\lambda} e^{-j2\pi y_n \cos\beta/\lambda}, \qquad (4)
\end{aligned}
$$

where the exponential term is written as the sum of factors related to α and β, respectively, hence the two parameters to be estimated can be decoupled from each other.

If there are no invalid elements, the complex amplitudes of sensors are $y(k) \triangleq [y_{1,1}(k), \cdots, y_{M,1}(k), \cdots, y_{1,N}(k), \cdots, y_{M,N}(k)]^T \in \mathbb{C}^{MN\times 1}$ and we have

$$y(k) = [a(\alpha_1, \beta_1), \cdots, a(\alpha_P, \beta_P)] s(k) + n(k), \quad (5)$$

where $n(k) = [n_{1,1}(k), \cdots, n_{M,1}(k), \cdots, n_{1,N}(k), \cdots, n_{M,N}(k)]^T \in \mathbb{C}^{MN\times 1}$, and $a(\alpha_p, \beta_p) = [a_{1,1}(\alpha_p, \beta_p), \cdots, a_{M,1}(\alpha_p, \beta_p), \cdots, a_{1,N}(\alpha_p, \beta_p), \cdots, a_{M,N}(\alpha_p, \beta_p)]^T \in \mathbb{C}^{MN\times 1}$ is the manifold vector.

According to (4), the manifold vector can be rewritten as the Kronecker product of two vectors:

$$a(\alpha_p, \beta_p) = b^L(\beta_p) \otimes a^L(\alpha_p), \qquad (6)$$

where $a^L(\alpha_p) \triangleq [e^{-j2\pi x_1 \cos\alpha_p/\lambda}, \cdots, e^{-j2\pi x_M \cos\alpha_p/\lambda}]^T \in \mathbb{C}^{M\times 1}$ and $b^L(\beta_p) \triangleq [e^{-j2\pi y_1 \cos\beta_p/\lambda}, \cdots, e^{-j2\pi y_N \cos\beta_p/\lambda}]^T \in {}^{N\times 1}$ denote the steering vectors of the linear sub-arrays that lie on the x-axis and y-axis, respectively. The superscript L denotes 'linear'. Thus, we can rewrite the signals received at the array in a decoupled form:

$$Y(k) = A^L X(k) + N(k), \qquad (7)$$

where $Y(k) \triangleq \begin{bmatrix} y_{1,1}(k) & \cdots & y_{1,N}(k) \\ \vdots & \ddots & \vdots \\ y_{M,1}(k) & \cdots & y_{M,N}(k) \end{bmatrix} \in \mathbb{C}^{M\times N}$ is the matrix form of $y(k)$, $A^L \triangleq [a^L(\alpha_1), \cdots, a^L(\alpha_P)] \in \mathbb{C}^{M\times P}$ denotes manifold matrix whose columns are comprised of P manifold vectors of the linear sub-array on the x-axis,

$$X(k) \triangleq \begin{bmatrix} s_1(k)(b^L(\beta_1))^T \\ \vdots \\ s_P(k)(b^L(\beta_P))^T \end{bmatrix} \in \mathbb{C}^{P\times N}$$ denotes the signal matrix in the decoupled signal model, and $N(k) \triangleq \begin{bmatrix} n_{1,1}(k) & \cdots & n_{1,N}(k) \\ \vdots & \ddots & \vdots \\ n_{M,1}(k) & \cdots & n_{M,N}(k) \end{bmatrix} \in \mathbb{C}^{M\times N}$ is the matrix form of $n(k)$.

Although the 2D DOA (α_p, β_p) of multiple sources are different from each other, it is possible that some of the incoming sources have the same α and there exist multiple identical columns in the manifold matrix A^L. Therefore, in order to make sure that all columns in the manifold matrix A^L are different from each other, the 2D DOA of incoming sources can be denoted as $(\alpha_1, \beta_{1,1}), \cdots, (\alpha_1, \beta_{1,f_1}), \cdots, (\alpha_P, \beta_{P,1}), \cdots, (\alpha_P, \beta_{P,f_P})$, where $f_1, \cdots, f_P$ are positive integers. So the total number of incoming sources is $P_0 \triangleq \sum_{p=1}^{P} f_p$. Correspondingly, the incoming signals are denoted as $s(k) = [s_{1,1}(k), \cdots, s_{1,f_1}(k), \cdots, s_{P,1}(k), \cdots, s_{P,f_P}(k)]^T \in \mathbb{C}^{P_0\times 1}$. Therefore, the signal matrix $X(k)$ in Equation 7 becomes

$$X(k) \triangleq \begin{bmatrix} \sum_{i=1}^{f_1} s_{1,i}(k)(b^L(\beta_{1,i}))^T \\ \vdots \\ \sum_{i=1}^{f_P} s_{P,i}(k)(b^L(\beta_{P,i}))^T \end{bmatrix} \in \mathbb{C}^{P\times N}.$$

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 4 of 16

From Equation 7, it can be observed that the information of α and β is in the column and row spaces of $Y(k)$, respectively. The manifold matrix A^L is decided by α and is not related to β, while the signal matrix $X(k)$ is determined by β. Equation 7 can also be explained in terms of the rectangular manifold. Alpha can be estimated by analyzing the samples of the sub-array that lies on the x-axis and consists of sensors at $\{(x_m, y_1)\}$, $m = 1, \cdots, M$. The linear sub-arrays that are parallel to x-axis in the rectangular array share the same manifold when only the estimation of α is taken into account. Therefore, in order to estimate α, each snapshot received by the rectangular array can be regarded as N correlated snapshots of the linear sub-array that lies on the x-axis. Then, 2D DOA can be estimated by solving two 1D DOA estimation problems successively instead of directly estimating 2D DOA. Since α is decoupled from β, we can also estimate β firstly and then estimate α. What is more, we will show in Section 3 that integrating the results of these two problems helps to improve the performance.

Now we still consider the rectangular array whose elements are all valid and concentrate on the elements indexed (m_d, n_d), $d = 1, \cdots, D$. Obviously, we have:

$$
\begin{bmatrix}
y_{1,1}(k) & \cdots & \cdots & \cdots & y_{1,N}(k) \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & y_{m_d,n_d}(k) & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
y_{M,1}(k) & \cdots & \cdots & \cdots & y_{M,N}(k)
\end{bmatrix}
-
\begin{bmatrix}
0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \ddots & 0 & 0 \\
\vdots & \ddots & y_{m_d,n_d}(k) & \ddots & \vdots \\
0 & 0 & \ddots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0
\end{bmatrix}
$$

$$
= A^L X(k) +
\begin{bmatrix}
n_{1,1}(k) & \cdots & \cdots & \cdots & n_{1,N}(k) \\
\vdots & \ddots & & \ddots & \vdots \\
\vdots & \ddots & n_{m_d,n_d}(k) - y_{m_d,n_d}(k) & \ddots & \vdots \\
\vdots & \ddots & & \ddots & \vdots \\
n_{M,1}(k) & \cdots & \cdots & \cdots & n_{M,N}(k)
\end{bmatrix},
\tag{8}
$$

which can be rewritten as:

$$
Y^f = A^L X(k) +
\begin{bmatrix}
n_{1,1}(k) & \cdots & \cdots & \cdots & n_{1,N}(k) \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & n^f_{m_d,n_d}(k) & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
n_{M,1}(k) & \cdots & \cdots & \cdots & n_{M,N}(k)
\end{bmatrix}.
\tag{9}
$$

where $Y^f = \begin{bmatrix}
y_{1,1}(k) & \cdots\cdots\cdots & y_{1,N}(k) \\
\vdots & \ddots\ddots\ddots & \vdots \\
\vdots & \ddots\ 0\ \ddots & \vdots \\
\vdots & \ddots\ddots\ddots & \vdots \\
y_{M,1}(k) & \cdots\cdots\cdots & y_{M,N}(k)
\end{bmatrix}.$

In Equation 9, the received signals at the elements indexed (m_d, n_d), $d = 1, \cdots, D$ are disregarded and set to zeros while the noise signals at these sensors change into $n^f_{m_d,n_d}(k) = n_{m_d,n_d}(k) - y_{m_d,n_d}(k)$. The superscript f denotes 'faulty'. The case of faulty sensors can be processed in the above form. If there are D invalid elements, then no restriction is placed on the range of $y_{m_d,n_d}(k)$, $d = 1, \cdots, D$ since there are no valid samples at these invalid sensors. As a result, $n^f_{m_d,n_d}(k)$, $d = 1, \cdots, D$ in Equation 9 is unconstrained. In this way, the complex amplitudes of the sensors at a faulty rectangular array can still be written in the form of the decoupled signal model (Equation 9). However, it should be taken into account that the noise signals at the faulty sensors are unconstrained and no longer distribute as noise signals at valid sensors. As a result, when the decoupled signal model is exploited, it should be noticed that the received signals at the faulty elements are set to be zero and the noise signals at these faulty elements are totally unknown and do not provide any additional information.

3 Proposed algorithm

Based on the above decoupled model, 2D DOA estimation can be achieved by two steps. Firstly, solve a 1D direction finding problem to estimate the first parameter α and get the information about the second parameter β. This process can be accomplished by 1D DOA estimator directly working on the input data instead of using the covariance estimator, and information about the second parameter (signal matrix in Equation 7) can be obtained as well. Secondly, solve another several 1D direction finding problems to estimate β based on the rows extracted from the signal matrix and the corresponding α can be obtained from the row number.

DOA estimation methods based on SSR have been found to have several advantages over conventional direction finding methods and many of them work directly on the input data. Here, we pick l_1-SVD [7] as our 1D DOA estimation algorithm and propose a new 2D DOA estimation method called 2D-l_1-SVD. Exploiting the decoupled signal model, 2D-l_1-SVD has a much lower computational load than l_1-SVD due to the successive estimation of 2D DOA parameters. Moreover, an improved version called enhanced-2D-l_1-SVD is also proposed in this section to deal with multiple sources that are close to each other in α or β domain but well separated in the other domain.

Obviously, l_1-SVD does not need specific array shape. However, rectangular arrays are more suitable to exploit the decoupled signal model. As a result, although techniques such as manifold separation [18] can be used for modeling the steering vector of antenna arrays with practical interest with arbitrary geometry, we only consider rectangular arrays and faulty rectangular arrays in this

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 5 of 16

paper. On the other hand, it is obvious that JLZA-DOA [8] and the other sparse reconstruction based methods that work directly on the input data can also be extended to 2D situation using the aforementioned decoupled signal model similarly.

3.1 2D-l_1-SVD algorithm

The steps of the 2D-l_1-SVD algorithm are illustrated in Figure 3.

The notations that are used in this section are given in Table 1. For each notation, the superscript denotes the description of the variable and the subscript denotes the index of the variable.

Step 1: Exploiting the decoupled signal model

The multiple snapshots received at the antenna array can be denoted as a $MN \times T$ data matrix $\boldsymbol{Y}^{\text{ant}} \triangleq \left[\boldsymbol{y}^{\text{ant}}(1), \cdots, \boldsymbol{y}^{\text{ant}}(T)\right] \in \mathbb{C}^{MN \times T}$ where T is the number of the snapshots. Here, $\boldsymbol{y}^{\text{ant}}(k) \triangleq \left[y_{1,1}^{\text{ant}}(k), \cdots, y_{1,N}^{\text{ant}}(k), \cdots, y_{M,1}^{\text{ant}}(k), \cdots, y_{M,N}^{\text{ant}}(k)\right]^{\text{T}} \in \mathbb{C}^{MN \times 1}$ denotes the signal sampled at time instant k ($1 \leq k \leq T$). The superscript ant denotes 'antenna.' According to the narrowband signal model, we have:

$$\boldsymbol{Y}^{\text{ant}} = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{N}^{\text{ant}}, \tag{10}$$

where $\boldsymbol{A} = \left[\boldsymbol{a}\left(\alpha_1, \beta_{1,1}\right), \cdots, \boldsymbol{a}\left(\alpha_1, \beta_{1,f_1}\right), \cdots, \boldsymbol{a}\left(\alpha_P, \beta_{P,1}\right), \cdots, \boldsymbol{a}\left(\alpha_P, \beta_{P,f_P}\right)\right] \in \mathbb{C}^{MN \times P_0}$ is the manifold matrix of P_0 incoming signals, $\boldsymbol{S} = \left[\boldsymbol{s}(1), \cdots, \boldsymbol{s}(T)\right] \in \mathbb{C}^{P_0 \times T}$ is the signal matrix of the incoming signals and $\boldsymbol{N}^{\text{ant}} = \left[\boldsymbol{n}^{\text{ant}}(1), \cdots, \boldsymbol{n}^{\text{ant}}(T)\right] \in \mathbb{C}^{MN \times T}$ is the noise-signal matrix at the sensor array. Here, $\boldsymbol{s}(k) = [s_{1,1}(k), \cdots, s_{1,f_1}(k), \cdots, s_{P,1}(k), \cdots, s_{P,f_P}(k)]^{\text{T}} \in \mathbb{C}^{P_0 \times 1}$ and $\boldsymbol{n}^{\text{ant}}(k) = [n_{1,1}^{\text{ant}}(k), \cdots, n_{1,N}^{\text{ant}}(k), \cdots, n_{M,1}^{\text{ant}}(k), \cdots, n_{M,N}^{\text{ant}}(k)]^{\text{T}} \in \mathbb{C}^{MN \times 1}$ denote the incoming signals and the noise signals sampled at time instant k ($1 \leq k \leq T$), respectively. As

mentioned before, faulty arrays can be treated in a similar way as a complete rectangular array. For rectangular arrays with faulty or missing elements, the manifold vector $\boldsymbol{a}(\alpha_p, \beta_{p,i})$ ($1 \leq i \leq f_p$) refers to the manifold vector of the complete rectangular array and the data matrix $\boldsymbol{Y}^{\text{ant}}$ satisfies $y_{m_d,n_d}(k) = 0$, $d = 1, \cdots, D$, $k = 1, \cdots, T$. Also, it should be noticed that noise signals at the faulty sensors $n_{m_d,n_d}(k)$ are unconstrained.

For practical direction finding problems, we use the SVD of the data matrix to reduce both the computational complexity and the sensitivity to noise, just like l_1-SVD [7]. The SVD of the data matrix is $\boldsymbol{Y}^{\text{ant}} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^{\text{H}}$, where the superscript H denotes the conjugate transpose operation. Therefore, the data matrix is decomposed into the signal and noise subspaces. Then, the signal subspace that contains most of the signal power is kept to reduce the dimension. Let $\boldsymbol{D}_K = \left[\boldsymbol{I}_K, \boldsymbol{0}_{K \times (T-K)}\right]^{\text{T}}$ where $\boldsymbol{I}_K$ is a $K \times K$ identity matrix. Here, K denotes the assumed number of sources and does not need to be equal to the actual number of sources P_0. And it is indicated in [7] that l_1-SVD maintains robustness to the assumed number of sources. Then, let $\boldsymbol{Y}^{\text{sg}} = \boldsymbol{Y}^{\text{ant}}\boldsymbol{V}\boldsymbol{D}_K \triangleq \left[\boldsymbol{y}^{\text{sg}}(1), \cdots, \boldsymbol{y}^{\text{sg}}(K)\right] \in \mathbb{C}^{MN \times K}$, $\boldsymbol{S}^{\text{sg}} = \boldsymbol{S}\boldsymbol{V}\boldsymbol{D}_K = \left[\boldsymbol{s}^{\text{sg}}(1), \cdots, \boldsymbol{s}^{\text{sg}}(K)\right] \in \mathbb{C}^{P_0 \times K}$ and $\boldsymbol{N}^{\text{sg}} = \boldsymbol{N}^{\text{ant}}\boldsymbol{V}\boldsymbol{D}_K \triangleq \left[\boldsymbol{n}^{\text{sg}}(1), \cdots, \boldsymbol{n}^{\text{sg}}(K)\right] \in \mathbb{C}^{MN \times K}$ where the superscript sg denotes 'signal'. More specifically, $\boldsymbol{y}^{\text{sg}}(k) \triangleq \left[y_{1,1}^{\text{sg}}(k), \cdots, y_{1,N}^{\text{sg}}(k), \cdots, y_{M,1}^{\text{sg}}(k), \cdots, y_{M,N}^{\text{sg}}(k)\right]^{\text{T}} \in \mathbb{C}^{MN \times 1}$, $\boldsymbol{s}^{\text{sg}}(k) = \left[s_{1,1}^{\text{sg}}(k), \cdots, s_{1,f_1}^{\text{sg}}(k), \cdots, s_{P,1}^{\text{sg}}(k), \cdots, s_{P,f_P}^{\text{sg}}(k)\right]^{\text{T}} \in \mathbb{C}^{P_0 \times 1}$ and $\boldsymbol{n}^{\text{sg}}(k) \triangleq \left[n_{1,1}^{\text{sg}}(k), \cdots, n_{1,N}^{\text{sg}}(k), \cdots, n_{M,1}^{\text{sg}}(k), \cdots, n_{M,N}^{\text{sg}}(k)\right]^{\text{T}} \in \mathbb{C}^{MN \times 1}$. Therefore:

$$\boldsymbol{Y}^{\text{sg}} = \boldsymbol{A}\boldsymbol{S}^{\text{sg}} + \boldsymbol{N}^{\text{sg}}. \tag{11}$$

The manifold vector of the rectangular array can be rewritten as Kronecker product of two manifold vectors
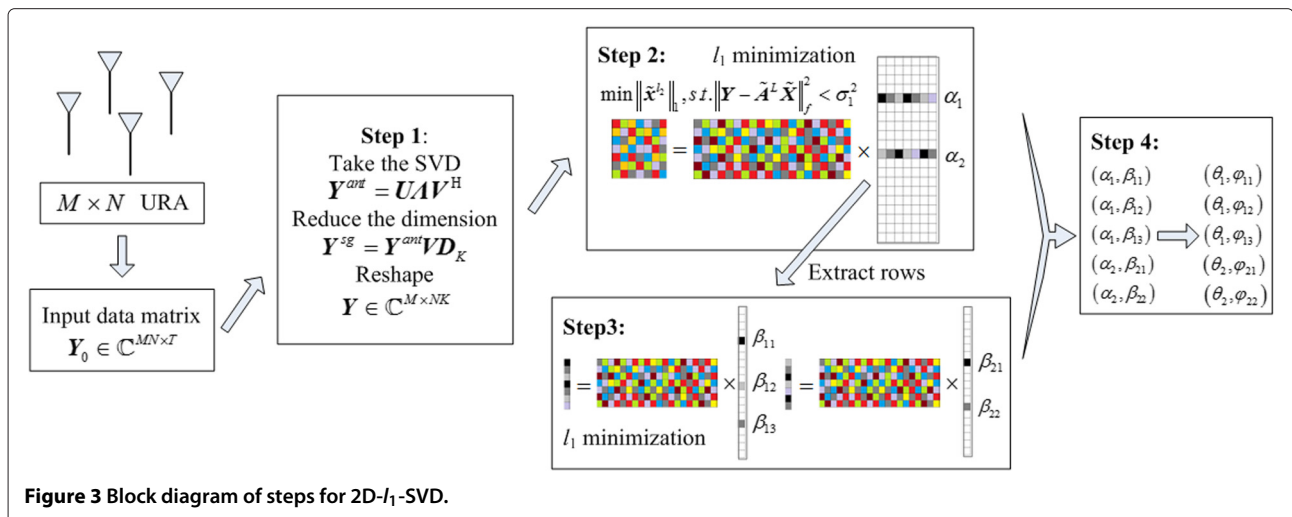

Figure 3 Block diagram of steps for 2D-l_1-SVD.

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 6 of 16

Table 1 Notations used in section 3.1

Notation	Size	Definition
MN	1×1	The number of all the elements in the rectangular array
D	1×1	The number of faulty elements in the rectangular array
P_0	1×1	The number of incoming sources
P	1×1	The number of different α
T	1×1	The number of snapshots
K	1×1	The assumed number of incoming sources
$\boldsymbol{Y}^{\text{ant}}$	$MN \times T$	The data matrix received at the array
$\boldsymbol{A}$	$MN \times P_0$	The manifold matrix of the rectangular array
$\boldsymbol{S}$	$P_0 \times T$	The signal matrix of incoming sources
$\boldsymbol{N}^{\text{ant}}$	$MN \times T$	The noise-signal matrix at the array
$\boldsymbol{Y}^{\text{sg}}$	$MN \times K$	The reduced matrix containing most of the signal power
$\boldsymbol{S}^{\text{sg}}$	$P_0 \times K$	The reduced signal matrix
$\boldsymbol{N}^{\text{sg}}$	$MN \times K$	The reduced noise signal matrix
$\boldsymbol{Y}$	$M \times NK$	The rearrangement of $\boldsymbol{Y}^{sg}$
$\boldsymbol{A}^L$	$M \times P$	The manifold matrix of the linear sub-array on x-axis
$\boldsymbol{X}$	$P \times NK$	The signal matrix in the decoupled model
$\boldsymbol{N}$	$M \times NK$	The rearrangement of $\boldsymbol{N}^{sg}$
$\tilde{\boldsymbol{A}}^L$	$M \times K_\alpha$	The dictionary of manifold vector of the linear sub-array on x-axis
$\tilde{\boldsymbol{X}}$	$K_\alpha \times NK$	The sparse signal matrix in the SSR in the α domain
$\tilde{\boldsymbol{B}}^L$	$N \times K_\beta$	The dictionary of manifold vector of the linear sub-array on y-axis
$\tilde{\boldsymbol{S}}_i$	$K_\beta \times K$	The sparse signal matrix in the SSR in the β domain

corresponding to the linear subarrays, i.e., $\boldsymbol{a}\left(\alpha_p, \beta_{p,i}\right) = \boldsymbol{b}^L\left(\beta_{p,i}\right) \otimes \boldsymbol{a}^L\left(\alpha_p\right)$ (see Equation 6). Then, expressing the equation above in a matrix form similarly as Equation 7, we have:

$$\boldsymbol{Y} = \boldsymbol{A}^L \boldsymbol{X} + \boldsymbol{N}, \qquad (12)$$

where $\boldsymbol{Y} \stackrel{\Delta}{=} [\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_K] \in \mathbb{C}^{M \times NK}, \boldsymbol{N} \stackrel{\Delta}{=} [\boldsymbol{N}_1, \cdots, \boldsymbol{N}_K] \in \mathbb{C}^{M \times NK}$ and:

$$\boldsymbol{X} = \begin{bmatrix} \sum\limits_{i=1}^{f_1} s_{1,i}^{\text{sg}}(1)\left(\boldsymbol{b}^L\left(\beta_{1,i}\right)\right)^{\text{T}} & \cdots & \sum\limits_{i=1}^{f_1} s_{1,i}^{\text{sg}}(K)\left(\boldsymbol{b}^L\left(\beta_{1,i}\right)\right)^{\text{T}} \\ \vdots & \ddots & \vdots \\ \sum\limits_{i=1}^{f_P} s_{P,i}^{\text{sg}}(1)\left(\boldsymbol{b}^L\left(\beta_{P,i}\right)\right)^{\text{T}} & \cdots & \sum\limits_{i=1}^{f_P} s_{P,i}^{\text{sg}}(K)\left(\boldsymbol{b}^L\left(\beta_{P,i}\right)\right)^{\text{T}} \end{bmatrix}$$
$$\in \mathbb{C}^{P \times NK}. \qquad (13)$$

$\boldsymbol{A}^L = \left[\boldsymbol{a}^L\left(\alpha_1\right), \cdots, \boldsymbol{a}^L\left(\alpha_P\right)\right] \in \mathbb{C}^{M \times P}$ is the manifold matrix of the linear sub-array on the x-axis and no longer depends on β-DOA while the information of β-DOA is now contained in the signal matrix $\boldsymbol{X}$. And:

$$\boldsymbol{Y}_k = \begin{bmatrix} y_{1,1}^{\text{sg}}(k) & \cdots & y_{1,N}^{\text{sg}}(k) \\ \vdots & \ddots & \vdots \\ y_{M,1}^{\text{sg}}(k) & \cdots & y_{M,N}^{\text{sg}}(k) \end{bmatrix} \in \mathbb{C}^{M \times N},$$

$$\boldsymbol{N}_k = \begin{bmatrix} n_{1,1}^{\text{sg}}(k) & \cdots & n_{1,N}^{\text{sg}}(k) \\ \vdots & \ddots & \vdots \\ n_{M,1}^{\text{sg}}(k) & \cdots & n_{M,N}^{\text{sg}}(k) \end{bmatrix} \in \mathbb{C}^{M \times N}.$$

It is easy to see that $y_{m_d,n_d}^{\text{sg}}(k) = 0$ and $n_{m_d,n_d}^{\text{sg}}(k)$ is unconstrained for $d = 1, \cdots, D, k = 1, \cdots, T$.

Step 2: Estimate α from multiple measurement vectors (MMV)

An overcomplete representation of $\boldsymbol{A}^L$ in terms of all possible α, which is denoted as $\tilde{\boldsymbol{A}}^L$, is introduced here. Let $\left\{\tilde{\alpha}_1, \cdots, \tilde{\alpha}_{K_\alpha}\right\}$ be a sampling grid of all directions of interest in the α domain. The number of potential directions in the α domain K_α is typically much greater than the number of different α, i.e., $K_\alpha \gg P$. Then, $\tilde{\boldsymbol{A}}^L$ is composed of steering vectors corresponding to each potential α as its columns. Here, the steering vectors in $\tilde{\boldsymbol{A}}^L$ correspond to the linear sub-array on the x-axis. Therefore, Equation 12 can be rewritten in a sparse reconstruction form:

$$\boldsymbol{Y} = \tilde{\boldsymbol{A}}^L \tilde{\boldsymbol{X}} + \boldsymbol{N}, \qquad (14)$$

where $\tilde{\boldsymbol{A}}^L \stackrel{\Delta}{=} \left[\boldsymbol{a}^L\left(\tilde{\alpha}_1\right), \cdots, \boldsymbol{a}^L\left(\tilde{\alpha}_{K_\alpha}\right)\right] \in \mathbb{C}^{M \times K_\alpha}$ and $\tilde{\boldsymbol{X}} \stackrel{\Delta}{=} \begin{bmatrix} \tilde{\boldsymbol{x}}_1^{\text{T}} \\ \vdots \\ \tilde{\boldsymbol{x}}_{K_\alpha}^{\text{T}} \end{bmatrix} \in \mathbb{C}^{K_\alpha \times NK}$. The k_αth row of the signal matrix $\tilde{\boldsymbol{X}}$ is:

$$\tilde{\boldsymbol{x}}_{k_\alpha}^{\text{T}} \stackrel{\Delta}{=} \left[\sum_{i=1}^{\tilde{f}_{k_\alpha}} \tilde{s}_{k_\alpha,i}^{\text{sg}}(1)\left(\boldsymbol{b}^L\left(\tilde{\beta}_{k_\alpha,i}\right)\right)^{\text{T}}, \cdots, \sum_{i=1}^{\tilde{f}_{k_\alpha}} \tilde{s}_{k_\alpha,i}^{\text{sg}}(K)\left(\boldsymbol{b}^L\left(\tilde{\beta}_{k_\alpha,i}\right)\right)^{\text{T}}\right],$$
$$(15)$$

where $\tilde{f}_{k_\alpha}, k_\alpha = 1, \cdots, K_\alpha$ are positive integers and $\tilde{f}_{k_\alpha} = f_p$ if $\tilde{\alpha}_{k_\alpha} = \alpha_p$. The source signal $\tilde{s}_{k_\alpha,i}^{\text{sg}}(k)$ is nonzero and equal to $s_{p,i}^{\text{sg}}(k)$ if $\tilde{\alpha}_{k_\alpha} = \alpha_p$ and zero otherwise. So each column in $\tilde{\boldsymbol{X}}$ is sparse.

As a result, the overcomplete representation in Equation 14 allows us to exchange the estimation of α for the problem of sparse spectrum of each column in $\tilde{\boldsymbol{X}}$, which can be solved via regularizing it to favor sparse signal fields using the l_1 methodology. In the case that there are no invalid

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 7 of 16

elements, estimation of α can be accomplished by solving a MMV problem based on l_1-norm minimization:

$$\min_{\tilde{X}} \left\| \tilde{x}^{l2} \right\|_1, s.t. \left\| Y - \tilde{A}^L \tilde{X} \right\|_F^2 < \sigma_1^2, \tag{16}$$

where the l_1-term enforces sparsity of the representation. $\tilde{x}^{l2} \triangleq \left(\left\| \tilde{x}_1 \right\|_2, \cdots, \left\| \tilde{x}_{K_\alpha} \right\|_2 \right) \in \mathbb{C}^{1 \times K_\alpha}$ and $\|\cdot\|_F$ denotes the Frobenius norm which is defined as $\left\| Y - \tilde{A}^L \tilde{X} \right\|_F^2 = \left\| \text{vec}\left(Y - \tilde{A}^L \tilde{X} \right) \right\|_2^2$. Let $\hat{X} \in \mathbb{C}^{K_\alpha \times NK}$ denote the result of the above MMV problem. Then, the pseudo spectrum of α is obtained by calculating the l_2-norm of each row in the signal matrix $\hat{X}$, i.e., the amplitude of the pseudo spectrum at $\tilde{\alpha}_{k_\alpha}$ is $\left\| \hat{x}_{k_\alpha} \right\|_2$.

In the MMV problems based on l_1-norm minimization mentioned above, l_1-term enforces sparsity while l_2-term forces the residual to be small. The residual specifies how much noise we wish to allow. For rectangular arrays with invalid elements, the noise signals or residuals of invalid elements are unknown while the residuals of valid elements are forced to be small. Hence, the constrained condition of the l_2-term should be revised, taking into consideration of invalid elements. On the other hand, the spatial sparsity still exists, so l_1-term remains unchanged.

When there are invalid elements, the columns of Y in the above MMV problem can be regarded as multiple snapshots of a linear subarray, in which there exists invalid data. Residual at these invalid elements should be unconstrained since no information of these elements is known *a priori*. Let $N^{\text{res}} \triangleq Y - \tilde{A}^L \tilde{X} \in \mathbb{C}^{M \times NK}$. Therefore, estimation of α can be accomplished by solving a modified MMV problem based on l_1-norm minimization:

$$\min_{\tilde{X}} \left\| \tilde{x}^{l2} \right\|_1, s.t. \sum_{\substack{1 \le m \le M, 1 \le n \le N, 1 \le k \le K, \\ (m,n) \ne (m_d, n_d), d=1, \cdots, D}} \left| N^{\text{res}}_{m, n+(k-1)*K} \right|^2 < \sigma_1^2, \tag{17}$$

where the l_1-term enforces sparsity of the representation. We will specify the sufficient condition of correct recovery of the signal matrix $\hat{X}$ in Section 4. Then, the pseudo spectrum of α is obtained by calculating the l_2 norm of each row in the signal matrix $\hat{X}$ similarly.

Just like [7], with the knowledge of the distribution of noise, we can find a confidence interval for $\|N\|_F^2$, then use its upper value for σ_1^2.

Step 3: Estimate β based on the signal matrix

Let P denote the number of different α of the sources. Then, the pseudo spectrum of α may have P local maxima which are denoted as $\{ \hat{\alpha}_1, \cdots, \hat{\alpha}_P \}$. The row number in the signal matrix $\hat{X}$ that corresponds to these local maxima is denoted as $\{ i_1, \cdots, i_P \}$. The rows of $\hat{X}$ corresponding to possible α contain information about β. Let the ith row of $\hat{X}$ be

$\left(\hat{x}_{1,1}(i), \cdots, \hat{x}_{N,1}(i), \cdots, \hat{x}_{1,K}(i), \cdots, \hat{x}_{N,K}(i) \right) \in \mathbb{C}^{1 \times NK}$,

and $\hat{X}_i \triangleq \begin{bmatrix} \hat{x}_{1,1}(i) & \cdots & \hat{x}_{1,K}(i) \\ \vdots & \ddots & \vdots \\ \hat{x}_{N,1}(i) & \cdots & \hat{x}_{N,K}(i) \end{bmatrix} \in \mathbb{C}^{N \times K}$. According to Equation 15, we have:

$$\hat{X}_{i_p} = B_p^L S_p, \tag{18}$$

where $B_p^L = \left[b^L(\beta_{p,1}), \cdots, b^L(\beta_{p,f_p}) \right] \in \mathbb{C}^{N \times f_p}$ and $S_p = \left[s_p^{sg}(1), \cdots, s_p^{sg}(K) \right] \in \mathbb{C}^{f_p \times K}$. Here, $s_p^{sg}(k) = \left[s_{p,1}^{sg}(k), \cdots, s_{p,f_p}^{sg}(k) \right]^{\mathrm{T}}$. Similarly, an overcomplete representation of B_p^L is introduced. Let $\tilde{B}^L = \left[b^L(\tilde{\beta}_1), \cdots, b^L(\tilde{\beta}_{K_\beta}) \right] \in \mathbb{C}^{N \times K_\beta}$ where $\{ \tilde{\beta}_1, \cdots, \tilde{\beta}_{K_\beta} \}$ are all potential β. And $K_\beta \gg f_p$ $(1 \le p \le P)$. Taking into consideration of the noises introduced in the above MMV problems, we can solve another P MMV problems to estimate β:

$$\min_{\tilde{S}_p} \left\| \tilde{s}_p^{l2} \right\|_1, s.t. \left\| \hat{X}_{i_p} - \tilde{B}^L \tilde{S}_p \right\|_F^2 < \sigma_2^2, p = 1, \cdots, P, \tag{19}$$

where $\tilde{S}_p = \begin{bmatrix} \tilde{s}_{p,1}^{\mathrm{T}} \\ \vdots \\ \tilde{s}_{p,K_\beta}^{\mathrm{T}} \end{bmatrix} \in \mathbb{C}^{K_\beta \times K}$ and $\tilde{s}_p^{l2}$ denotes the vector whose k_βth element is the l_2-norm of the k_βth row in $\tilde{S}_p$. Then, the pseudo spectrum of β is $\| \tilde{s}_p^{l2} \|$ and the estimation of β is accomplished.

Although the distribution of the elements in $\hat{X}_i$ is difficult to estimate, σ_2^2 can be determined by the upper value of $K \left\| \left(n_{m1}^{sg}(k), \cdots, n_{mN}^{sg}(k) \right) \right\|_2^2$ empirically.

Step 4: Calculate (θ, φ) based on (α, β) if needed. The effects of invalid elements on the degree of freedom will be studied in Section 4.

3.2 Enhanced 2D-l_1-SVD algorithm

We will show in Section 5 that 2D-l_1-SVD works properly when the sources are well separated in both α and β domains. As illustrated in [14], there exists a source of bias inherent in the nature of the sparsity enforcing functionals. For example, consider a 1D case:

$$X = AS + N, \tag{20}$$

where $A \triangleq \left[a(\tilde{\theta}_1), \cdots, a(\tilde{\theta}_{K_\theta}) \right]$ and $\left\{ \tilde{\theta}_1, \cdots, \tilde{\theta}_{K_\theta} \right\}$ are all potential θ. $a(\theta)$ is the steering vector of a linear array. It is assumed that there are only two sources, which are from θ_1 and θ_2, impinging on the linear array. Obviously, the sparsity condition is satisfied with proper $\left\{ \tilde{\theta}_1, \cdots, \tilde{\theta}_{K_\theta} \right\}$ and the sources can be well resolved if they are not too close to each other. However, there is notable bias when

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 8 of 16

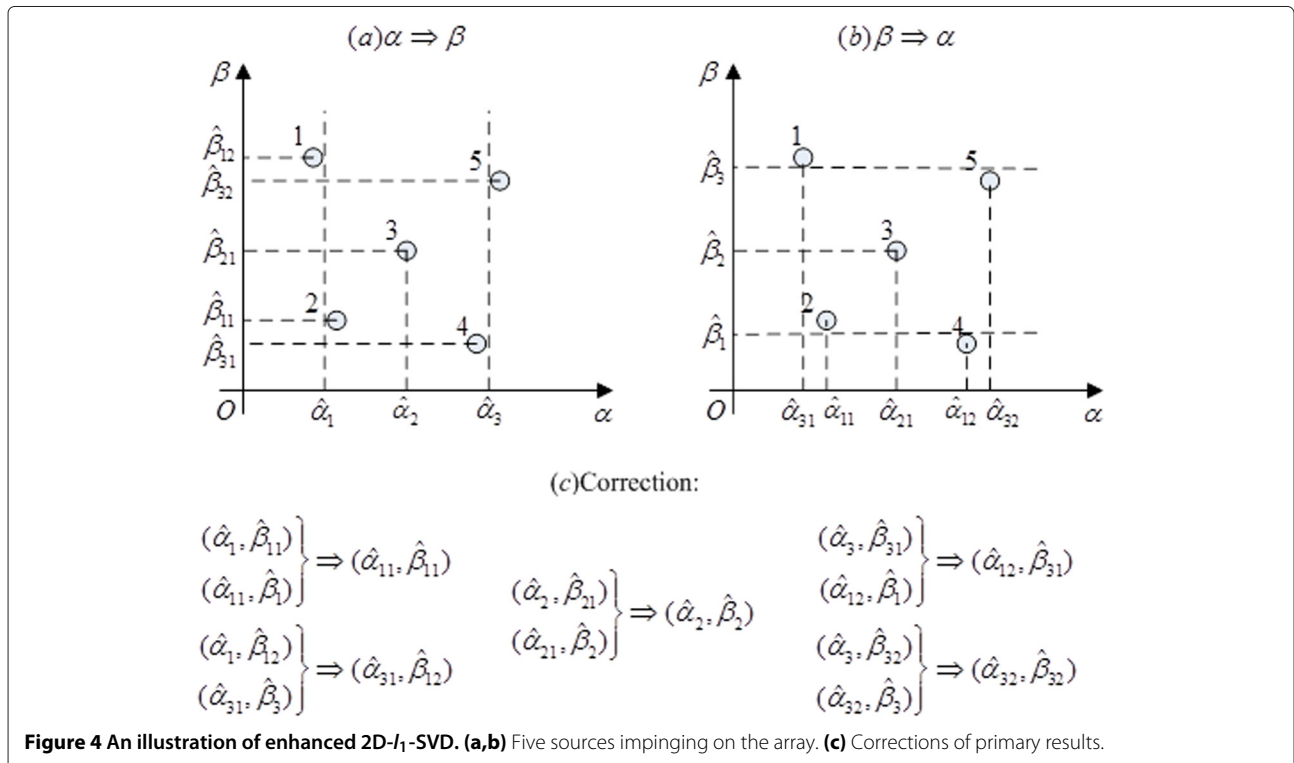two sources are too close to each other although the sparsity condition is still satisfied [14].

The problem still exists when l_1-SVD is extended to 2D cases. If two sources, which are from (α_1, β_1) and (α_2, β_2), are close to each other in both α and β domains so $\boldsymbol{a}(\alpha_1, \beta_1)$ is quite similar to $\boldsymbol{a}(\alpha_2, \beta_2)$, then l_1-SVD will get biased results. However, l_1-SVD can work properly and get nearly unbiased results when distinguished signals are close to each other in α domain but well separated to each other in β domain while 2D-l_1-SVD gets biased results in the α domain. An enhanced 2D-l_1-SVD is proposed to solve this problem, and it still has a much lower complexity than l_1-SVD.

We demonstrate the problem of the primary 2D-l_1-SVD algorithm with an example and then illustrate the main idea of the enhanced-2D-l_1-SVD. As illustrated in Figure 4a,b, there are five sources impinging on the array. The true DOAs of these five sources are (α_p, β_p), $p = 1, \cdots, 5$. And $\alpha_1 \approx \alpha_2$, $\alpha_4 \approx \alpha_5$, $\beta_1 \approx \beta_5$, and $\beta_2 \approx \beta_4$. Using 2D-l_1-SVD, we estimate α firstly and then estimate β (denoted as 2D-l_1-SVD-α), as shown in Figure 4a. Since α_1 is very close to α_2, they cannot be identified from each other in α domain. α_4 and α_5 cannot be identified either. As a result, only three estimates, $\hat{\alpha}_1 \approx (\alpha_1 + \alpha_2)/2$, $\hat{\alpha}_2 \approx \alpha_3$, and $\hat{\alpha}_3 \approx (\alpha_4 + \alpha_5)/2$, are obtained by solving a MMV problem due to the signals gathering in the α domain. Then, another three MMV problems are solved to get the estimates in the β domain. For example, we can extract the row information that corresponds to $\hat{\alpha}_1$ from the signal matrix $\hat{\boldsymbol{X}}$ and get two estimates, $\hat{\beta}_{1,1}$ and $\hat{\beta}_{1,2}$. So the estimates of DOAs of sources 1 and 2 are $(\hat{\alpha}_1, \hat{\beta}_{1,2})$ and $(\hat{\alpha}_1, \hat{\beta}_{1,1})$, as illustrated in Figure 4a. Similarly, the estimates of the DOAs of sources 3 to 5 are $(\hat{\alpha}_2, \hat{\beta}_{2,1})$, $(\hat{\alpha}_3, \hat{\beta}_{3,1})$, and $(\hat{\alpha}_3, \hat{\beta}_{3,2})$. Apparently, the estimation of the DOAs of sources 1, 2, 4, and 5 in α domain is not accurate. If we estimate β firstly and then estimate α(denoted as 2D-l_1-SVD-β), there is a similar problem since $\beta_1 \approx \beta_5$ and $\beta_2 \approx \beta_4$, as shown in Figure 4b. And the estimates of the DOAs of sources 1 to 5 using 2D-l_1-SVD-β are denoted as $(\hat{\alpha}_{3,1}, \hat{\beta}_3)$, $(\hat{\alpha}_{1,1}, \hat{\beta}_1)$, $(\hat{\alpha}_{2,1}, \hat{\beta}_2)$, $(\hat{\alpha}_{1,2}, \hat{\beta}_1)$, and $(\hat{\alpha}_{3,2}, \hat{\beta}_3)$, respectively. However, these sources can be distinguished from each other using l_1-SVD since the DOAs of the sources are well separated either in α domain or β domain.

Although 2D-l_1-SVD-α cannot accurately estimate α when there exist multiple sources close to each other in α domain, it can provide precise β-DOA if the sources are not too close to each other in β domain since the information of β is in the column of the input data matrix. Similar results can be obtained if β is firstly estimated using 2D-l_1-SVD-β. So the enhanced algorithm is designed to combine the estimation results in conjunction with a selection strategy.

The selection strategy is based on the condition that any two sources are not too close to each other in both α and β domains. The main idea is to make pairs of the estimation

Figure 4 An illustration of enhanced 2D-l_1-SVD. (a,b) Five sources impinging on the array. **(c)** Corrections of primary results.

Table 2 Arithmetic complexity of 2D-l_1-SVD and l_1-SVD

Methods		Arithmetic complexity	
l_1-SVD			$O\left(K^3 K_\alpha^3 K_\beta^3\right)$
2D-l_1-SVD	α	$O\left(K^3 N^3 K_\alpha^3\right)$	$O\left(K^3 N^3 K_\alpha^3\right) + PO\left(K^3 K_\beta^3\right)$
	β	$O\left(K^3 K_\beta^3\right)$	
Enhanced-2D-l_1-SVD		$O\left(K^3 N^3 K_\alpha^3\right) + P_\alpha O\left(K^3 K_\beta^3\right) + O\left(K^3 M^3 K_\beta^3\right) + P_\beta O\left(K^3 K_\alpha^3\right)$	

results of 2D-l_1-SVD-α and 2D-l_1-SVD-β so that in each pair, the DOA estimated by 2D-l_1-SVD-α and 2D-l_1-SVD-β corresponds to the same source. In Figure 4a, sources 1 and 2 are close to each other in α domain, so $\hat{\alpha}_1$ is inaccurate while $\hat{\beta}_{1,2}$ and $\hat{\beta}_{1,1}$ are precise. In Figure 4b, we get three different β, in which $\hat{\beta}_3$ is closest to $\hat{\beta}_{1,2}$ and $\hat{\beta}_1$ is closest to $\hat{\beta}_{1,1}$. Consider the estimation of α corresponding to $\hat{\beta}_3$ and we find that $\hat{\alpha}_{3,1}$ is the closest to $\hat{\alpha}_1$. Therefore, $(\hat{\alpha}_1, \hat{\beta}_{1,2})$ and $(\hat{\alpha}_{3,1}, \hat{\beta}_3)$ are considered to be a pair. Similarly, we have the other four pairs. For pairs corresponding to sources 1, 2, 4, and 5, 2D-l_1-SVD-α has got more precise estimation of β and 2D-l_1-SVD-β has got more precise estimation of α. Therefore, primary results of 2D-l_1-SVD-α and 2D-l_1-SVD-β are corrected, as illustrated in Figure 4c. It should be noticed that the pairing process in enhanced-2D-l_1-SVD is quite different from the pair-matching in [16]. Given the set of elevation angles and azimuth angles, which is $\{\theta_1, \cdots, \theta_P\}$ and $\{\varphi_1, \cdots, \varphi_P\}$, respectively, the conventional pair-matching process in [16] choose P final estimates from P^2 possible pairs by making use of the statistical analysis of the received data. In enhanced-2D-l_1-SVD, final DOA estimates are extracted from two existed set of 2D DOA without using information of the received data.

Another issue that needs to be taken into account is that spurious peaks may appear in the pseudo spectrum of l_1-SVD due to inappropriate regularization parameter but the selection strategy in enhanced-2D-l_1-SVD provides a detection of spurious peaks. For a certain $(\hat{\alpha}_i, \hat{\beta}_i)$ got by 2D-l_1-SVD-α, a corresponding $(\hat{\alpha}_j, \hat{\beta}_j)$ got by 2D-l_1-SVD-β can be found to make a pair with $(\hat{\alpha}_i, \hat{\beta}_i)$. Then, $(\hat{\alpha}_i, \hat{\beta}_i)$ is considered to be a false peak if:

$$\sqrt{\left|\hat{\alpha}_i - \hat{\alpha}_j\right|^2 + \left|\hat{\beta}_i - \hat{\beta}_j\right|^2} > \delta, \tag{21}$$

where δ is a threshold and the beamwidth of the array pattern is a reasonable choice for δ. The overall procedure of the enhanced 2D-l_1-SVD is summarized as follows:

Step 1: Complete the 2D-l_1-SVD-α.
Estimate α firstly and then estimate β. There are P different α and the estimation results are denoted as $\Omega_\alpha = \{(\hat{\alpha}_1, \hat{\beta}_{1,1}), \cdots, (\hat{\alpha}_1, \hat{\beta}_{1,f_1}), \cdots, (\hat{\alpha}_P, \hat{\beta}_{P,1}), \cdots, (\hat{\alpha}_P, \hat{\beta}_{P,f_P})\}$. For $\hat{\alpha}_p$, there are f_p different β.

Step 2: Complete the 2D-l_1-SVD-β.
Estimate β firstly and then estimate α. There are Q different β and the estimation results are denoted as $\Omega_\beta = \{(\hat{\alpha}_{1,1}, \hat{\beta}_1), \cdots, (\hat{\alpha}_{1,g_1}, \hat{\beta}_1), \cdots, (\hat{\alpha}_{Q,1}, \hat{\beta}_Q), \cdots, (\hat{\alpha}_{Q,g_Q}, \hat{\beta}_Q)\}$. For $\hat{\beta}_q$, there are g_q different α.

Step 3: Selection strategy.
For each $(\hat{\alpha}_p, \hat{\beta}_{pq}) \in \Omega_\alpha$, $1 \leq q \leq g_p$, find $\hat{\beta}_q$ in $\{\hat{\beta}_1, \cdots, \hat{\beta}_Q\}$ that is closest to $\hat{\beta}_{p,q}$. The estimation of α corresponding to $\hat{\beta}_q$ is $\{\hat{\alpha}_{q,1}, \cdots, \hat{\alpha}_{q,f_q}\}$. Find $\hat{\alpha}_{qk}$ in $\{\hat{\alpha}_{q,1}, \cdots, \hat{\alpha}_{q,f_q}\}$ that is closest to $\hat{\alpha}_p$. Then, $(\hat{\alpha}_p, \hat{\beta}_{p,q})$ and $(\hat{\alpha}_{q,k}, \hat{\beta}_q)$ are considered to be a pair. If $\sqrt{\left|\hat{\alpha}_p - \hat{\alpha}_{q,k}\right|^2 + \left|\hat{\beta}_{p,q} - \hat{\beta}_q\right|^2} > \delta$, then, $(\hat{\alpha}_{q,k}, \hat{\beta}_{p,q})$ was considered to be a false peak. If $\sqrt{\left|\hat{\alpha}_p - \hat{\alpha}_{q,k}\right|^2 + \left|\hat{\beta}_{p,q} - \hat{\beta}_q\right|^2} \leq \delta$, then $(\hat{\alpha}_{q,k}, \hat{\beta}_{p,q})$ is the correct DOA.

4 Performance

Using l_1-SVD as 1D DOA estimators during successive estimation, 2D-l_1-SVD has several advantages over conventional 2D methods, including high resolution, robustness to the number of snapshots, low SNR, and coherent sources due to the use of sparse reconstruction [7]. In this section, we compare the performance of 2D-l_1-SVD and l_1-SVD and study about the effect of invalid elements. Firstly, we demonstrate that the computational complexity of 2D-l_1-SVD is much lower than that of l_1-SVD. Secondly, we study about the degree of freedom of 2D-l_1-SVD. Then, we will show that 2D-l_1-SVD keeps better robustness over the assumed number of sources in the

Table 3 Degrees of freedom of 2D-l_1-SVD and l_1-SVD

Methods	Degrees of freedom			
	Single snapshot		Multiple snapshots	
l_1-SVD	$MN/2$		$MN - 1$	
2D-l_1-SVD-α	α	$M/2$	α	$M - 1$
	β for each α	$N/2$	β for each α	$N - 1$
	Maximum	$MN/4$	Maximum	$(M-1)(N-1)$
2D-l_1-SVD-β	β	$N/2$	β	$N - 1$
	α for each β	$M/2$	α for each β	$M - 1$
	Maximum	$MN/4$	Maximum	$(M-1)(N-1)$

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 10 of 16

Table 4 Time cost of different methods

Time (s)	$K_{\alpha\beta}$			
	46	91	181	361
2D-MUSIC	0.3315	1.2961	5.1043	20.2029
l_1-SVD	2.7922	12.9736	59.9812	397.105
2D-l_1-SVD	0.4381	0.4929	0.5964	0.8356
Enhanced-2D-l_1-SVD	1.0704	1.4574	1.6928	2.5747

presence of multiple sources. And finally, we investigate the number of sources that 2D-l_1-SVD can process with both rectangular arrays and faulty rectangular arrays.

4.1 Computational complexity
Sparse reconstruction methods based on l_1-norm minimization can be achieved by second-order cone

programming (SOCP). For optimizing the joint optimization problem over K vectors in SOCP framework using an interior point method, the arithmetic complexity is $O\left(K^3 K_\theta^3\right)$, where K_θ is the number of potential directions [7,19]. Therefore, we have the computational load of 2D-l_1-SVD and l_1-SVD in Table 2. Since K_α and K_β are always much bigger than M, N, K, the arithmetic cost of 2D-l_1-SVD is much smaller than l_1-SVD.

4.2 Degrees of freedom
A necessary and sufficient condition [20] from the measurements $X = AS$, $|\text{supp}(S)| = k$, to uniquely determine S is ($|\text{supp}(S)|$ denotes the union over all the individual supports $\cup_i \text{supp}(s_i)$ for $S = [s_1, \cdots, s_l]$):

$$|\text{supp}(S)| < (\text{spark}(A) - 1 + \text{rank}(X))/2, \quad (22)$$
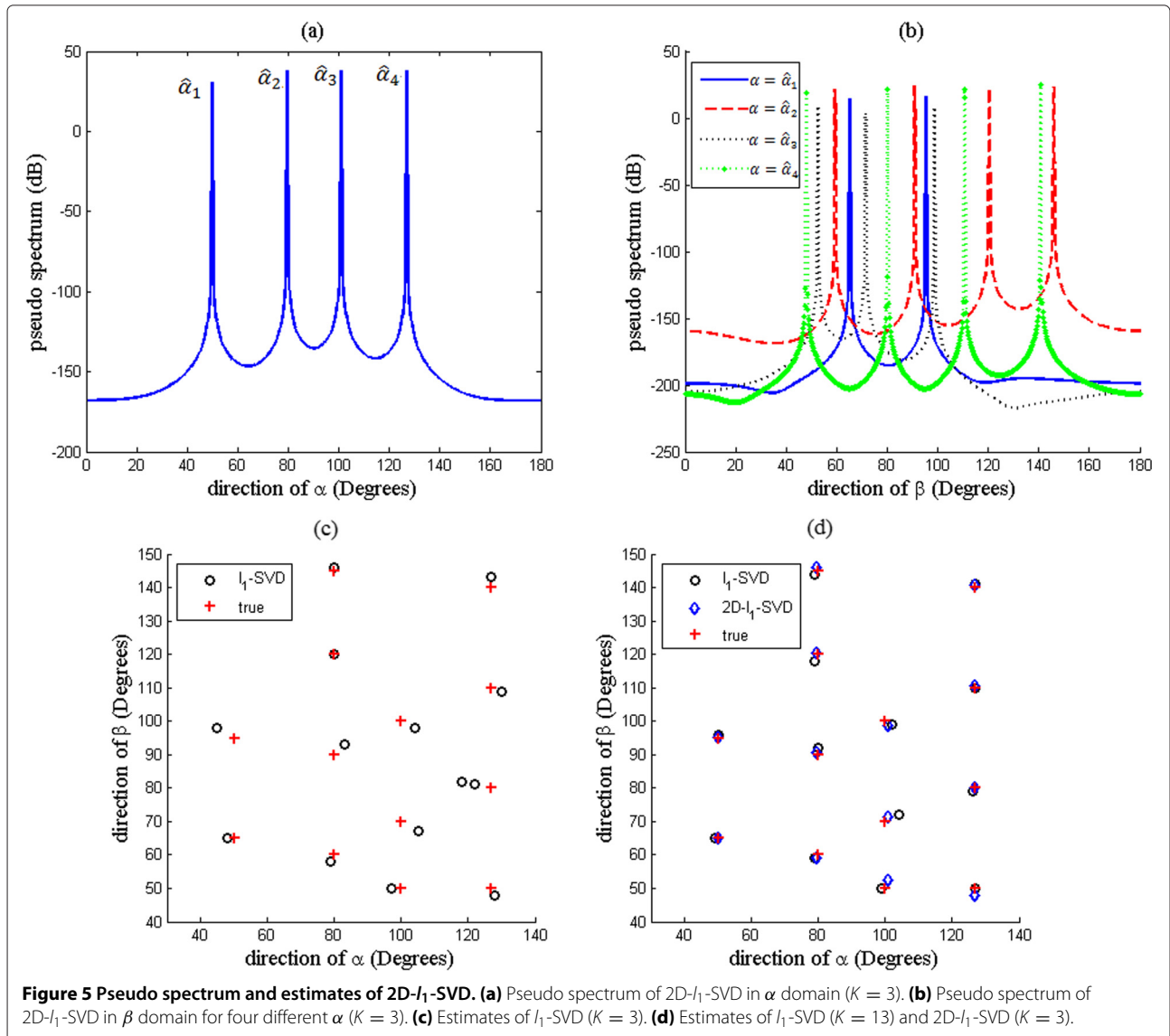
Figure 5 Pseudo spectrum and estimates of 2D-l_1-SVD. (a) Pseudo spectrum of 2D-l_1-SVD in α domain ($K=3$). **(b)** Pseudo spectrum of 2D-l_1-SVD in β domain for four different α ($K=3$). **(c)** Estimates of l_1-SVD ($K=3$). **(d)** Estimates of l_1-SVD ($K=13$) and 2D-l_1-SVD ($K=3$).

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 11 of 16

where the spark of A is defined as the smallest number of columns in A that are linearly dependent. For DOA estimation problems, the measurement matrix X and the signal matrix S stand for the sampled snapshots at the array and the incoming source signals, respectively. The dictionary A is a matrix composed of the steering vectors corresponding to each potential DOA as its columns. And the number of columns in A is much bigger than the number of rows in A. Let L_0 denote the number of sensors in the array. For linear or rectangular arrays whose manifold matrix is similar to a Vandermonde matrix, spark(A) is equal to $L_0 + 1$ if the potential DOAs in the dictionary are not too close to each other so that the mutual coherence [20] is small. On the hand, rank(X) = 1 if the number of snapshots is only one and rank(X) = K if the number of singular vectors used is sufficient, e.g., equal to the number of sources. So empirically, the l_1-SVD technique can resolve $M-1$ sources with an M-sensors array if they are not located too close to each other [7]. This holds under the assumption that the number of singular vectors used in l_1-SVD is sufficient, e.g., equal to the number of sources. When fewer singular vectors are taken than the number of sources, the number of resolvable sources may decrease.

For a $M \times N$ rectangular array, with no consideration of array ambiguity and only one snapshot available, l_1-SVD can process up to $MN/2$ sources. For 2D-l_1-SVD-α, there exists a coherence between the samples of sub-arrays at different columns of the array as far as the estimation of α is concerned, so 2D-l_1-SVD-α can process $M/2$ different α and up to $N/2$ different β for each α. When there are enough snapshots and the number of uncorrelated columns in the data matrix X is not less than the number of sensors, l_1-SVD can process up to $MN - 1$ sources, while 2D-l_1-SVD-α can process up to $M-1$ different α and $N-1$ different β for each certain α. Similarly, 2D-l_1-SVD-β can process up to $N/2$ different β and $M/2$ different α with only one snapshot. When there are enough snapshots, 2D-l_1-SVD-β can process up to $N - 1$ different β and $M - 1$ different α. So the number of sources that 2D-l_1-SVD or enhanced-2D-l_1-SVD can process is less than l_1-SVD when the sources are well separated in both α and β domains. However, when the sources are gathering in α or β domains and several sources may be treated as one group in successive parameter estimation, enhanced-2D-l_1-SVD can resolve up to $(M - 1)(N - 1)$ sources, which is close to l_1-SVD if M and N are sufficiently large in practical application. The degrees of freedom of different methods are illustrated in Table 3.

4.3 The assumed number of sources
The l_1-SVD technique works on the K singular vectors where K is the assumed number of sources. It has been illustrated in [7] that l_1-SVD has a low sensitivity to K.

In [7], it is illustrated that l_1-SVD can resolve $M - 1$ sources using an M-sensor array. However, when fewer singular vectors are taken than the number of sources, the condition in (22) may be not satisfied and the number of resolvable sources may decrease. This limitation still exists in 2D situation. Assume that the 2D DOA of incoming sources are $\{(\alpha_1, \beta_{1,1}), \cdots, (\alpha_1, \beta_{1,f_1}), \cdots, (\alpha_P, \beta_{P,1}), \cdots, (\alpha_P, \beta_{P,f_P})\}$, and $P_0 \triangleq \sum_{p=1}^{P} f_p$. To resolve P_0 sources that have P different α, the assumed number of sources for l_1-SVD should be not less than P_0 if P_0 is sufficiently large (for example, P_0 is close to the number of sensors). The number of singular vectors should not be less than P or f_p to resolve P different α and f_p different β, correspondingly. As a result, if the assumed number of sources for 2D-l_1-SVD is not less than $\max\{P, f_1, \cdots, f_P\}$, the 2D-$l_1$-SVD can resolve P_0 sources. So 2D-l_1-SVD may resolve more sources with fewer singular vectors. Moreover, 2D-l_1-SVD may have an even smaller computational complexity by taking a smaller number of singular vectors.

4.4 Faulty or nonrectangular arrays
When there are only a small number of elements with failure in the array, the faulty elements have little effect on the performance of 2D-l_1-SVD. The faulty array is no longer rectangular while it can be regarded as a rectangular array with missing elements. Under this circumstance, the degree of freedom may be affected while 2D-l_1-SVD is still able to produce correct DOA estimates. Empirically, 2D-l_1-SVD-α can resolve $M_0 - 1$ different α and $N_0 - 1$ different β for each α when the sources are not too close to
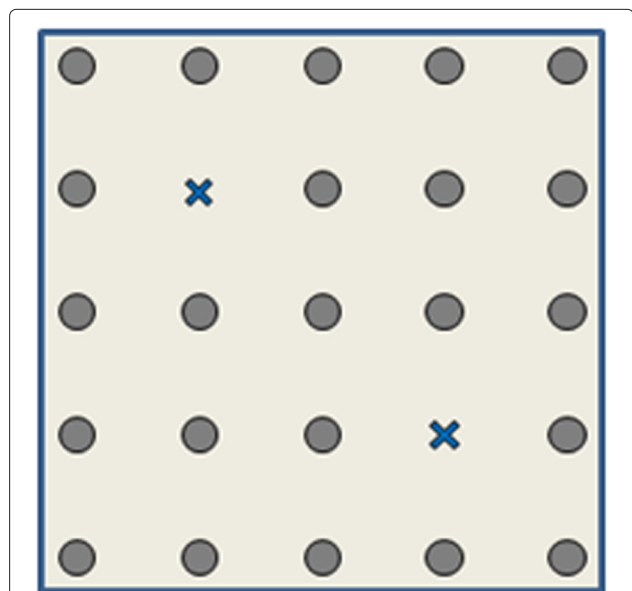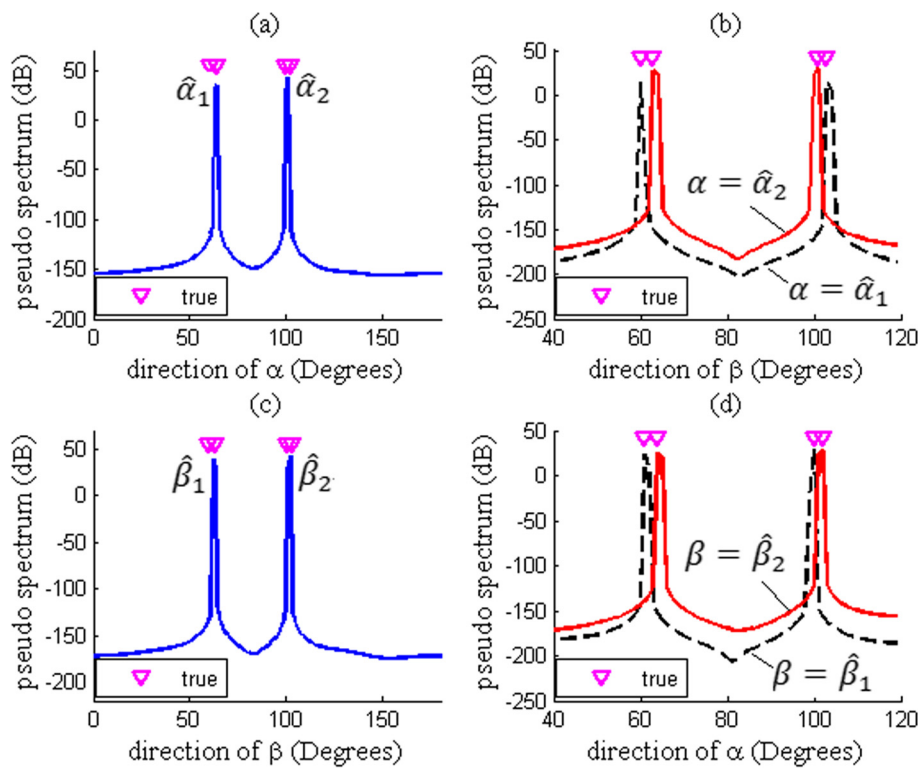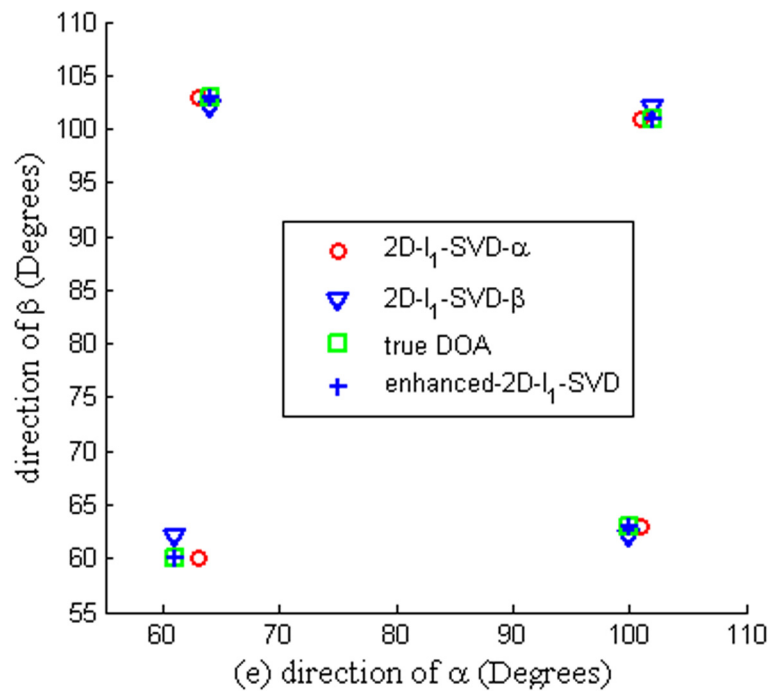

Figure 6 Faulty array.

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 12 of 16

Figure 7 Pseudo spectrum. (a) α in 2D-l_1-SVD-α. **(b)** β for two different α in 2D-l1-SVD-α. **(c)** β in 2D-l1-SVD-β. **(d)** α for two different β in 2D-l1-SVD-β.

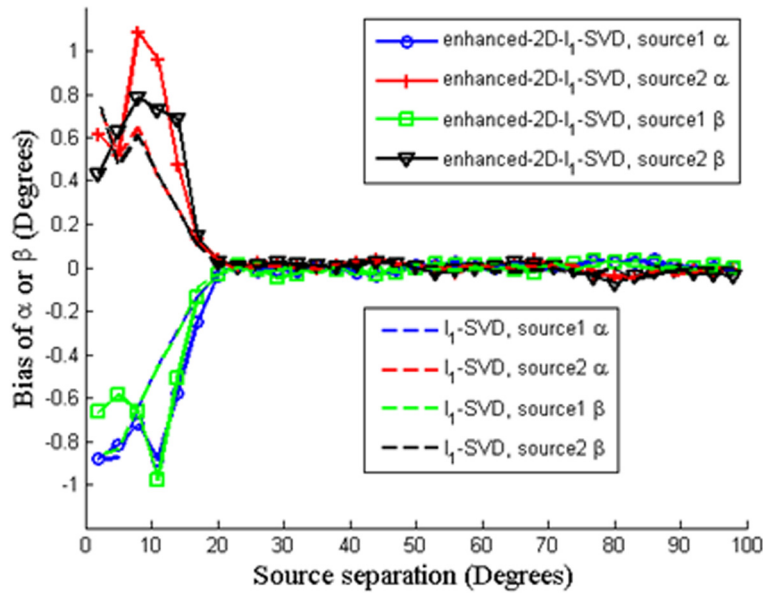

Figure 8 Selection strategy of enhanced-2D-l_1-SVD.

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 13 of 16

Figure 9 Bias in α and β domain of 2D-l_1-SVD and l_1-SVD in localizing two sources as a function of the source separation.

each other. Here, M_0 and N_0 denote the minimum number of valid elements in each row or column of the rectangular array. On the other hand, there should not be too many invalid elements, or the distances between valid elements become so large that there will be grating lobes.

5 Simulation results

In this section, several simulations are conducted to validate the advantages of 2D-l_1-SVD. All simulations are performed using MATLAB 2012a running on an Intel Core i7 3770 CPU @ 3.4 GHz with 16 GB RAM, under Windows 7.

5.1 Time cost

Firstly, we compare the computation time of 2D-MUSIC, l_1-SVD, 2D-l_1-SVD, and enhanced 2D-l_1-SVD. Consider a 5×5 URA with sensors spaced half a wavelength apart and there are three uncorrelated sources with 100
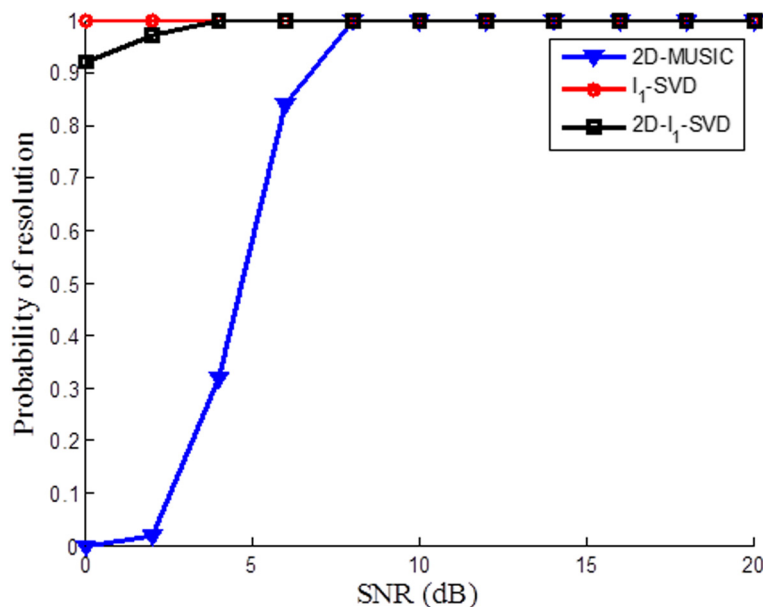

Figure 10 The probability of resolution as a function of SNR.

snapshots. The DOAs of the three sources are (120°, 70°), (80°, 50°), and (50°, 130°). And SNR = 10 dB. The time that the four methods cost against $K_{\alpha\beta}$, which represents the density of the grids or the number of potential source locations (the grids or potential locations are set as $\left[0, \pi / \left(K_{\alpha\beta} - 1\right), \cdots, \left(K_{\alpha\beta} - 2\right) \pi / \left(K_{\alpha\beta} - 1\right), \pi\right]$ for both α and β domains), is demonstrated in Table 4. Each value in the table is an average over 50 trials. Table 4 shows that 2D-l_1-SVD and enhanced 2D-l_1-SVD have a much lower computational load than other methods. The cost time of enhanced 2D-l_1-SVD is nearly twice of that of 2D-l_1-SVD, and it is affected by the number of different α and β.
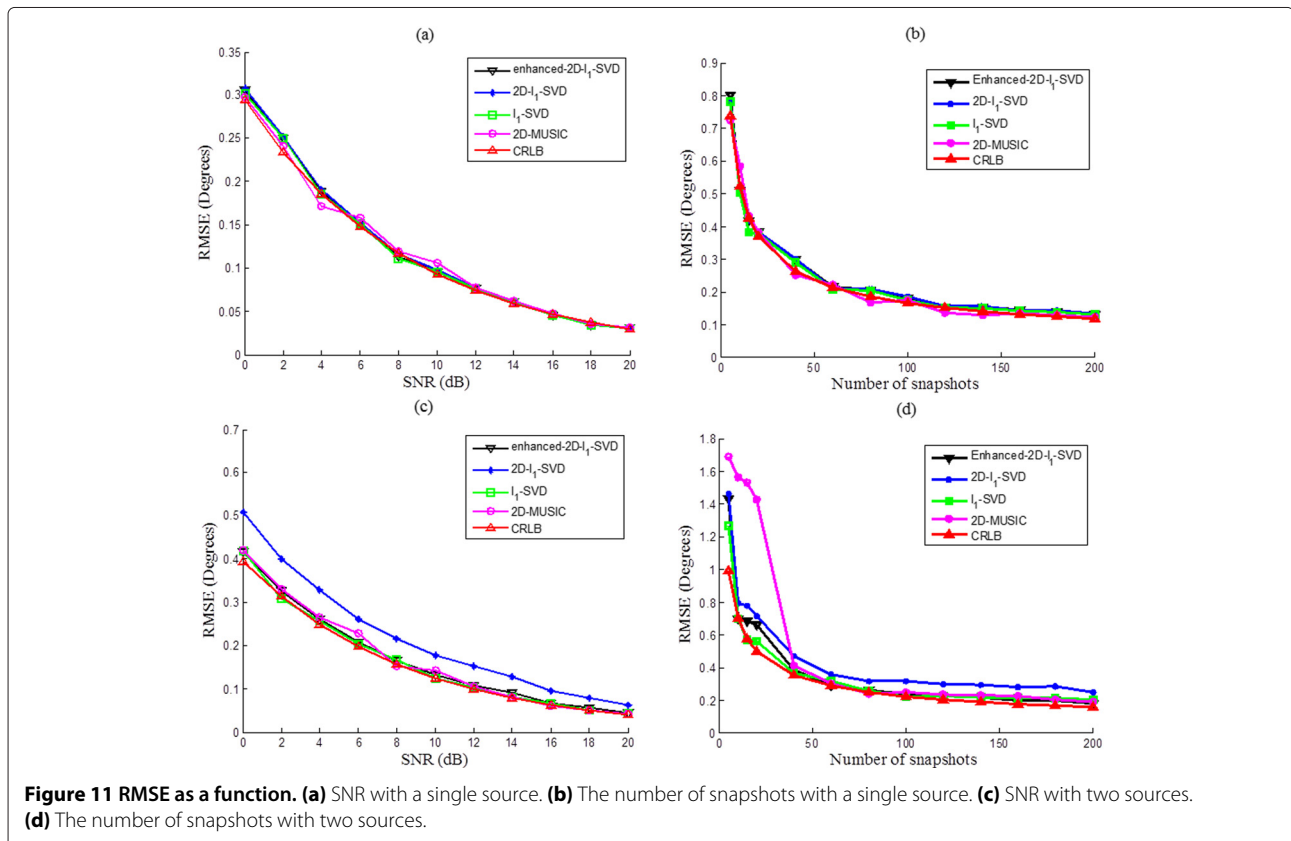
5.2 Pseudo spectrum with multiple sources

The pseudo spectrum of 2D-l_1-SVD and l_1-SVD when there are 13 independent sources (four different α) impinging on the 5×5 URA with sensors spaced half a wavelength apart is illustrated in Figure 5. The true DOAs are (50°, 65°), (50°, 95°), (80°, 60°), (80°, 90°), (80°, 120°), (80°, 145°), (100°, 50°), (100°, 70°), (100°, 100°), (127°, 50°), (127°, 80°), (127°, 110°), and (127°, 140°). SNR = 10 dB, and there are 100 snapshots. 2D-l_1-SVD is able to identify four different α and give correct DOA estimation when the assumed number of sources is $K = 3$ (Figure 5a,b,d). l_1-SVD has given correct estimation if $K = 13$ (Figure 5d)

while there are spurious peaks in the spectrum of l_1-SVD if $K = 3$ (Figure 5c). So 2D-l_1-SVD may perform better with fewer singular vectors.

5.3 Bias

Many DOA estimation methods may have difficulty resolving closely spaced sources, and there is bias inherent in the nature of sparsity enforcing functionals [14]. Consider the faulty array in Figure 6 with sensors spaced half a wavelength apart. There are four uncorrelated sources with a SNR = 10 dB. And there are 100 snapshots. The pseudo spectrum of 2D-l_1-SVD-α and 2D-l_1-SVD-β when the true DOAs are (61°, 60°), (64°, 103°), (100°, 63°), and (102°, 101°) is illustrated in Figure 7. We can see that the peaks of the spectrum in Figure 7a,c are biased. A final estimation of DOA is given in Figure 8 by combining the results got by 2D-l_1-SVD-α and 2D-l_1-SVD-β, and the results are correct.

The bias of the DOA estimation of two sources with different angular separation between them is investigated here. Consider the faulty array in Figure 6. The source 1 is held fixed at (62.6°, 58.7°) and the source 2 is moving from the location of source 1 linearly. And SNR = 10 dB. Bias in the α or β domain of 2D-l_1-SVD and l_1-SVD as a function of the source separation δ_0 (degrees) is demonstrated in Figure 9 while source 2 is located at (62.6° + δ_0, 58.7° +

Figure 11 RMSE as a function. (a) SNR with a single source. **(b)** The number of snapshots with a single source. **(c)** SNR with two sources. **(d)** The number of snapshots with two sources.

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 15 of 16

δ_0). The values on each curve are an average over 50 trials. Figure 9 shows the presence of bias for low separations using both enhanced-2D-l_1-SVD and l_1-SVD, and the bias disappears when $\delta_0 > 20°$ for both algorithms.

5.4 Resolution capability

The resolution performance of the proposed algorithms and the 2D-MUSIC algorithm is compared here. Consider the faulty array in Figure 6. There are two incoming sources from (80.3°, 106.4°) and (85.3°, 111.4°). There are 100 snapshots. The probability of resolution [21] as a function of SNR is illustrated in the Figure 10, and it is based on 100 trials. We can see that both l_1-SVD and 2D-l_1-SVD outperform 2D-MUSIC when the SNR is less than 8 dB. And 2D-l_1-SVD performs close to l_1-SVD. Note that the resolution probability of enhanced-2D-l_1-SVD is the same to that of 2D-l_1-SVD since the results of the enhanced algorithm are based on the estimates of 2D-l_1-SVD.

5.5 RMSE

The root mean squared error (RMSE) is defined as

$$\text{RMSE} = \frac{1}{P} \sum_{p=1}^{P} \left(\text{RMSE}\left(\alpha_p\right) + \text{RMSE}\left(\beta_p\right) \right), \quad (23)$$

where $\text{RMSE}\left(\alpha_p\right) = \left[\sum_{l=1}^{L} \left(\hat{\alpha}_p\left(l\right) - \alpha_p\right)^2 / L \right]^{1/2}$, $\text{RMSE}\left(\beta_p\right) = \left[\sum_{l=1}^{L} \left(\hat{\beta}_p\left(l\right) - \beta_p\right)^2 / L \right]^{1/2}$ and $\left(\hat{\alpha}_p\left(l\right), \hat{\beta}_p\left(l\right)\right)$ is the lth estimates of the DOA of pth signal $\left(\alpha_p, \beta_p\right)$. L is the number of realizations.

Consider the faulty array in Figure 6. The RMSE of different methods against SNR based on 100 realizations compared with the Cramer-Rao lower bound (CRLB) [22] when there are a single source from (80.3°, 106.4°) with 100 snapshots is given in Figure 11a. And the RMSE against the number of snapshots when SNR of the single source is 5 dB is given in Figure 11b. It shows that 2D-MUSIC, 2D-l_1-SVD, and enhanced-2D-l_1-SVD perform close to the CRLB when there is only one source. The RMSE as a function of SNR based on 100 realizations when there are two uncorrelated sources from (80.3°, 106.4°) and (100.9°, 127.1°) is given in Figure 11c, and the RMSE against the number of snapshots when SNR of the two sources is 5 dB is given in Figure 11d. We can see that the performance of 2D-MUSIC deteriorates when there are not enough snapshots (less than twice of the number of antennas) while SSR based algorithms work properly.

6 Conclusions

In this paper, a new 2D estimation method called 2D-l_1-SVD and its improved version enhanced-2D-l_1-SVD

are proposed for rectangular arrays. They are proved to be able to work for rectangular arrays even with missing or faulty elements. Theoretical analysis and simulation results indicate that 2D-l_1-SVD has a much lower arithmetic complexity due to successive parameter estimation while performing close to the popular l_1-SVD. What is more, 2D-l_1-SVD has better robustness to the assumed number of sources and enhanced-2D-l_1-SVD is able to detect spurious peaks which are caused by inappropriate regularization parameter.

References
1. H Krim, M Viberg, Two decades of array signal processing research: the parametric approach. IEEE Signal Process. Mag. **13**(4), 67–94 (1996)
2. E Tuncer, B Friedlander, *Classical and Modern Direction-of-Arrival Estimation* (Elsevier Academic Press, Burlington, USA, 2009)
3. J Capon, High-resolution frequency-wavenumber spectrum analysis. Proc IEEE. **57**(8), 1408–1418 (1969)
4. R Schmidt, Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. **34**(3), 276–280 (1986)
5. R Roy, T Kailath, ESPRIT-estimation of signal parameters via rotational invariance techniques. IEEE Trans. Acoustics Speech Signal. Process. **37**(7), 984–995 (1989)
6. P Stoica, K Sharman, Maximum likelihood methods for direction-of-arrival estimation. IEEE Trans. Acoustics Speech Signal. Process. **38**(7), 1132–1143 (1990)
7. D Malioutov, M Cetin, A Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans. Signal Process. **53**(8), 3010–3022 (2005)
8. M Hyder, K Mahata, Direction-of-arrival estimation using a mixed $l_{2,0}$ norm approximation. IEEE Trans. Signal Process. **58**(9), 4646–4655 (2010)
9. P Stoica, P Babu, J Li, A sparse covariance-based estimation method for array processing. IEEE Trans. Signal Process. **59**(2), 629–638 (2011)
10. J Zheng, M Kaveh, Sparse spatial spectral estimation: a covariance fitting algorithm, performance and regularization. IEEE Trans. Signal Process. **61**(11), 2767–2777 (2013)
11. X Xu, X Wei, Z Ye, DOA estimation based on sparse signal recovery utilizing weighted l_1-norm penalty. IEEE Signal Process. Lett. **19**(3), 155–158 (2012)
12. A Gershman, M Rubsamen, M Pesavento, One-and two-dimensional direction-of-arrival estimation: an overview of search-free techniques. Signal Process. **90**, 1338–1349 (2010)
13. Y Wang, L Lee, A tree structure one-dimensional based algorithm for estimating the two-dimensional direction of arrivals and its performance analysis. IEEE Trans. Antennas Propag. **56**(1), 178–188 (2008)
14. D Malioutov, A sparse signal reconstruction perspective for source localization with sensor arrays. PhD thesis, Massachusetts Institute of Technology (2003)
15. Y Liu, M Wu, S Wu, Fast OMP algorithm for 2D angle estimation in MIMO radar. Electronics Lett. **46**(6), 444–445 (2010)
16. S Kikuchi, H Tsuji, A Sano, Pair-matching method for estimating 2-D angle of arrival with a cross-correlation matrix. IEEE Antennas Wireless Propag. Lett. **5**(1), 35–40 (2006)
17. Q Liu, S OuYang, L Jin, Two-dimensional DOA estimation with L-shaped array based on a jointly sparse representation. Inf. Technol. J. **12**, 2037–2042 (2013)
18. F Belloni, A Richter, V Koivunen, DOA estimation via manifold separation for arbitrary array structures. IEEE Trans. Signal Process. **55**(10), 4800–4810 (2007)

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2015) 2015:7

Page 16 of 16

19. M Lobo, L Vandenberghe, S Boyd, H Lebret, Applications of second-order cone programming. Linear Algebra Its Applicat. Special Issue Linear Algebra Control Signals Image Process. **284**, 193–228 (1998)
20. M Davies, Y Eldar, Rank awareness in joint sparse recovery. IEEE Trans. Inf. Theory. **58**(2), 1135–1146 (2012)
21. Q Zhang, Probability of resolution of the MUSIC algorithm. IEEE Trans. Signal Process. **43**(4), 978–987 (1995)
22. Y Hua, T Sarkar, A note on the Cramer-Rao bound for 2-D direction finding based on 2-D array. IEEE Trans. Signal Process. **39**(5), 1215–1218 (1991)