

RESEARCH

Open Access

Subspace weighted $\ell_{2,1}$ minimization for sparse signal recovery

Chundi Zheng, Gang Li*, Yimin Liu and Xiqin Wang

Abstract

In this article, we propose a weighted $\ell_{2,1}$ minimization algorithm for jointly-sparse signal recovery problem. The proposed algorithm exploits the relationship between the noise subspace and the overcomplete basis matrix for designing weights, i.e., large weights are appointed to the entries, whose indices are more likely to be outside of the row support of the jointly sparse signals, so that their indices are expelled from the row support in the solution, and small weights are appointed to the entries, whose indices correspond to the row support of the jointly sparse signals, so that the solution prefers to reserve their indices. Compared with the regular $\ell_{2,1}$ minimization, the proposed algorithm can not only further enhance the sparseness of the solution but also reduce the requirements on both the number of snapshots and the signal-to-noise ratio (SNR) for stable recovery. Both simulations and experiments on real data demonstrate that the proposed algorithm outperforms the ℓ_1 -SVD algorithm, which exploits straightforwardly $\ell_{2,1}$ minimization, for both deterministic basis matrix and random basis matrix.

Keywords: sparse signal recovery, weighted $\ell_{2,1}$ minimization, multiple measurement vectors (MMV), direction-of-arrival estimation

1 Introduction

In recent years, sparse signal recovery has attracted a great deal of attention from the signal processing society [1-16]. Using an overcomplete basis $\mathbf{A} \in \mathbb{C}^{M \times K}$ ($M \ll K$) and the sparsity prior on signal \mathbf{x} , the sparse representation problem of the noiseless measurements with single measurement vector (SMV) $\mathbf{y} = \mathbf{A}\mathbf{x}$ can be solved by a combinatorial ℓ_0 problem

$$\min \|\mathbf{x}\|_0 \quad \text{s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}, \quad (1)$$

where the sparse signal \mathbf{x} has only P nonzero components and $\|\mathbf{x}\|_0 = P$ represents the number of nonzero components of \mathbf{x} . Unfortunately, the minimization problem (1) is NP-hard. A practicable way of solving the sparse representation problem is to employ the following convex optimization

$$\min \|\mathbf{x}\|_1 \quad \text{s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}, \quad (2)$$

where $\|\mathbf{x}\|_1 = \sum_{i=1}^K |x_i|$. As a surrogate of the ℓ_0 norm, the regular ℓ_1 norm is tractable, but it depends on signal's coefficient values and attenuates the nature of literal ℓ_0 sparsity count, which may cause performance degrading in some situations. To avoid the disadvantage of the dependence on magnitude of the regular ℓ_1 minimization, Candès et al. designed an iterative reweighted formulation of ℓ_1 minimization to more democratically penalize nonzero coefficients, namely, large weights will discourage to reserve those entries who are more likely to be zero in recovered signal, whereas small weights will encourage to reserve larger entries [12]. In other words, the essence of the iterative reweighted ℓ_1 minimization lies in that large weights are appointed to those elements of \mathbf{x} , whose indices are more likely to be outside of the support [14,16], which expels their indices from the support in the sparse solution and further consolidates the sparsity-encouraging nature of regular ℓ_1 minimization [12-16]. The support of \mathbf{x} is defined as $\text{Supp}(\mathbf{x}) = \{k | x_k \neq 0\}$. Incidentally, it was proved that the iterative reweighted ℓ_1 minimization can indeed improve both the recoverable sparsity thresholds and the recovery accuracy upon the regular ℓ_1 minimization [13-15].

* Correspondence: gangli@tsinghua.edu.cn
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

It is worth noting that the iterative reweighted ℓ_1 minimization was designed for the SMV problem [12]. In fact, the multiple measurement vectors (MMV) problem is encountered in many applications of sparse signal representation such as array processing [1,6-11], magne-toencephalography [1], nonparametric spectrum analysis of time series [17], equalization of sparse communication channels [18] and so on. In the MMV case the ℓ_1 -SVD method [7-9] replaces the ℓ_1 norm minimization with the mixed norm $\ell_{2,1}$ norm minimization. Similar to the regular ℓ_1 norm minimization, the $\ell_{2,1}$ minimization also meets the disadvantage of the dependence on magnitude. Therefore, in the MMV case how to design an appropriate weighting vector to cope with the disadvantage of the dependence on magnitude of the regular $\ell_{2,1}$ minimization is an interesting issue. In this article, we focus on the noisy MMV case and propose an algorithm of the jointly-sparse signal recovery based on the relationship between the noise subspace and the overcomplete basis for weighting the jointly sparse signals, which extends the essence of the iterative reweighted ℓ_1 minimization in [12] from SMV to MMV. The measurements with MMV can be written as

$$\mathbf{y}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{n}(t), \quad t = 1, \dots, T, \quad (3)$$

where the vector $\mathbf{n}(t)$ denotes an additive noise vector with zero-mean and variance σ^2 , the vector $\mathbf{x}(t)$ is the jointly-sparse signals and the support is independent of the snapshot t [1]. Without loss of generality, the additive noise $\mathbf{n}(t)$ is assumed to be uncorrelated with the jointly-sparse signals $\mathbf{x}(t)$. The row support of the jointly sparse signals \mathbf{X} (\mathbf{X} denotes the matrix form of $\mathbf{x}(t)$) plays a key role in the sparse signal recovery with MMV, and it can be defined as $\text{Supp}_{\text{row}}(\mathbf{X}) = \{k | \mathbf{X}_k^{(\ell_2)} \neq \mathbf{0}\} \triangleq \Lambda$ [6], where $\mathbf{X}_k^{(\ell_2)}$ denotes the k th entry of $\mathbf{X}^{(\ell_2)}$, and $\mathbf{X}^{(\ell_2)}$ is a column vector whose k th elements denotes the ℓ_2 norm of k th row of \mathbf{X} . It is obvious that the index set $\Lambda \subseteq \{1, \dots, K\}$ and its cardinality $|\Lambda| = P$. Considering the relationship between the indices of the columns of \mathbf{A} and the row support Λ , the overcomplete basis \mathbf{A} can be divided into two submatrix, i.e., $\mathbf{A} = [\mathbf{A}_\Lambda \mathbf{A}_{\Lambda^c}]$, where the indices of the columns of the submatrix \mathbf{A}_Λ constitute the row support Λ , and the indices of the columns of the submatrix \mathbf{A}_{Λ^c} constitute the complement of Λ , i.e., $\Lambda \cup \Lambda^c = \{1, \dots, K\}$ and $\Lambda \cap \Lambda^c = \emptyset$. On the other hand, the subspace decomposition on $\{\mathbf{y}(t), t = 1, \dots, T\}$ provides the signal subspace and the noise subspace. It is noted that the noise subspace is orthogonal to the column space of \mathbf{A}_Λ [19-21] but not to the submatrix \mathbf{A}_{Λ^c} .

From this observation, this article designs a subspace weighted (SW) $\ell_{2,1}$ minimization algorithm in the MMV case, in which small and large weights are generated by using the orthogonality between noise subspace and \mathbf{A}_Λ

and the isomorphism between noise subspace and \mathbf{A}_{Λ^c} . We will show that the designed weights can force the entries whose indices are more likely to be outside of the row support to be close to zero in the solution, and, therefore further promotes the sparseness of the solution and improves the recovery accuracy.

Although our proposed algorithm also exploits the singular value decomposition (SVD), compared with the ℓ_1 -SVD algorithm [7-9], the key difference is that we not only use the SVD to reduce the computation complexity but also employ it to obtain the signal subspace and the noise subspace and, accordingly, to design the weights. Thus, we call the proposed algorithm as the SW $\ell_{2,1}$ -SVD algorithm. The experiments prove that the SW $\ell_{2,1}$ -SVD algorithm can achieve better estimation performance than ℓ_1 -SVD algorithm that exploits straightforwardly $\ell_{2,1}$ minimization. In addition, simulations and experiments on real data also demonstrate that the SW $\ell_{2,1}$ -SVD algorithm can be applied to the DOA estimation, high resolution radar imaging, and other sparse recovery related problems with the random basis matrix.

The remainder of this article is organized as follows. In the following section, we describe the sparse signal representation framework in the MMV case. In Section 3, we formulate the SW $\ell_{2,1}$ -SVD algorithm. In Section 4, the performance of the proposed method is explored with some examples. The summary is given in Section 5.

2 The ℓ_1 -SVD algorithm

For recovering the jointly-sparse signals \mathbf{X} , a feasible way is that the row support of the jointly sparse signals is first determined and then the signals can be recovered by solving a least square (LS) problem [1]. In addition, in some applications the problem of interest is to determine the row support rather than recover \mathbf{X} oneself. Therefore, in this article we consider the problem of determining the row support of the jointly sparse signals.

The Equation (3) can be expressed with matrix form:

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N}. \quad (4)$$

The truncated SVD can be exploited to hold the principal components on the measurements \mathbf{Y} [7-9]:

$$\mathbf{Y}_{\text{SV}} = \mathbf{U}\mathbf{\Sigma}\mathbf{D}_P = \mathbf{Y}\mathbf{V}\mathbf{D}_P, \quad (5)$$

where $\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$, the superscript H denotes the conjugate transpose, the non-zero entries of $\mathbf{\Sigma}$ are equal to the singular values of \mathbf{Y} and they are sorted in descending order on the diagonal; the columns of \mathbf{U} and \mathbf{V} are, respectively, left singular vectors and right singular vectors for corresponding singular values; $\mathbf{D}_P = [\mathbf{I}_P; \mathbf{0}]$, where \mathbf{I}_P is a $P \times P$ identity matrix and $\mathbf{0}$ is a $(T - P) \times P$ matrix of zeros. Moreover, let

$$\mathbf{X}_{SV} = \mathbf{XVD}_P, \quad (6)$$

and

$$\mathbf{N}_{SV} = \mathbf{NVD}_P. \quad (7)$$

Obviously, \mathbf{X}_{SV} and \mathbf{X} have the same row support. The ℓ_1 -SVD algorithm can be described as [7-9]:

$$\min \|\mathbf{X}_{SV}\|_{2,1} \quad \text{s.t.} \quad \|\mathbf{Y} - \mathbf{AX}_{SV}\|_F^2 \leq \beta^2, \quad (8)$$

where $\beta^2 \geq \|\mathbf{N}_{SV}\|_F^2$ is a regularization parameter, the mixed norm $\ell_{2,1}$ norm is defined as $\|\cdot\|_{2,1} \triangleq \sum_i (\sum_j |[\cdot]_{i,j}|^2)^{1/2}$ [22], and $\|\cdot\|_F$ denotes Frobenius norm, respectively. In practice, we select the set of the indices of the P peaks in the solution as the estimate of the row support set $\hat{\Lambda}$.

3 Subspace weighted $\ell_{2,1}$ -SVD algorithm

Here, given the MMV case we exploit the relationship between the noise subspace and the overcomplete basis to construct a weighting vector that can improve the performance of the ℓ_1 -SVD method. Incidentally, we already presented the SW method in [23], where we used an extra eigendecomposition of sample correlation matrix to obtain the noise subspace. However it is not necessary to employ the extra eigendecomposition, because the SVD of measurements \mathbf{Y} has revealed how to obtain subspace decomposition [17]. In addition, we address some interesting issues and extend the application of the SW method.

Retracing the SVD mentioned as (5), we have

$$\mathbf{U} = [\mathbf{U}_s \ \mathbf{U}_n], \quad (9)$$

where $\mathbf{U}_s = [\mathbf{u}_1, \dots, \mathbf{u}_p]$ and $\mathbf{U}_n = [\mathbf{u}_{p+1}, \dots, \mathbf{u}_M]$, which correspond to the signal subspace and noise subspace, respectively, [17].

In [19,20], it is proved that

$$\mathbf{A}_\Lambda^H \mathbf{U}_n = \mathbf{B}, \quad (10)$$

where $\mathbf{B} = [b_{i,j}]$ and $b_{i,j} \rightarrow 0$ as the number of snapshots $T \rightarrow \infty$. As a result, we have

$$\mathbf{A}^H \mathbf{U}_n = \begin{bmatrix} \mathbf{A}_\Lambda^H \mathbf{U}_n \\ \mathbf{A}_{\Lambda^c}^H \mathbf{U}_n \end{bmatrix} = \begin{bmatrix} \mathbf{B} \\ \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_\Lambda \\ \mathbf{W}_{\Lambda^c} \end{bmatrix} = \mathbf{W}, \quad (11)$$

where $\mathbf{C} = \mathbf{A}_{\Lambda^c}^H \mathbf{U}_n$. We can express the weighted vector as

$$\mathbf{W}^{(\ell_2)} = \begin{bmatrix} \mathbf{W}_\Lambda^{(\ell_2)} \\ \mathbf{W}_{\Lambda^c}^{(\ell_2)} \end{bmatrix}, \quad (12)$$

where $\mathbf{W}_{\Lambda,i}^{(\ell_2)} \rightarrow 0$ and $\mathbf{W}_{\Lambda^c,i}^{(\ell_2)} \rightarrow C_i^{(\ell_2)} > 0$ as $T \rightarrow \infty$ [19] $\mathbf{W}_{\Lambda,i}^{(\ell_2)}$, $\mathbf{W}_{\Lambda^c,i}^{(\ell_2)}$ and $C_i^{(\ell_2)}$ denote the i th entry of

$\mathbf{C}^{(\ell_2)}$, and $\mathbf{C}^{(\ell_2)}$, respectively. This is consistent with the methodology of the iterative reweighted ℓ_1 minimization, i.e., large weights are assigned to the entries whose indices are more likely to be outside of the row support, whereas small weights are assigned to the entries whose indices are inside of the row support [12,14,16]. When the limited snapshots are used in actual application, it is also guaranteed that the entries of $\mathbf{W}_\Lambda^{(\ell_2)}$ are much smaller than those of $\mathbf{W}_{\Lambda^c}^{(\ell_2)}$ [19-21].

We define

$$\mathbf{w} = \mathbf{W}^{(\ell_2)}. \quad (13)$$

The sparse solution can be found by^a

$$\min \|\mathbf{X}_{SV}\|_{w;2,1} \quad \text{s.t.} \quad \|\mathbf{Y} - \mathbf{AX}_{SV}\|_F^2 \leq \beta^2, \quad (14)$$

where $\|\cdot\|_{w;2,1} \triangleq \sum_i w_i (\sum_j |[\cdot]_{i,j}|^2)^{1/2}$, w_i denotes the i th entry of \mathbf{w} .

Some related issues are discussed as follows.

Discussion 1 : An interesting issue raised by the Equation (12) is how many snapshots are enough for the SW method to work. The weighted process can be seen as a preprocessing that obtains the rough information about the row support and the weighted values. The SW method employs the methodology of the MUSIC method [19] to achieve the preprocessing. Therefore, the SW method is consistent with the MUSIC method on the requirement of the number of snapshots. The theoretical limitation of the requirement of the number of snapshots $T \geq P$ is showed in [24,25] for the MUSIC method. In other words, the SW method is able to work with very small number of snapshots.

Discussion 2 : The prior information about the number of sources plays a key role in partitioning the noise subspace and the signal subspace. The right partition of the noise subspace and the signal subspace is beneficial to accomplish the optimal weights for the SW method. In practice, the number of sources can be determined by exploiting the information theoretic criterion such as the Akaike's information criterion (AIC) [26] and the minimum description length (MDL) criterion [27]. These methods require the eigenvalue of the sample correlation matrix $\hat{\mathbf{R}}$, where $\hat{\mathbf{R}} = \frac{1}{T} \mathbf{Y} \mathbf{Y}^H$ is a Hermitian matrix. We can use the SVD of \mathbf{Y} to obtain the eigenvalue because the eigenvalue decomposition (EVD) of a Hermitian matrix is a special case of the SVD of a general matrix [17]. As a result, we have $\Sigma_e = \frac{1}{T} \Sigma \Sigma^H$, where the elements on the diagonal of Σ_e are the eigenvalue of $\hat{\mathbf{R}}$. Therefore, the number of sources can be determined by combining the SVD of \mathbf{Y} with the information theoretic criterion. However, in some situations, for example, when the signal-to-noise ratio (SNR) is very low or the number of snapshots is very

small, the classical AIC and MDL rules are likely to overestimate or underestimate the number of sources. Thus, another interesting issue is the robustness of the proposed SW $\ell_{2,1}$ -SVD algorithm to the estimate of the number of sources. Here we give a brief explanation about the robustness of SW $\ell_{2,1}$ -SVD to the estimate of the number of sources and leave a detailed discussion to future work. For one thing, both the proposed SW $\ell_{2,1}$ -SVD algorithm and the original ℓ_1 -SVD algorithm [8] use information about the number of sources to reduce the computational complexity, in which the incorrect determination of the number of sources does not incur catastrophic consequences [8]. For another, the weighted $\ell_{2,1}$ minimization processing is not very sensitive to the determination of the number of sources. Considering two cases, i.e., the estimate of the number of sources $\hat{P} = 0$ (the extreme underestimation case) and $P < \hat{P} \leq M - 1$ (the overestimation case), where the real value of the number of sources is assumed as $0 < P < M - 1$. For the former the estimate of the noise subspace $\hat{\mathbf{U}}_n$ is equal to \mathbf{U} , and then all weights are identical because the matrix \mathbf{U} is a unitary matrix, i.e., $\|\mathbf{a}_i^H \hat{\mathbf{U}}_n\|_2 = \|\mathbf{a}_j^H \hat{\mathbf{U}}_n\|_2$ for $i \neq j$. As a result, SW $\ell_{2,1}$ -SVD becomes ℓ_1 -SVD in the extreme underestimation case. For the latter $\hat{\mathbf{U}}_n = [\mathbf{u}_{\hat{P}+1}, \dots, \mathbf{u}_M]$ and the subspace that its columns span is a true subset of the noise subspace. Therefore, the orthogonality between $\hat{\mathbf{U}}_n$ and \mathbf{A}_Λ still exists and it only incurs gradual degradation of performance because the shrunk subspace dimension will weaken the multiple averaging effect [28]. Obviously, SW $\ell_{2,1}$ -SVD can cope with the overestimation case. We illustrate this conclusion in Section 4.

4 Examples

In this section, we present some examples to demonstrate the performance of the proposed SW $\ell_{2,1}$ -SVD algorithm. We first address source localization with a uniform linear array (ULA) and a nonuniform linear array (NULA). Then we consider sparse signal recovery problem with the random basis matrix in the presence of noise. Lastly, we employ the real data to illustrate the performance of the proposed method. Here we use the CVX package for solving the convex optimization problem [29].

4.1 Source localization with ULA

We consider the ULA composed of $M = 10$ sensors separated by half a wavelength for the source localization problem. The grid is uniformly sampled with 0.1° from -90° to 90° (unless specifically stated). The overcomplete basis matrix $\mathbf{A} = [\mathbf{a}(\varphi_1), \dots, \mathbf{a}(\varphi_K)]$ is a deterministic basis matrix under this condition, where the vector $\mathbf{a}(\varphi_k)$ denotes the array steering vector and φ_k is the k th sampling grid.

4.1.1 Localization accuracy

In the first experiment we suppose that there are three signals impinging on the array from $\theta_1 = 12^\circ$, $\theta_2 = 43^\circ$, and $\theta_3 = 67^\circ$. The number of snapshots is $T = 200$. We compare the RMSE of the DOA estimates yielded by SW $\ell_{2,1}$ -SVD with those of ℓ_1 -SVD, Root-MUSIC [30], and CRB [20]. In Figure 1a, three sources are assumed to be uncorrelated; and in Figure 1b, the sources at $\theta_1 = 12^\circ$ and $\theta_2 = 43^\circ$ are coherent, whereas the source at $\theta_3 = 67^\circ$ is uncorrelated to the first two sources. The spatial smoothing technique with a 4-element smoothing subarray is employed for Root MUSIC to decorrelate the coherent signals. As can be seen from Figure 1, the Root-MUSIC algorithm that can be seen as the typical representative of the subspace-like algorithm can provide good accuracy in uncorrelated sources case, whereas it need use the spatial smoothing technique to obtain competitive performance in coherent sources case. As for the ℓ_1 -SVD algorithm that employs the regular $\ell_{2,1}$ minimization, it can yield acceptable DOA estimates in coherent sources case but does not compete with the subspace-like algorithm in uncorrelated sources case. Since the weighted $\ell_{2,1}$ minimization further consolidates the sparsity-encouraging nature of regular $\ell_{2,1}$ minimization, the SW $\ell_{2,1}$ -SVD can improve the recovery accuracy. As a result, the presented SW $\ell_{2,1}$ -SVD algorithm gives competitive DOA estimates that are closer to the CRB for both uncorrelated and coherent sources.

4.1.2 DOA tracking for mobile sources

An advantage of sparse signal representation methodology is that it has robustness to limited number of snapshots for DOA estimation [8]. Then, we design a scenario to validate the ability of the presented SW $\ell_{2,1}$ -SVD, which there are two uncorrelated moving sources in the array's viewing field and we need estimate their DOAs using a few snapshots.^b We consider two moving sources in this simulation. A source moves linearly from 30° to 21° ; the other one moves first from 0° to 3° , and then from 3° to 0° , and last from 0° to 3° . It is assumed that the moving step is 0.03° per snapshot over a course of 300 data snapshots. We use the most recent three snapshots to estimate DOAs for SNR = 12 dB. As can be seen from Figure 2, the Root-MUSIC algorithm has some strong outliers; especially some strong outliers confuse two sources over some periods. Although it is also demonstrated that the ℓ_1 -SVD algorithm can work with a few snapshots, there are some outliers, which causes degradation in the performance of the trajectories. The presented SW $\ell_{2,1}$ -SVD has a few slight outliers, which do not affect DOA tracking for mobile sources. This shows that the weighted $\ell_{2,1}$ minimization outperforms the regular $\ell_{2,1}$ minimization in the sense

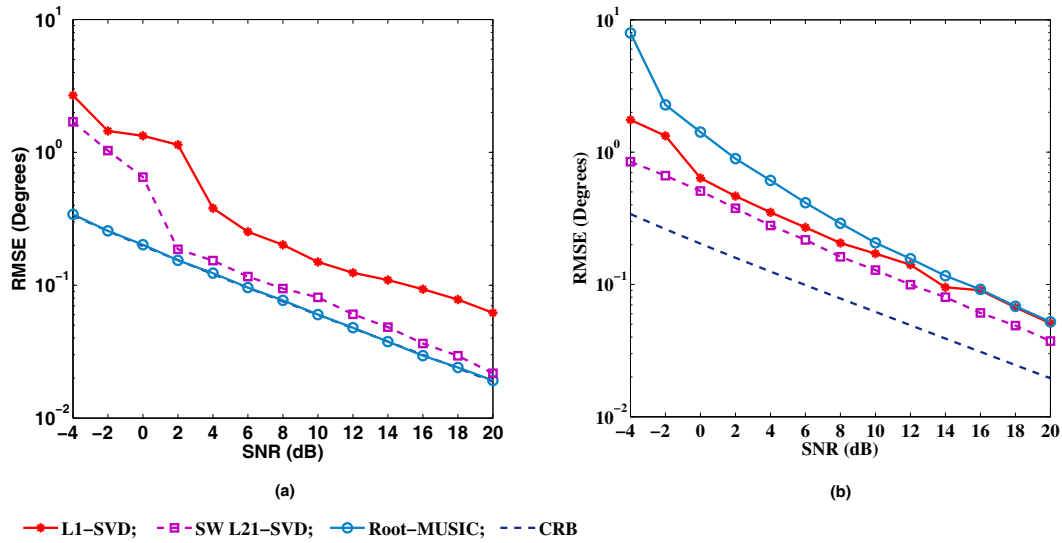


Figure 1 RMSE of angle estimation. (a) Uncorrelated sources. (b) Coherent sources. The number of snapshots is $T = 200$. Each point is average of 500 Monte-Carlo trials. Asterisk-solid curve: ℓ_1 -SVD; Square-dash curve: SW $\ell_{2,1}$ -SVD; Circle-solid curve: Root-MUSIC; Dash curve: CRLB.

that a few snapshots are employed to obtain exact sparse recovery.

4.2 Source localization with NULA

In this section, we consider the NULA composed of $M = 10$ sensors that are randomly selected from a ULA

with 20 sensors for the source localization problem. Here we only consider the coherent sources case. Again, the sources at $\theta_1 = 12^\circ$ and $\theta_2 = 43^\circ$ are coherent, and the source at $\theta_3 = 67^\circ$ is uncorrelated to other sources. In Figure 3, we show the spatial spectrum obtained with MUSIC, ℓ_1 -SVD, and SW $\ell_{2,1}$ -SVD in 100 Monte Carlo

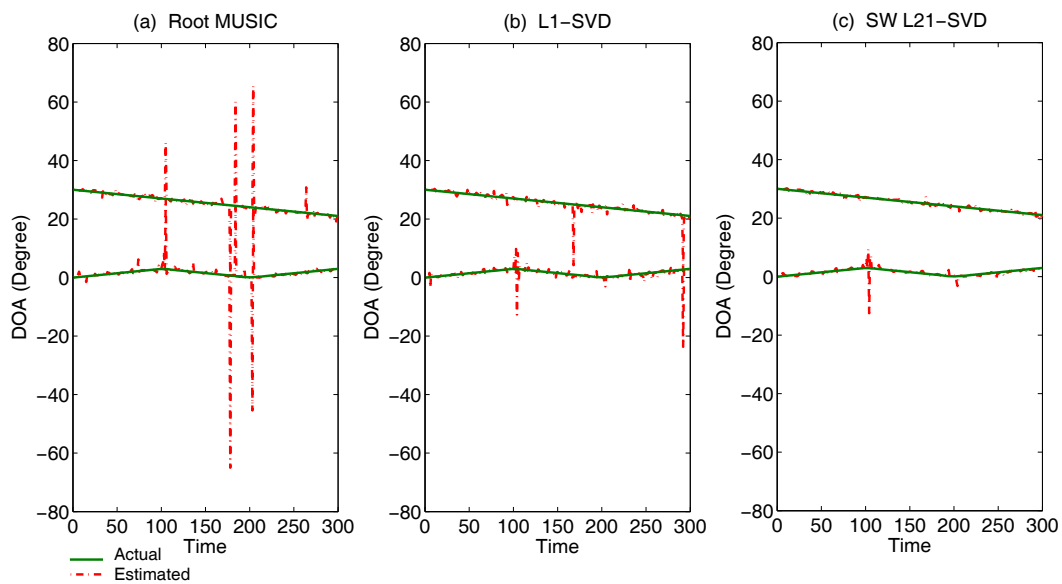
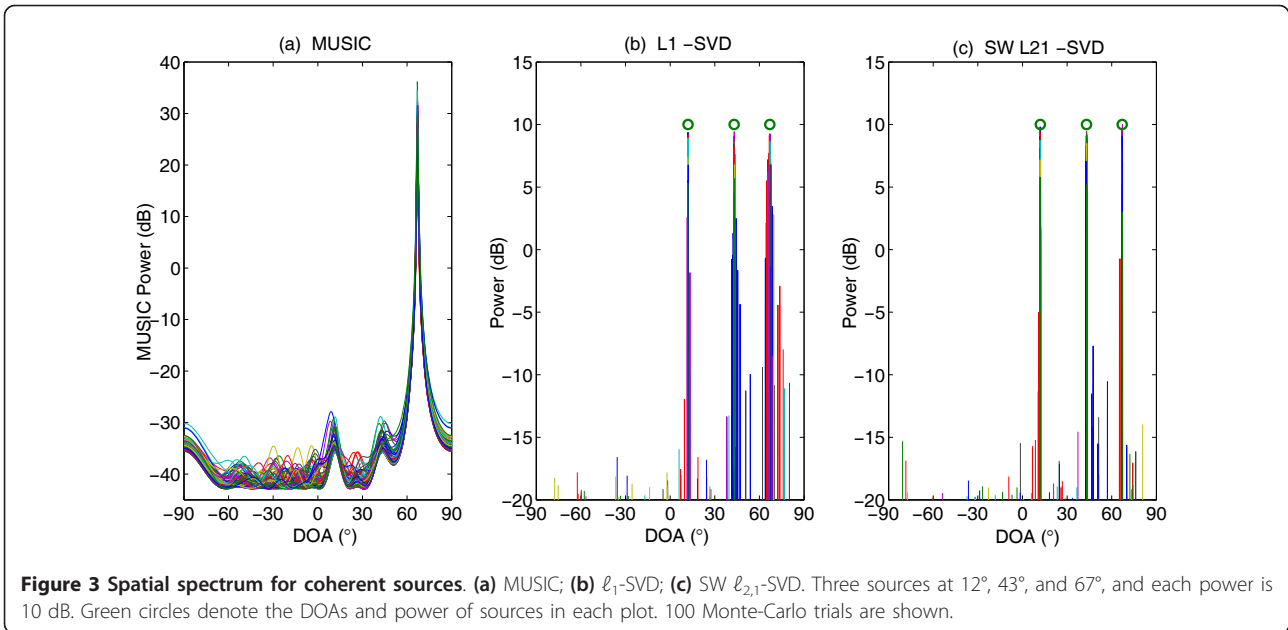


Figure 2 DOA tracking performance for two uncorrelated mobile sources. (a) Root-MUSIC; (b) ℓ_1 -SVD; (c) SW $\ell_{2,1}$ -SVD. SNR is 12 dB, the number of snapshots $T = 3$. The solid lines and “-” denote the trajectories of the true DOAs and the estimates, respectively.



runs. In this experiment SNR is 10 dB, the number of snapshots is 200. The spatial smoothing technique is valid for an ULA, but not for NULA [17]. Therefore, as it is shown in Figure 3a, MUSIC has only one significant peak because spatial smoothing does not work for NULA. However, both ℓ_1 -SVD and SW $\ell_{2,1}$ -SVD still provide good estimates. In addition, it also is noted that the proposed SW $\ell_{2,1}$ -SVD algorithm has smaller variance than that of the ℓ_1 -SVD algorithm, which is consistent with the conclusion of the Section 4.1.1, i.e., the SW $\ell_{2,1}$ -SVD algorithm has better localization accuracy than that of the ℓ_1 -SVD algorithm.

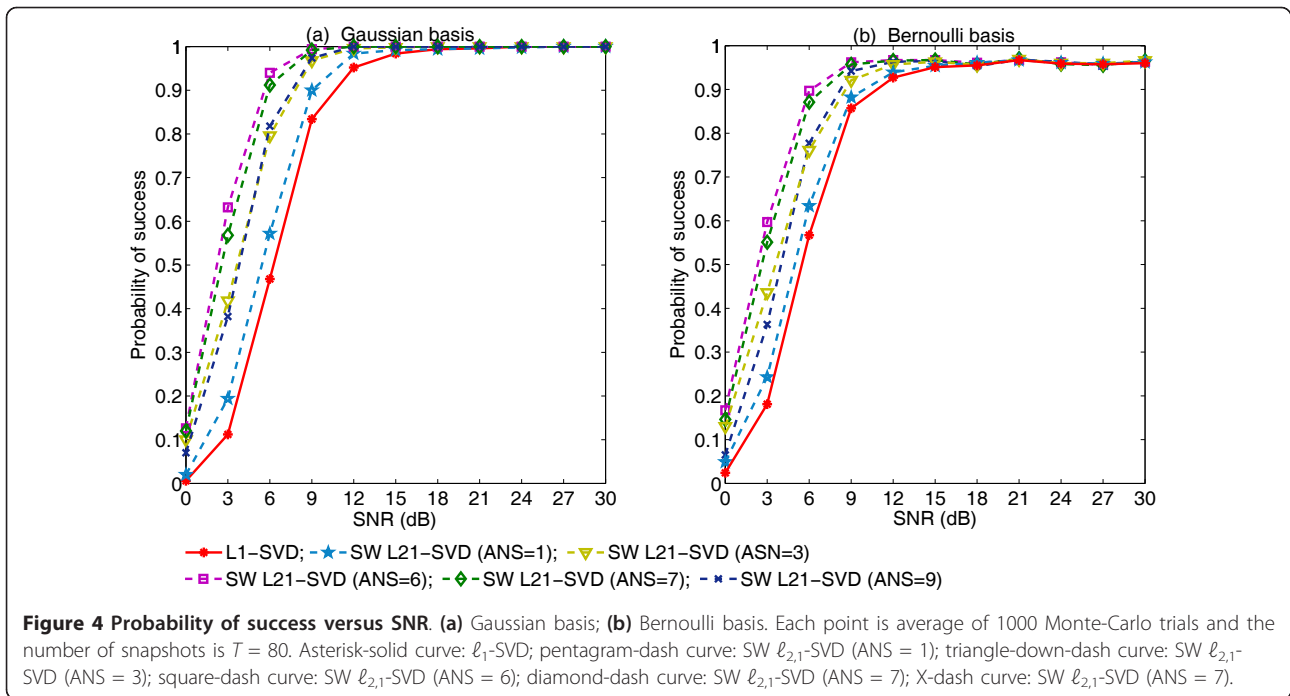
4.3 Sparse recovery for random basis matrix

A related problem is that the presented SW $\ell_{2,1}$ -SVD can be extended to other instances with a random basis matrix that accord with the model (3) in the sparse representation framework. Here, we give some examples to demonstrate the validity of the extension. We assume a MMV sparse matrix $\mathbf{X} \in \mathbb{R}^{K \times T}$ with P non-zero rows that their indices are chosen randomly, and their amplitudes is chosen randomly from a standard normal distribution. The overcomplete basis matrix $\mathbf{A} \in \mathbb{C}^{M \times K}$ is a random matrix with i.i.d. Gaussian entries or i.i.d symmetric Bernoulli ± 1 entries, and its columns are normalized. Our object is to estimate the row support of the sparse signal \mathbf{X} with the measurements $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N}$, where \mathbf{N} is an additive Gaussian white noise matrix and its variance σ^2 is determined from a specified SNR level as $\sigma^2 = \frac{\|\mathbf{X}\|_F^2}{T \times P} \times 10^{-\text{SNR}/10}$. In the experiment parameter settings are as follows: $M = 12$, $K = 30$, and $P = 6$. The set of the indices corresponding to P peaks in the

estimate $\hat{\mathbf{X}}_{\text{SV}}^{(\ell_2)}$ are regarded as the estimate of the row support of the jointly sparse signals. The estimate of the row support $\hat{\Lambda}$ is considered to be correct if and only if it is fully consistent with the true row support Λ . As demonstrated in Figures 4 and 5, the SW $\ell_{2,1}$ -SVD algorithm improves the recovery performance, especially reducing both the SNR requirement and the required number of snapshots for stable recovery. In addition, for exploring robustness to the number of sources, we employ the assumed number of sources (ANS) to perturb the SW $\ell_{2,1}$ -SVD algorithm. It is shown in Figures 4 and 5 that the SW $\ell_{2,1}$ -SVD algorithm accomplishes the optimal performance when ANS is equal to the number of sources P (i.e., $\text{ANS} = P = 6$). Furthermore, they also demonstrate that the SW $\ell_{2,1}$ -SVD algorithm can cope with both the overestimation case and the underestimation case and it still excels the ℓ_1 -SVD algorithm in these cases.

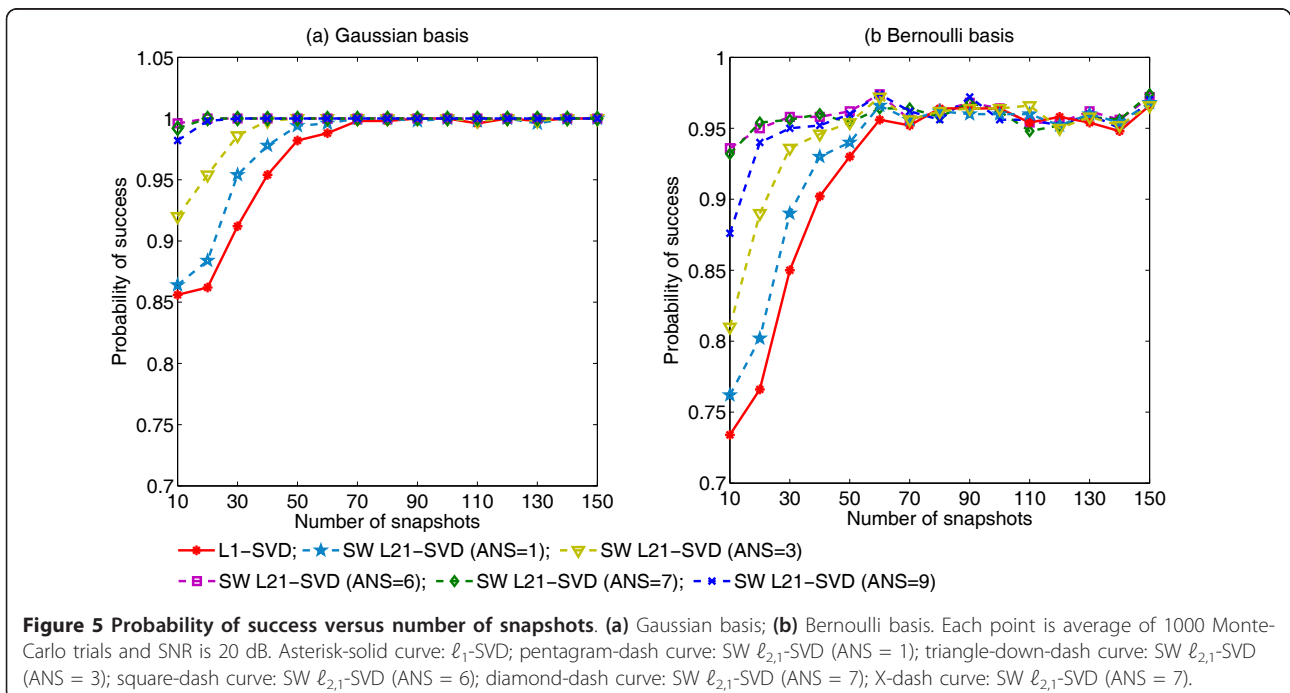
4.4 High resolution radar imaging via sparse recovery

Here we attempt to obtain high range resolution in data collected by a real stepped frequency radar. The radar is Ka band and the frequency step size Δf is 8 MHz, and the pulse repeat interval (PRI) is 0.15 ms. In the observed scene two Corner Reflectors separated by 0.4 m are fixed on a straight road and they are collinear with radar. The width of the transmitted pulses is $1 \mu\text{s}$ and 64 pulses data are collected. The data model can be written as $y(n) = \sum_{i=1}^P \beta_i r_i(n) + w(n)$, where β_i is the complex scattering intensity of the i th target, $r_i(n) = \exp[j\frac{4\pi}{v_c}(f_0 + n\Delta f)R_i]$, v_c is the speed of light, R_i is



the distance between the i th target and radar, f_0 is the initial frequency, $w(n)$ denotes the noise. Without loss of generality, in the recovery the item $\frac{2}{v_c}(f_0 + n\Delta f)R_i$ is normalized (it is called as the normalized frequency in the context) and the interval $[0 \ 1]$ is uniformly sampled with 1024 grids.

In Figure 6, the normalized frequency spectrum obtained from MUSIC, ℓ_1 -SVD, and SW $\ell_{2,1}$ -SVD are displayed, where the data window size is 32, i.e., 33 snapshots can be obtained by sliding the window. In this case the mentioned algorithms clearly discern the two targets. Compared with the sparse recovery



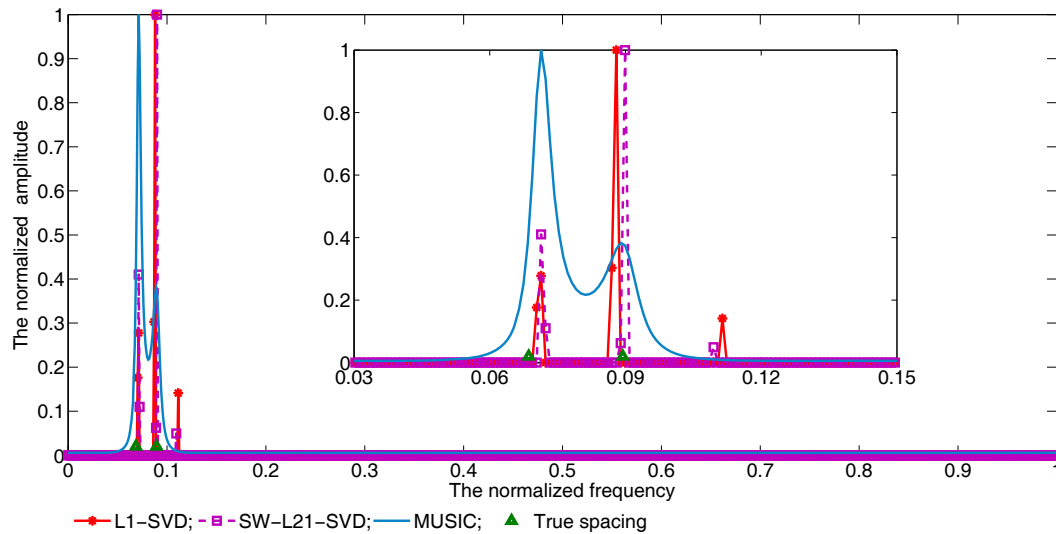


Figure 6 The normalized frequency spectrum for real measurement data. The data window size is 32. Asterisk-solid curve: ℓ_1 -SVD; square-dash curve: SW $\ell_{2,1}$ -SVD; solid curve: MUSIC; triangle-up: true spacing.

algorithm, however, the peak obtained from MUSIC is obtuse. In addition, ℓ_1 -SVD has a stronger spurious peak that may confuse the true target with spurious peaks. The confidence interval exerts an influence on how many spurious peaks exist in the solution. Although increasing the confidence interval can

suppress spurious peaks to some extent, as it is shown in Figure 7, we may adopt a pessimistic attitude for ℓ_1 -SVD. It is worthwhile to note that SW $\ell_{2,1}$ -SVD has a very slight spurious peak so that we can ignore it, especially for a higher confidence interval. Furthermore, the confidence interval determines the size of the

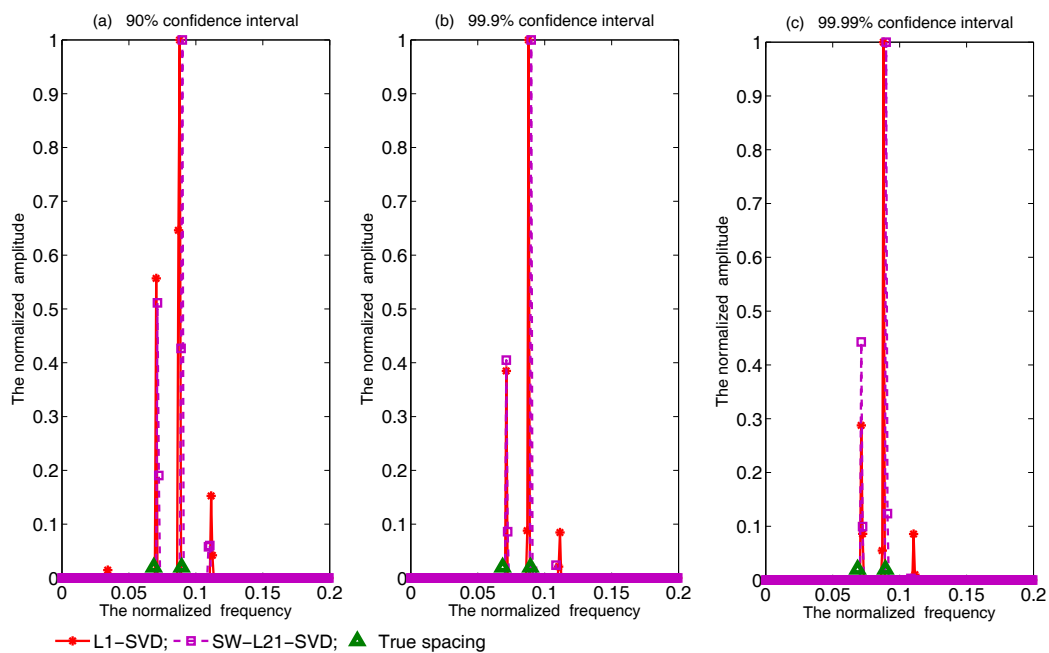


Figure 7 Robustness of SW $\ell_{2,1}$ -SVD to the confidence interval. (a) 90% confidence interval; (b) 99.9% confidence interval; (c) 99.99% confidence interval. The data window size is 32. Asterisk-solid curve: ℓ_1 -SVD; square-dash curve: SW $\ell_{2,1}$ -SVD; triangle-up: true spacing.

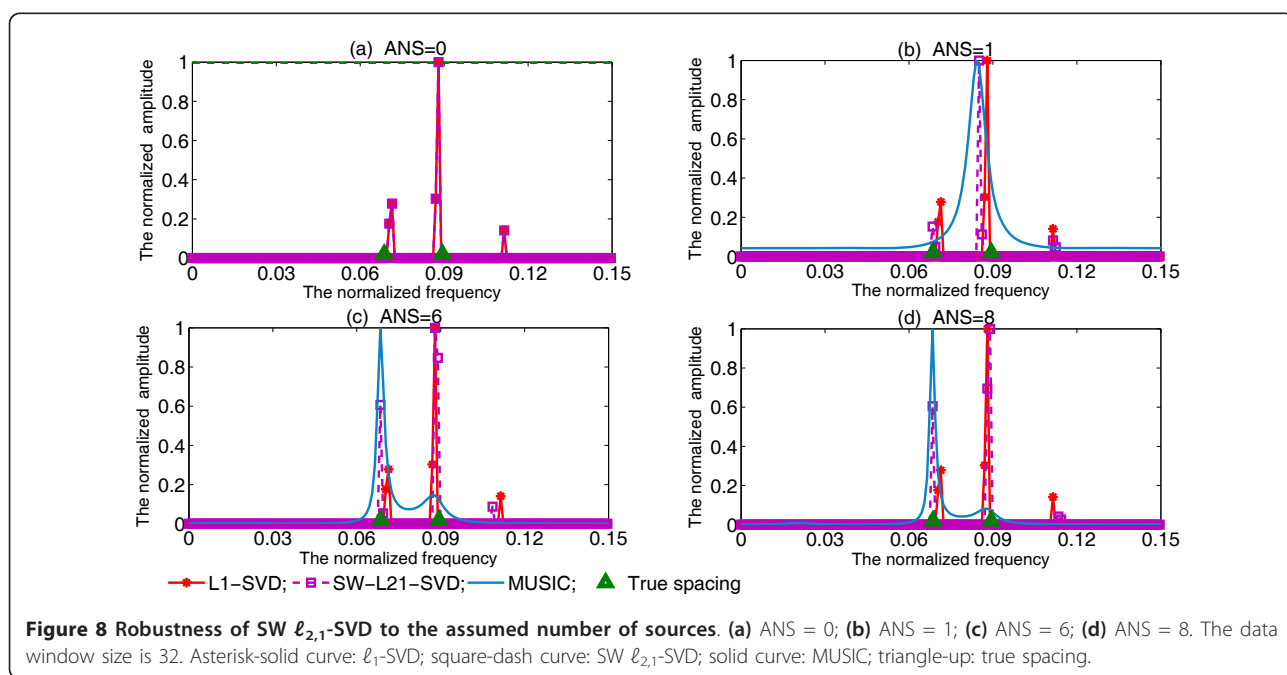
regularization parameter, i.e., the higher the confidence interval, the larger the regularization parameter. Figures 6 and 7 show clearly SW $\ell_{2,1}$ -SVD has a good performance when the size of the regularization parameter varies widely.

Next, we illustrate the robustness of the proposed SW $\ell_{2,1}$ -SVD algorithm to the estimate of the number of sources. In Figure 8, we artificially adjust the number of sources so that we can explore what extent underestimating or overestimating the number of sources affects the performance of the SW $\ell_{2,1}$ -SVD algorithm. The data window size is fixed at 32 and the assumed number of sources (ANS) is exploited to determine the dimension of the noise subspace and reduce the computational complexity (unless specially stated). Note that ANS = 0 is used only for determining the dimension of the noise subspace and $\mathbf{D}_p = [\mathbf{I}_2; \mathbf{0}]$ is utilized to reduce the computational complexity in Figure 8a. It is illustrated in Figure 8, in the extreme underestimation case, i.e., ANS = 0, the result of SW $\ell_{2,1}$ -SVD and ℓ_1 -SVD is identical, which corroborates ℓ_1 -SVD is a special example of SW $\ell_{2,1}$ -SVD. In addition, we can observe the fact SW $\ell_{2,1}$ -SVD can clearly discern the two targets in both underestimation and overestimation cases, whereas MUSIC is invalid in the underestimation case. Therefore, this illustration shows that underestimation and overestimation of the number of sources do not incur catastrophic consequences for SW $\ell_{2,1}$ -SVD.

In Figure 9, we compare the estimation accuracy of the distance between two Corner Reflector obtained from MUSIC, ℓ_1 -SVD, and SW $\ell_{2,1}$ -SVD as a function of the data window size M (the number of snapshots $T = 64 - M + 1$). When the data window size is too small or too large, MUSIC does not provide reliable results. Therefore, MUSIC need carefully select the data window size so as to it can obtain reliable estimates. Although the performance of the ℓ_1 -SVD algorithm is better than that of MUSIC for large data window size (i.e., $M \geq 61$), it not give reliable results for small data window size (i.e., $M \leq 7$). For the same experiment context, as expected, SW $\ell_{2,1}$ -SVD yields competitive performance and it is robust to the data window size.

5 Conclusion

In this article, we proposed an effective weighted $\ell_{2,1}$ minimization algorithm that exploits the relationship between the noise subspace and the overcomplete basis matrix to obtain the weights for the jointly-sparse signal recovery problem. The proposed SW $\ell_{2,1}$ -SVD algorithm appoints the large weights to those, whose indices are more likely to be outside of the support so that their indices are banished from the support. This can further promote the sparseness of the solution at right positions. We provided experimental results to testify that, for both deterministic basis matrix and random basis matrix, the proposed SW $\ell_{2,1}$ -SVD algorithm can obtain



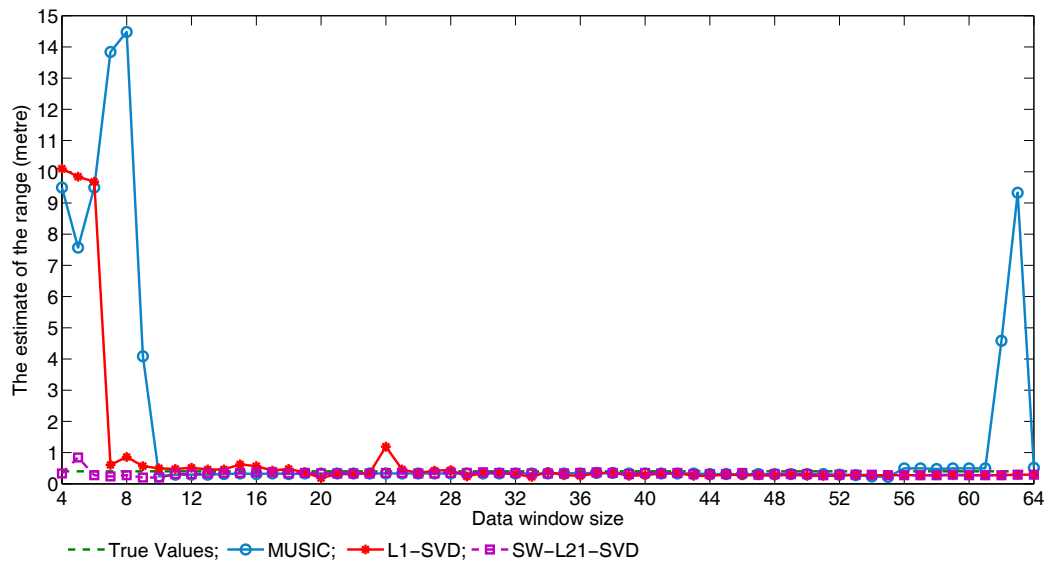


Figure 9 The estimation accuracy of the distance versus the data window size. Asterisk-solid curve: ℓ_1 -SVD; square-dash curve: SW $\ell_{2,1}$ -SVD; solid-circles curve: MUSIC; dash curve: true distance.

better performance than that of the ℓ_1 -SVD algorithm with fewer number of snapshots and lower SNR.

Endnotes

The material in this article was presented in part at the 2011 International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2011), May 2011, Prague, Czech Republic. ^aIn the context, we use the method introduced in [8] to determine the regularization parameter β^2 , and the confidence interval that controls the size of β^2 is set to 99% (unless specially stated). ^bAlthough the indices of non-zero rows of $\mathbf{x}(t)$ are dependent of the snapshot t for the moving sources, $\mathbf{x}(t)$ can still be seen as the jointly-sparse signals providing the number of snapshots T is very small.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grant 40901157, and in part by the National Basic Research Program of China (973 Program) under Grant 2010CB731901, and in part by the Doctoral Fund of Ministry of Education of China under Grant 200800031050, and in part by Tsinghua National Laboratory for Information Science and Technology (TNList) Cross-discipline Foundation.

Competing interests

The authors declare that they have no competing interests.

Received: 30 August 2011 Accepted: 2 May 2012 Published: 2 May 2012

References

1. S Cotter, B Rao, K Engan, K Kreutz-Delgado, Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Trans Signal Process.* **53**(7), 2477–2488 (2005)
2. J Chen, X Huo, Theoretical results on sparse representations of multiple-measurement vectors. *IEEE Trans Signal Process.* **54**(12), 4634–4643 (2006)

3. A Fletcher, S Rangan, V Goyal, K Ramchandran, Denoising by sparse approximation: Error bounds based on rate-distortion theory. *EURASIP J Appl Signal Process.* **2006**, 1–19 (2006)
4. M Mishali, Y Eldar, Reduce and boost: Recovering arbitrary sets of jointly sparse vectors. *IEEE Trans Signal Process.* **56**(10), 4692–4702 (2008)
5. J Lv, Y Fan, A unified approach to model selection and sparse recovery using regularized least squares. *Annals Stat.* **37**(6A), 3498–3528 (2009). doi:10.1214/09-AOS683
6. E Berg, M Friedlander, Joint-sparse recovery from multiple measurements. *Computing Research Repository - CORR abs/0904.2* 1–19 (2009).
7. D Malioutov, M Cetin, A Willsky, Source localization through a Laplacian prior: an SVD-based approach, in *IEEE Workshop on Statistical Signal Processing*, (St Louis, MO., USA, 2003), pp. 573–576
8. 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)
9. D Malioutov, A sparse signal reconstruction perspective for source localization with sensor arrays, in *Master's thesis*, (Mass. Inst. Technol., Cambridge, MA, 2003), http://ssg.mit.edu/~dmm/publications/malioutov_MS_thesis.pdf
10. J Zheng, M Kaveh, H Tsuji, Sparse spectral fitting for direction of arrival and power estimation, in *IEEE/SP 15th Workshop on Statistical Signal Processing*, (Cardiff University and City Hall Cardiff, U.K., 2009), pp. 429–432
11. Z Tang, G Blacchiere, G Leus, Aliasing-free wideband beamforming using sparse signal representation. *IEEE Trans Signal Process.* **59**(7), 3464–3469 (2011)
12. E Candes, M Wakin, S Boyd, Enhancing sparsity by reweighted ℓ_1 minimization. *J Fourier Anal Appl.* **14**(5), 877–905 (2008). doi:10.1007/s00041-008-9045-x
13. D Wipf, S Nagarajan, Iterative reweighted ℓ_1 and ℓ_2 methods for finding sparse solutions. *IEEE J Sel Top Signal Process.* **4**(2), 317–329 (2010)
14. W Xu, M Khajehnejad, A Avestimehr, B Hassibi, Breaking through the thresholds: an analysis for iterative reweighted ℓ_1 minimization via the grassmann angle framework, in *2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Dallas, TX, USA, 5498–5501 (2010)
15. D Needell, Noisy signal recovery via iterative reweighted ℓ_1 -minimization, in *Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, Calif, USA, 113–117 (2009)
16. M Khajehnejad, W Xu, A Avestimehr, B Hassibi, Analyzing weighted ℓ_1 minimization for sparse recovery with nonuniform sparse models. *IEEE Trans Signal Process.* **59**(5), 1985–2001 (2011)

17. P Stoica, R Moses, *Spectral Analysis of Signals* (Pearson/Prentice Hall, Upper Saddle River, NJ, 2005)
18. I Fevrier, S Gelfand, M Fitz, Reduced complexity decision feedback equalization for multipath channels with large delay spreads. *IEEE Trans Commun.* **47**(6), 927–937 (1999). doi:10.1109/26.771349
19. R Schmidt, Multiple emitter location and signal parameter estimation. *IEEE Trans Antennas Propag.* **34**(3), 276–280 (1986). doi:10.1109/TAP.1986.1143830
20. P Stoica, N Arye, MUSIC, maximum likelihood, and Cramer-Rao bound. *IEEE Trans Acoustics Speech Signal Process.* **37**(5), 720–741 (1989). doi:10.1109/29.17564
21. M Kaveh, A Barabell, The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise. *IEEE Trans Acoustics Speech Signal Process.* **34**(2), 331–341 (1986). doi:10.1109/TASSP.1986.1164815
22. M Kowalski, Sparse regression using mixed norms. *Appl Comput Harmonic Anal.* **27**(3), 303–324 (2009). doi:10.1016/j.acha.2009.05.006
23. C Zheng, G Li, H Zhang, X Wang, An approach of DOA estimation using noise subspace weighted ℓ_1 minimization, in *2011 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Prague Czech, 2856–2859 (2011)
24. A Paulraj, V Reddy, T Shan, T Kailath, Performance analysis of the MUSIC algorithm with spatial smoothing in the presence of coherent sources, in *IEEE Military Communications Conference-Communications-Computers: Teamed for the 90's, 1986 MILCOM*, Monterey, Calif, USA, **3**, 41.5.1–41.5.5 (1986)
25. E Sason, Source localization based on sparse signal representation, http://www.sipl.technion.ac.il/new/Archive/Annual_Proj_Pres/sipl2010/Posters/Source%20localization.pdf
26. H Akaike, A new look at the statistical model identification. *IEEE Trans Automatic Control.* **19**(6), 716–723 (1974). doi:10.1109/TAC.1974.1100705
27. J Rissanen, Modeling by shortest data description. *Automatica.* **14**(5), 465–471 (1978). doi:10.1016/0005-1098(78)90005-5
28. SJ Orfanidis, *Optimum Signal Processing*, 2nd edn. (2007), <http://eceweb1.rutgers.edu/~orfanidi/osp2e/>
29. M Grant, S Boyd, CVX, Matlab software for disciplined convex programming, version 1.21 (2010), <http://cvxr.com/cvx/>
30. A Barabell, Improving the resolution performance of eigenstructure-based direction-finding algorithms, in *IEEE International Conference on Acoustics, Speech, Signal Processing*, Boston, Mass., USA, **8**, 336–339 (1983)

doi:10.1186/1687-6180-2012-98

Cite this article as: Zheng et al.: Subspace weighted $\ell_{2,1}$ minimization for sparse signal recovery. *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:98.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
