General similar sensing matrix pursuit: An efficient and rigorous reconstruction algorithm to cope with deterministic sensing matrix with high coherence

Jing Liu a,*, Mahendra Mallick b, Feng Lian a, ChongZhao Han a, MingXing Sheng a, XiangHua Yao a

a School of Electronics and Information Engineering, Xi’an Jiaotong University, Xi’an, PR China
b Independent Consultant, Anacortes, WA, USA

A R T I C L E   I N F O

Article history:
Received 31 March 2014
Received in revised form 1 December 2014
Accepted 2 March 2015
Available online 11 March 2015

Keywords:
Compressed sensing
Deterministic sensing matrix with high coherence
General similar sensing matrix pursuit (GSSMP) algorithm

A B S T R A C T

In this paper, a novel algorithm, called the general similar sensing matrix pursuit (GSSMP), is proposed to deal with the deterministic sensing matrix with high coherence. First, the columns of the sensing matrix are divided into a number of similar column groups based on the similarity distance. Each similar column group presents a set of coherent columns or a single incoherent column, which provides a unified framework to construct the similar sensing matrix. The similar sensing matrix is with low coherence provided that the minimum similar distance between any two condensed columns is large. It is proved that under appropriate conditions the GSSMP algorithm can identify the correct subspace quite well, and reconstruct the original $K$-sparse signal perfectly. Moreover, we have enhanced the proposed GSSMP algorithm to cope with the unknown sparsity level problem, by testing each individual contributing similar column group one by one to find the true vectors spanning the correct subspace. The simulation results show that the modified GSSMP algorithm with the contributing similar column group test process can estimate the sparse vector representing the radar scene with an unknown number of targets successfully.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Compressed sensing has received considerable attention recently, and has been studied in diverse fields, e.g., image processing [1], underwater acoustic communication [2], wireless communication [3] and radar [4–6]. The central goal of compressed sensing is to capture attributes of a signal using very few measurements. In most work to date, this broader objective is exemplified by the important special case in which a $K$-sparse vector $x \in \mathbb{R}^N$ (with $N$ large) is to be reconstructed from a small number $M$ of linear measurements with $K < M < N$. $K$-sparse signals are the signals that can be represented by $K$ significant coefficients over an $N$-dimensional basis. This can be compactly described by

$$y = \Phi x + e,$$

where $y \in \mathbb{R}^M$ denotes a measurement vector, $\Phi$ represents an $M \times N$ sensing matrix, and $e \in \mathbb{R}^M$ is an arbitrary noise vector with $\|e\|_2 \leq \epsilon$, where $\epsilon$ is the bound on the noise magnitude.

In compressed sensing, one of the well-studied conditions on the sensing matrix which guarantees stable recovery for a number of reconstruction algorithms is the...
Coherence plays a central role in the sensing matrix construction, because small coherence implies the RIP [9]. In the early work of compressed sensing, the entries of the sensing matrix are generated by an independent identically distributed (i.i.d.) Gaussian or Bernoulli process, or from random Fourier ensembles [10–12]. The role of random measurement provides the worst case performance that guarantees in the context of an adversarial signal/error model. Random sensing matrices are easy to construct, and are 2K-RIP with high probability [12]. However, the random matrices are often not feasible for real-world applications due to the cost of multiplying arbitrary matrices with signal vectors of high dimension, and there is no guarantee that a specific realization works perfectly for the reconstruction.

With the application area of compressed sensing extended to wider fields, the random sensing matrix is being replaced by the deterministic sensing matrix. Recent research has focused on the construction of the deterministic matrix which often exhibits considerable structure [13]. In [14], the authors propose a new deterministic low-storage construction of compressive sampling matrices based on classical finite-geometry generalized polygons. This can be seen as a foundation on deterministic sensing matrices and reconstruction. The connection between sensing matrices and coding theory can be traced to [15], where the theory of finite classical generalized polygons is utilized to derive and study low-density parity-check (LDPC) codes. Among the approaches for deterministic matrix construction, the Vandermonde matrices seem to be good options, since any $K$ columns of an $M \times N$ Vandermonde matrix are linearly independent. However, when $N$ increases, the constant $\delta_K$ rapidly approaches 1 and some of the $M \times K$ submatrices become ill-conditioned [16]. In [17], the second order Reed–Muller codes are used to construct bipolar matrices. However, they lack a guarantee on the RIP order. In [18], the authors propose a series of deterministic sensing matrices, the binary, bipolar, and ternary compressed sensing matrices which satisfy the RIP condition.

The key concept of coherence is extended to pairs of orthonormal bases. This enables a new choice for the sensing matrices: one simply selects an orthonormal basis that is incoherent with the sparsity basis, and obtains measurements by selecting a subset of the coefficients of the signal in the chosen basis [19]. This approach has applications in radar systems [6,20], where an additional sensing matrix $H$ is introduced and the received signal is compressed further by making nonadaptive, linear projections of the direct data sampled at the Nyquist frequency. However, neither of these algorithms mention the hardware implementation of the additional sensing matrix, which is very complex and expensive.

In this work, we focus on the deterministic sensing matrix built directly on the real acquisition systems, e.g., radar systems [21], and sensor array systems [22], which have been widely used in underwater acoustics and wireless communications. In the context of phased array radar system based on space time adaptive processing (STAP) technique [21], the sensing matrix is composed of spatial-Doppler steering vectors in columns, which is deterministic in nature. As the resolution of the angle-Doppler plane becomes finer, the coherence between the columns of the sensing matrix increases, thereby degrading the reconstruction reliability and performance. Similarly, for the sensor array system [22], the sensing matrix is composed of DOA steering vectors in columns. Increasing the resolution of the DOA angle plane leads to finer gridding, which increases the correlation between the basis elements of the sensing matrix.

Only a few papers have been published on the deterministic sensing matrix with high coherence [23–25]. Sparse Bayesian learning (SBL) algorithm [26–28] is capable of handling the sensing matrix with high coherence, and has been applied in the passive SAR radar system to improve the imaging resolution [23]. In [24], a novel approach based on the SBL algorithm is proposed for sparse nonstationary signal reconstruction using multiple windows.

In our previous work [25], a novel algorithm, called the similar sensing matrix pursuit (SSMP), is proposed to cope with the deterministic sensing matrix with high coherence. The proposed algorithm builds a similar sensing matrix based on the original sensing matrix, which has low coherence. A subspace pursuit (SP) algorithm is then used to find a rough estimate of the true support set, which contains the indices of the columns that contribute to the original sparse vector. Three kinds of structures of the estimated support set are considered, and three individual refined estimation processes are carried out under these three conditions. The proposed algorithm obtains much better performance while coping with a deterministic sensing matrix with high coherence, compared with the SP and basis pursuit (BP) algorithms. However, the proposed algorithm is heuristic in nature, which implies that the original $K$-sparse vector can be reconstructed based on the similar sensing matrix with high probability, and there is no rigorous proof for this conclusion. Moreover, two thresholds have to be set in the algorithm. One threshold is used to distinguish the incoherent columns from the coherent columns, and the other is used to divide the coherent columns into separate similar column groups. The setting of the two thresholds is very tight, which limits the application area of the SSMP algorithm. This method is complex and impractical for real-world problems, since the incoherent and coherent columns are treated separately and three individual refined estimation processes under three different conditions have to be considered.

This paper presents a rigorous version of the SSMP algorithm in a unified framework, which is named as general...
similar sensing matrix pursuit (GSSMP) algorithm. First, a natural clustering method is used to divide the columns of the sensing matrix into a number of similar column groups based on the similarity distance, avoiding using two rigid thresholds as in the SSMP algorithm. Each similar column group presents a set of coherent columns or a single incoherent column, and this provides a unified framework to construct the similar sensing matrix. The similar sensing matrix is then built by drawing a condensed column from each similar column group. It is proved that the similar sensing matrix has low coherence provided that the minimum similar distance between any two condensed columns is large.

We use the theory of subspace pursuit to analyze the convergence of the GSSMP algorithm. This approach is different from our previous work on developing SSMP. In compressed sensing, the major challenge associated with \( K \)-sparse signal reconstruction is to identify in which subspace (generated by not more than \( K \) columns of the sensing matrix \( \Phi \)) the measured signal \( y \) lies [29]. Once the correct subspace is determined, the nonzero signal coefficients are calculated by applying the pseudo-inverse. The key procedure of the greedy algorithms, e.g., the matching pursuit (MP) [30], orthogonal matching pursuit (OMP) [31], regularized OMP (ROMP) [32], stagewise OMP (StOMP) [33], SP [29], and compressive sampling matching pursuit (CoSaMP) [34] algorithms, is the method used in finding the columns that span the correct subspace. In this paper, we prove that under appropriate conditions the GSSMP algorithm can identify the correct subspace quite well, and reconstruct the original \( K \)-sparse signal perfectly.

Moreover, the proposed GSSMP algorithm is modified to cope with the unknown sparsity level problem, by changing the original refined estimation process, i.e., the way of finding the columns spanning the correct subspace. The new refined estimation algorithm is named as contributing similar column group test algorithm, which tests each individual contributing similar column group one by one to find the true vectors spanning the correct subspace.

Finally, we present two simulation examples to verify the performance of the proposed GSSMP algorithm. In the first example, the chirp matrix with low dimension is chosen as the sensing matrix, which has high coherence. The simulation results demonstrate that the proposed GSSMP algorithm outperforms the BP and SP algorithms in reconstruction accuracy. Moreover, since the deterministic sensing matrix always depends on specific applications, the direction-of-arrival (DOA) estimation problem using compressed sensing for sensor array processing is addressed in the second example. The proposed GSSMP algorithm can reconstruct the sparse DOA angle plane with high resolution and identify multiple targets more accurately than the SBL [26], spatial compressed sensing (SCS) [22], multiple signal classification (MUSIC) [35] and focal underdetermined system solver (FOCUSS) [36] algorithms.

The paper is organized as follows. Section 2 introduces the proposed GSSMP algorithm. Section 3 contains the main result of the paper: a formal proof of the guaranteed reconstruction performance of the GSSMP algorithm. Section 4 analyzes the computational complexity of the proposed algorithm. Section 5 introduces the modified GSSMP algorithm with the contributing similar column group test process, and the simulation results are presented in Section 6. Finally, Section 7 summarizes conclusions.

2. General similar sensing matrix pursuit (GSSMP) algorithm

2.1. Intuition

In compressed sensing, the major challenge associated with \( K \)-sparse signal reconstruction is to identify in which subspace (generated by not more than \( K \) columns of the sensing matrix \( \Phi \)) the measured signal \( y \) lies [29]. Once the correct subspace is determined, the nonzero signal coefficients are calculated by applying the pseudo-inverse. The key procedure of the greedy algorithms is the method used for finding the columns that span the correct subspace. An important prior for the greedy algorithms is that the sensing matrix satisfies the RIP or has low coherence, which implies that the columns of the sensing matrix are locally near-orthogonal (incoherent with each other). However, in practice, it is challenging to design a deterministic sensing matrix having low coherence. The greedy algorithms always fail in reconstructing the \( K \)-sparse vector based on the deterministic sensing matrix with high coherence, since they cannot find the right column from a group of coherent (similar) columns to form the correct subspace.

In this paper, we consider a general condition when a deterministic sensing matrix has high coherence. A novel GSSMP algorithm is proposed to reconstruct the \( K \)-sparse signal, which is to build a similar sensing matrix based on the original sensing matrix. Each condensed column of the similar sensing matrix represents the characteristics of a group of coherent columns or a single incoherent column in the original sensing matrix. It can be easily proved that the similar sensing matrix has low coherence provided that the minimum similar distance between any two condensed columns is large. A rough estimate of the correct subspace is obtained based on the condensed columns of the newly built similar sensing matrix. Each condensed column of the estimated subspace indicates a group of coherent columns or a single incoherent column in the original sensing matrix, which contribute to the correct subspace and are termed as candidate columns. A combinatorial search is then performed among all the candidate columns to find the columns forming the correct subspace. It is proved that under appropriate conditions the GSSMP algorithm can identify the correct subspace quite well.

In the following sections, the similarity analysis of the original sensing matrix is performed in Section 2.2, the construction process of the similar sensing matrix is introduced in Section 2.3, and the complete GSSMP algorithm is introduced in Section 2.4.

2.2. Similarity analysis of the original sensing matrix

The construction process of the similar sensing matrix is based on the similarity analysis of the original sensing matrix. In this paper, similarity is defined as the absolute and normalized inner product between any two different
columns in the original sensing matrix \( \Phi \),
\[
\lambda(\phi_i; \phi_j) = \frac{||\phi_i^T \phi_j||}{||\phi_i||_2 \cdot ||\phi_j||_2}, \quad 1 \leq i, j \leq N \text{ and } i \neq j.
\] (3)

The number of similarity values of \( \Phi \) is \( N(N-1)/2 \). It can be seen that coherence is the largest similarity among the columns of a matrix. Any two columns with high similarity are coherent with each other, and vice versa.

The relationship between the columns of the sensing matrix is efficiently represented using the similarity. A scenario is considered where each column is represented as a point in a \( \mathbb{R}^M \) space, named the similarity space, where \( M' \) denotes the dimension of the similarity space, with \( M' \leq M \). The “distance” between any two points (columns) is equal to the “distance” related with similarity rather than the traditional Euclidean distance.

The detailed procedure is as follows. All the columns are mapped to a similarity space to obtain a number of column points, which are indicated as \( \{P_{\phi_1}, \ldots, P_{\phi_i}, \ldots, P_{\phi_N}\} \), with \( P_{\phi_i} \) denoting the column point corresponding to the column \( \phi_i \). The similarity distance \( d_{\text{similar}} \) is defined as the “distance” between any two column points \( P_{\phi_i} \) and \( P_{\phi_j} \) as
\[
d_{\text{similar}}(P_{\phi_i}, P_{\phi_j}) = 1 - \lambda(\phi_i; \phi_j). \tag{4}
\]

Therefore the distribution range of the similarity distance is \([0, 1]\), and the closer the two column points locate in the similarity space (with smaller distance), the more similar their corresponding columns are (with larger similarity).

A hierarchical clustering method is then used to cluster the column points. First, the similarity distance between every pair of column points is calculated. Second, the column points are grouped into a binary, hierarchical cluster tree, where pairs of column points that are in close proximity in similarity distance are linked to each other. As column points are paired into binary clusters, the newly formed clusters are grouped into larger clusters until a hierarchical tree is formed. Finally, the cutoff criterion is set to prune branches off the bottom of the hierarchical tree, and all the column points below each cutoff are assigned to a single cluster, resulting in a number of clusters. Therefore the cluster number is determined once the cutoff criterion is set. The setting of the cutoff criterion is based on the minimum similarity distance between clusters (similar column groups), which will be discussed in Section 3.

Each cluster contains one or more column points. For the cluster containing more than one column points, the similarity distance between any two column points is small. Accordingly, the columns from the same cluster are similar (coherent) with each other, which are defined as a similar column group. It should be noted that a similar column group may only contain one column.

Fig. 1 shows the classification results in a similarity space. Here we choose a two dimensional representation of the similarity space for clarity and simplicity. The column point and the similar column group are indicated by small square and large circle, respectively. \( d_{\text{similar}} \) indicates the similarity distance between any two column points. A similar column group may contain one or more columns. Any two columns in a similar column group are with small similarity distance, which are very similar with each other. Next, we build the similar sensing matrix based on the similar column groups.

2.3. Construction of the similar sensing matrix

All the columns of the sensing matrix are mapped to a similarity space to obtain a number of column points \( \{P_{\phi_1}, \ldots, P_{\phi_i}, \ldots, P_{\phi_N}\} \). A hierarchical clustering method is then used to cluster the column points based on the similarity distance, and we can obtain a number of clusters containing one or more column points. The columns from the same cluster constitute a similar column group. As a result, all the columns of the sensing matrix are divided into \( D \) similar column groups, \( \{\Gamma_1, \Gamma_2, \ldots, \Gamma_i, \ldots, \Gamma_D\} \). We assume that \( M/D < N \). Each similar column group contains one or more columns, e.g., \( \Gamma_i = \{\gamma_1, \ldots, \gamma_{N_i}\} \), where \( \gamma_1 \) represents the first column of the similar column group \( \Gamma_i \) and \( N_i \) indicates the number of columns in \( \Gamma_i \).

The columns in a similar column group are very similar with each other. We think of finding a single column from each similar column group, which represents the characteristics of all the columns in a group. The column is named as a condensed column. There are different ways to calculate the condensed column. Here the average of the columns in a similar column group is chosen as the condensed column, e.g., for the \( i \)th similar column group \( \Gamma_i \), its condensed column \( \gamma_C^i \) is defined as
\[
\gamma_C^i = \frac{1}{N_i} \sum_{j=1}^{N_i} \gamma_j^i. \tag{5}
\]

The similar sensing matrix is then built by combining the condensed columns from all the similar column groups, as \( \Psi = [\gamma_C^1, \ldots, \gamma_C^i, \ldots, \gamma_C^D] \), with size \( M \times D \).

**Remarks.** It should be noted that the original sensing matrix and the similar sensing matrix are connected via the condensed columns: each condensed column of the similar sensing matrix represents the characteristics of a similar column group in the original sensing matrix. The similar sensing matrix can be seen as a “compact” version of the original sensing matrix. Next, we show that the similar sensing matrix has low coherence provided that the similarity distance between its any two condensed columns is large.
Property 1. Let $T$ denote the minimum similarity distance between any two condensed columns of the similar sensing matrix $\Psi$, e.g., for any two condensed columns, e.g., $r_L^1$ and $r_L^c$, the similarity distance between their corresponding column points is larger than $T$.

\[
d_{\text{similar}}(P_{r_L^1}, P_{r_L^c}) \geq T.
\]  

(6)

Therefore, the similarity between any two condensed columns, $r_L^1$ and $r_L^c$, is no larger than $1 - T$:

\[
\lambda(r_L^1, r_L^c) \leq 1 - T.
\]  

(7)

Consequently, we have the following proposition for the similar sensing matrix.

Proposition 1. The coherence of the similar sensing matrix is less than or equal to $1 - T$.

Proof. The similar sensing matrix is built by combining the condensed columns from all the similar column groups, as $\Psi = [r_L^1, \ldots, r_L^c]$. According to Property 1, the similarity between any two condensed columns is no larger than $1 - T$. Therefore, the coherence of the similar sensing matrix, which is the largest similarity, is less than or equal to $1 - T$.

\[
\rho(\Psi) \leq 1 - T. \quad \Box
\]  

(8)

2.4. The general similar sensing matrix pursuit algorithm

The proposed GSSMP algorithm consists of six major steps which are described next.

Algorithm 1. General similar sensing matrix pursuit algorithm.

Input: Sensing matrix $\Phi$, measurement vector $y$.

Output: The estimated signal $x$.

1. Constructing the similar sensing matrix: The process is the same as that described in Section 2.3.  

2. Obtaining an initial estimate of the correct subspace: The OMP algorithm is used to find an estimate of the support set based on the measurement vector $y$ and the similar sensing matrix $\Psi$. The estimated support set is represented as $\hat{a} = [\hat{a}^1(r_L^1), \hat{a}^2(r_L^c), \ldots, \hat{a}^{K'}(r_L^c)]$, $K' \leq K$, where $\hat{a}^i(r_L^c)$ indicates that the first element of $\hat{a}$ corresponds to the $i$th condensed column $r_L^c$. We obtain an initial estimate of the correct subspace, $\hat{S}_\text{ini}$, spanned by the columns $r_L^1, r_L^c, \ldots, r_L^{c}$, defined as

\[
\hat{S}_\text{ini} = \text{SPAN}(r_L^1, r_L^c, \ldots, r_L^{c}).
\]

3. Finding the candidate columns in the original sensing matrix: Each column of the estimated subspace $\hat{S}_\text{ini}$ corresponds to a similar column group of the original sensing matrix, e.g., $r_L^c$ is the condensed column from the similar column group $G_1$. We can then obtain a set $\hat{\Lambda}$ containing the indices of $K'$ similar column groups corresponding to the columns in the estimated subspace $\hat{S}_\text{ini}$, as $\hat{\Lambda} = \{1, I, \ldots, I_l\}$. All the columns in each similar column group from $\hat{\Lambda}$ are listed and form a set of candidate columns $f$, where $f = [r_L^1, \ldots, r_L^{N_c}, \ldots, r_L^1, \ldots, r_L^{N_c}, \ldots, r_L^1, \ldots, r_L^{N_c}]$. We assume that the total number of the columns in $f$ is $H_{N_c}$.

4. Calculating the candidate subspaces: List $C_{H_{N_c}}$ combinations based on the columns in $f$. Each combination of $K$ columns spans a candidate subspace, e.g., the $p$th candidate subspace is represented as $\Psi^p = \text{SPAN}(r_L^1, r_L^c, \ldots, r_L^c)$. $p = 1, 2, \ldots, N_c$, where $N_c$ indicates the total number of candidate subspaces.

5. Calculating the candidate estimates: Based on each candidate subspace, the proposed algorithm solves a least squares problem to approximate the nonzero entries of the original $K$-sparse vector $x$ (9), and sets other entries as zero (10), resulting in a candidate estimate $\hat{x}^p, p = 1, 2, \ldots, N_c$ as

\[
\hat{x}^p_{[1,2,\ldots,N]} = 0,
\]

(10)

where $\Psi^p$ represents the pseudo-inverse operation, $\hat{x}^p_{[1,2,\ldots,N]}$ is composed of the entries of $\hat{x}^p$ indexed by $i \in \Psi^p$, and $\hat{x}^p_{[1,2,\ldots,N]} = 0$ is composed of the entries of $\hat{x}^p$ indexed by $i \in \{1, 2, \ldots, N\} - \Psi^p$ [31,29].  

6. Outputting the final estimates: For each candidate estimate $\hat{x}^p (p = 1, \ldots, N_c)$, calculate the residual $\hat{r}^p_{[1,2,\ldots,N]} = \hat{x}^p_{[1,2,\ldots,N]} - y$ via (11):

\[
\hat{r}^p_{[1,2,\ldots,N]} = y - \Phi \hat{x}^p.
\]

(11)

The $l_2$ norm of $\hat{r}^p$ is indicated as $\|\hat{r}^p\|_2$. Among the residuals, $\hat{r}^p_{[1,2,\ldots,N]}$, find the residual with the least $l_2$ norm, $\hat{r}_{\text{min}}$, and its associate candidate estimate, which is denoted as $\hat{x}_{\text{min}}$. Set $\hat{x}_{\text{min}}$ and its associate subspace $\hat{S}_{\text{min}}$ as the final estimates of the $K$-sparse signal $x$ and the correct subspace, respectively.

\[
\hat{x} = \hat{x}_{\text{min}},
\]

(12)

\[
\hat{S} = \hat{S}_{\text{min}}.
\]

(13)

3. Proof of the guaranteed reconstruction performance of the GSSMP algorithm

A sufficient condition for the perfect reconstruction of arbitrary $K$-sparse signal $x$ is stated in the following theorem.

Theorem 1. Let $x \in \mathbb{R}^N$ be a $K$-sparse signal, and let its corresponding measurement be $y = \Phi x + e \in \mathbb{R}^M$, where $e$ denotes the noise vector. Let $\Psi$ be a similar sensing matrix built based on the original sensing matrix $\Phi$. Suppose that the minimum similarity distance $T$ between any two condensed columns of the similar sensing matrix $\Psi$ satisfies

\[
T \geq 1 - 1/(4K' - 1),
\]

(14)

where $K'$ ($K' \leq K$) is the sparsity level of the intermediate vector $\hat{x}$ defined in Proposition 2. Then the GSSMP algorithm can identify the correct subspace, and reconstruct the $K$-sparse vector $x$ perfectly with reconstruction error that satisfies

\[
\|x - \hat{x}\|_2 \leq \frac{\|e\|_2}{\sqrt{1 - \mu(\Phi)(K' - 1)}}
\]

(15)

We now commence the proof of Theorem 1. First, we prove that the measurement vector $y$ can be represented using the condensed columns of the similar sensing matrix $\Psi$. At the same time, we can obtain an initial estimate of the correct subspace.
Proposition 2. The measurement vector $y$ can be represented using the condensed columns of the similar sensing matrix, as $y = \Phi x' + \epsilon$, where $x'$ is a $K'$-sparse vector ($K' < K$), and $\epsilon$ is an equivalent measurement noise vector. At the same time, we can obtain the initial estimate of the correct subspace $S_{\text{ini}}$. i.e., $S_{\text{ini}} = \text{SPAN}(\gamma_1^c, ..., \gamma_{C-1}^c, \gamma_C^c)$, where $\gamma_i^c$ is defined as the condensed column of the similar column group $\Gamma_i, i = 1, ..., K'$.

Proof. Let $(\tau_1, \tau_2, ..., \tau_K)$ be the $K$ randomly distributed nonzero elements of the original $K$-sparse vector $x$. The vectors $(b_1, b_2, ..., b_K)$ are defined as the columns of the sensing matrix $\Phi$, which correspond to the nonzero elements $(\tau_1, \tau_2, ..., \tau_K)$. The correct subspace $S$ is spanned by the vectors $(b_1, b_2, ..., b_K)$ as $S = \text{SPAN}(b_1, b_2, ..., b_K)$.

We consider two cases: (a) the vectors $(b_1, b_2, ..., b_K)$ are from $K$ different similar column groups of the original sensing matrix; (b) the vectors $(b_1, b_2, ..., b_K)$ are from $K'$ ($K' < K$) similar column groups.

For case (a), $b_i$ is from $\Gamma_i, i = 1, ..., K$. The vectors $(\gamma_1^c, \gamma_2^c, ..., \gamma_{C-1}^c)$ are the condensed columns of the similar column groups $\{\Gamma_1, \Gamma_2, ..., \Gamma_K\}$, respectively. The difference between $b_i$ and $\gamma_i^c$ is defined as $\Delta b_i = b_i - \gamma_i^c, i = 1, ..., K$. According to the definitions of the vectors $(b_1, b_2, ..., b_K)$, we have

$$y = \Phi x + \epsilon = r_1 b_1^c + r_2 b_2^c + \cdots + r_K b_K^c + \epsilon = r_1 (\gamma_1^c + \Delta b_1) + r_2 (\gamma_2^c + \Delta b_2) + \cdots + r_K (\gamma_K^c + \Delta b_K) + \epsilon = r_1 \gamma_1^c + r_2 \gamma_2^c + \cdots + r_K \gamma_K^c + r_1 (\Delta b_1) + r_2 (\Delta b_2) + \cdots + r_K (\Delta b_K) + \epsilon = r_1 \gamma_1^c + r_2 \gamma_2^c + \cdots + r_K \gamma_K^c + \epsilon', \quad (16)$$

where $\epsilon'$ is the equivalent measurement noise, $\epsilon' = r_1 (\Delta b_1) + r_2 (\Delta b_2) + \cdots + r_K (\Delta b_K) + \epsilon$. Considering that $(\gamma_1^c, \gamma_2^c, ..., \gamma_{C-1}^c)$ are from the similar sensing matrix $\Psi$, (16) naturally results in

$$y = \Psi x' + \epsilon', \quad (17)$$

where $x'$ is a $D \times 1$ sparse vector with $K'$ ($K' = K - 1$) nonzero elements $r_1, r_2, ..., r_{K'}$. We obtain the initial estimate of the correct subspace $S_{\text{ini}} = \text{SPAN}(\gamma_1^c, \gamma_2^c, ..., \gamma_{K-1}^c)$.

For case (b), without loss of generality, we assume that $K' = K - 1$, and that the two vectors $b_1$ and $b_2$ are both from the same group $\Gamma_1$ with condensed column $\gamma^c_1$. The difference between $b_1$ and $\gamma_1^c$ is defined as $\Delta b_1 = b_1 - \gamma_1^c$, and the difference between $b_2$ and $\gamma_2^c$ is defined as $\Delta b_2 = b_2 - \gamma_2^c$. Moreover, $b_1$ is from $\Gamma_{i-1}, i = 3, ..., K$. We have

$$y = \Phi x + \epsilon = r_1 b_1^c + r_2 b_2^c + \cdots + r_K b_K^c + \epsilon = r_1 (\gamma_1^c + \Delta b_1) + r_2 (\gamma_2^c + \Delta b_2) + r_3 (\gamma_3^c + \Delta b_3) + \cdots + r_K (\gamma_K^c + \Delta b_K) + \epsilon = r_1 \gamma_1^c + r_2 \gamma_2^c + \cdots + r_K \gamma_K^c + r_1 (\Delta b_1) + r_2 (\Delta b_2) + \cdots + r_K (\Delta b_K) + \epsilon = r_1 \gamma_1^c + r_2 \gamma_2^c + \cdots + r_K \gamma_K^c + \epsilon' = \Psi x' + \epsilon'. \quad (18)$$

where $x'$ is a $D \times 1$ sparse vector with $K'$ ($K' = K - 1$) nonzero elements $r_1, r_2, ..., r_{K'}$. Further, we have $r_i = r_1 + r_2$ and $r_i = r_{i+1}, i = 2, ..., K'$. The vector $\epsilon'$ is the equivalent measurement noise, which is defined as $\epsilon' = r_1 (\Delta b_1) + r_2 (\Delta b_2) + \cdots + r_K (\Delta b_K) + \epsilon$. We obtain the initial estimate of the correct subspace $S_{\text{ini}} = \text{SPAN}(\gamma_1^c, \gamma_2^c, ..., \gamma_{K-1}^c)$. This completes the proof. □

Next, we show that the $K'$-sparse vector $x'$ can be reconstructed perfectly based on the similar sensing matrix $\Psi$ and the measurement vector $y$ provided that the minimum similarity distance between any two condensed columns of the similar sensing matrix is large, and the amplitude of the equivalent measurement noise $\epsilon'$ is small enough.

Proposition 3. We have $y = \Psi x' + \epsilon' \in \mathbb{R}^M$, where $x' \in \mathbb{R}^D$ is a $K'$-sparse signal, $\epsilon'$ denotes the equivalent noise vector as $\epsilon' = r_1 (\Delta b_1) + r_2 (\Delta b_2) + \cdots + r_K (\Delta b_K) + \epsilon$, according to Proposition 2. Suppose that the minimum similarity distance $T$ of the similar sensing matrix $\Psi$ satisfies

$$T \geq 1 - 1/(4K')$$

and the amplitude of the equivalent noise $\epsilon'$ satisfies

$$\|\epsilon'\|_2 \leq A(1 - \mu(\Psi)(2K' - 1))/2,$$

where $A$ is a positive lower bound on the magnitude of the nonzero entries of $x'$, and $\mu(\Psi)$ denotes the coherence of the similar sensing matrix $\Psi$. The $K'$-sparse vector $x'$ can be reconstructed using the OMP algorithm, with reconstruction error

$$\|x' - \hat{x}'\|_2 \leq \frac{\|\epsilon'\|_2}{\sqrt{1 - \mu(\Psi)(2K' - 1)}},$$

Proof. According to Proposition 1, the coherence of the similar sensing matrix is less than or equal to $1 - T$:

$$\mu(\Psi) \leq 1 - T \quad (22)$$

Moreover, according to Theorem 3.1 in [37], in order to reconstruct the $K'$-sparse vector $x'$ perfectly using the OMP algorithm, the coherence $\mu(\Psi)$ should satisfy

$$\mu(\Psi) \leq 1/(4K' - 1).$$

The combination of (22) and (23) results in (24),

$$T \geq 1 - 1/(4K'),$$

which is a sufficient condition for the perfect reconstruction of the $K'$-sparse vector $x'$. Moreover, the reconstruction error satisfies [37, Theorem 3.1]

$$\|x' - \hat{x}'\|_2 \leq \frac{\|\epsilon'\|_2}{\sqrt{1 - \mu(\Psi)(2K' - 1)}},$$

This completes the proof. □

Though we can reconstruct $x'$ perfectly, we cannot obtain the exact value of the original signal $x$. There is no direct transform from $x$ ($D \times 1$ vector) to $x$ ($N \times 1$ vector). However, we can still identify the correct subspace based on the initial estimate of the correct subspace $S_{\text{ini}}$ provided $x'$ is perfectly reconstructed.

Proposition 4. Based on the initial estimate of the correct subspace $S_{\text{ini}}$, we can identify the correct subspace $S$. 

Proof. According to Proposition 3, the $K^\prime$-sparse signal $\mathbf{x}$ can be reconstructed perfectly with its $K^\prime$ nonzero elements and their associated vectors $[\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_{K^\prime}]$, which span the initial estimated subspace $\mathbf{S}_{\text{ini}} = \text{SPAN}(\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_{K^\prime})$. The vectors $[\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_{K^\prime}]$ correspond to a set of similar column groups in the original sensing matrix $(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, ..., \mathbf{\Gamma}_{K^\prime})$. According to Proposition 2, the vectors spanning the correct subspace, $(\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_K)$, are contained in the similar column groups $(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, ..., \mathbf{\Gamma}_{K^\prime})$, which correspond to the $K^\prime$ condensed columns spanning the initial estimated subspace $\mathbf{S}_{\text{ini}}$. As a result, a combinatorial search among all the columns contained in the similar column groups $(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, ..., \mathbf{\Gamma}_{K^\prime})$ can identify the vectors spanning the correct subspace, $(\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_K)$. This completes the proof. □

Remark. The combinatorial search procedures to identify the vectors spanning the correct subspace are described in Appendix A.

Proof of Theorem 1. We now complete the proof of Theorem 1. According to Proposition 4, the correct subspace $\mathbf{S}$ can be accurately identified based on the initial estimate of the correct subspace $\mathbf{S}_{\text{ini}}$ provided $\mathbf{x}$ is perfectly reconstructed. We can obtain an estimate of the original $K$-sparse vector $\mathbf{x}$, by solving a least squares problem to approximate the nonzero entries of $\mathbf{x}$ (26), and set other entries as zero (27):

$$\mathbf{x} = (\mathbf{S})^\dagger \mathbf{y}. \quad (26)$$

$$\hat{\mathbf{x}}_{1,2, \ldots, N, s} = 0. \quad (27)$$

The reconstruction error satisfies [37, Theorem 3.1]

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \leq \frac{\|\mathbf{e}\|_2}{\sqrt{1 - \mu(\Phi)(K - 1)}}. \quad (28)$$

Discussion: The accurate identification of the correct subspace relies on the perfect reconstruction of the $K^\prime$-sparse vector $\mathbf{x}^\prime$. In order to reconstruct $\mathbf{x}^\prime$ perfectly, the similar sensing matrix should have low coherence and the amplitude of the equivalent noise vector $\mathbf{e}$ should be small enough.

According to Proposition 2, we have

$$\mathbf{e}^\prime = r_1(\mathbf{\Delta b}_1) + r_2(\mathbf{\Delta b}_2) + \cdots + r_K(\mathbf{\Delta b}_K) + \mathbf{e}. \quad (29)$$

Let $B$ be the maximum difference between any column $\mathbf{b}_j$ and its corresponding condensed column $\mathbf{\gamma}_j^C$ within a similar column group $\mathbf{\Gamma}_i$. We then have

$$\|\mathbf{e}\|_2 = \|r_1(\mathbf{\Delta b}_1) + r_2(\mathbf{\Delta b}_2) + \cdots + r_K(\mathbf{\Delta b}_K) + \mathbf{e}\|_2$$

$$\leq |r_1| \cdot \|\mathbf{\Delta b}_1\|_2 + |r_2| \cdot \|\mathbf{\Delta b}_2\|_2 + \cdots + |r_K| \cdot \|\mathbf{\Delta b}_K\|_2 + \|\mathbf{e}\|_2$$

$$\leq B(|r_1| + |r_2| + \cdots + |r_K|) + \|\mathbf{e}\|_2$$

$$\leq B(K) \cdot \frac{\|\mathbf{e}\|_2}{\sqrt{1 - \mu(\Phi)(K - 1)}}. \quad (30)$$

where $A$ is a positive lower bound on the magnitude of the nonzero entries of $\mathbf{x}^\prime$. Eq. (30) implies that a small $B$ results in a small amplitude of the equivalent noise $\|\mathbf{e}\|_2$. Therefore, a column $\mathbf{b}_j$ should be close to its corresponding condensed column $\mathbf{\gamma}_j^C$ to obtain a small value of $B$.

### 4. Complexity analysis

In this section we analyze the computational complexity of the proposed GSSMP algorithm, and further make a comparison to the SP and BP algorithms based on the computational complexity.

The proposed GSSMP algorithm consists of offline processing and online processing. The offline processing transforms the $M \times N$ original sensing matrix to a $M \times D$ ($D < N$) similar sensing matrix based on similarity analysis. The computational complexity of the offline processing mainly focuses on the computation of the similarity between any two columns of the original sensing matrix, which is of the order of $O((N - 1)N/2)$ in the refined estimation process. In the rough estimation process, an OMP algorithm is used to find a rough estimate of the true support set. The computational complexity of the rough estimation process is the same as that of the OMP algorithm and is of the order of $O(\mu(\Phi))$ according to [34]. In the refined estimation process, $C_{S\text{cc}}$ combinations are listed out as the candidate subspaces, where $H_{cc}$ indicates the total number of columns in the $K^\prime$ contributing similar column groups. Assuming $G_{\text{max}}$ as the size of the largest similar column group, i.e., the number of columns contained in the largest similar column group, we have $H_{cc} \leq K^\prime G_{\text{max}}$. Based on each candidate subspace, the nonzero entries of the estimated sparse vector are calculated using the least squares algorithm. The computational cost of this refined estimation process is of the order of $O(C_{S\text{cc}} \cdot O(\text{LS}))$, where $O(\text{LS})$ indicates the order of the computational complexity for the least squares algorithm and $O(\text{LS}) = O(KM)$ according to [34]. The computational complexity of the proposed GSSMP algorithm is summarized in Table 1. In general, the

<table>
<thead>
<tr>
<th>Table 1: Computational complexity of GSSMP algorithm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offline processing</td>
</tr>
<tr>
<td>$O(\frac{(N-1)N}{2}M)$</td>
</tr>
</tbody>
</table>
computational complexity grows exponentially with the sparsity level $K$.

The computational costs of the SP and BP algorithms are of the order of $O(MN)$ and $LP(M,N)$, respectively, according to [34]. $LP(M,N)$ indicates the cost of solving a linear program with $N$ variables and $M$ constraints, which is $O(M^2N^{1.5})$ for an interior-point algorithm. The computational complexity of the SP and BP algorithms is shown in Table 2.

The proposed GSSMP algorithm is compared with the SP and BP algorithms based on the computational complexity of the online processing considering that the offline processing is carried out before the measurements arrive and is performed only once during the whole process with fixed original sensing matrix. The computational complexity of the proposed algorithm relies mainly on $H_{cc}$. If the sensing matrix is highly coherent with a large $G_{max}$, $H_{cc}$ may be very large, resulting in high computational complexity. On the contrary, if the sensing matrix is less coherent with a small $G_{max}$, $H_{cc}$ will be small and the computational complexity of the proposed algorithm will be comparable to that of the SP algorithm and lower than that of the BP algorithm.

### 5. Coping with unknown sparsity level

In practice, the sparsity level $K$ is always unknown. In this section, the GSSMP algorithm is enhanced to cope with the unknown sparsity level $K$ by changing the original refined estimation process, i.e., the method of finding the columns spanning the correct subspace $S$. Assuming that the initial estimate of the correct subspace, $\hat{S}_{ini}$, is spanned by $K$ condensed columns $\gamma_1^c, \gamma_2^c, \ldots, \gamma_K^c$, i.e., $\hat{S}_{ini} = \text{SPAN}(\gamma_1^c, \gamma_2^c, \ldots, \gamma_K^c)$. Each condensed column of $\hat{S}_{ini}$ corresponds to a similar column group of the original sensing matrix, which is named as the contributing similar column group, e.g., $\gamma_C^c$ is the condensed column from the contributing similar column group $\Gamma_i, i = 1, \ldots, K$. Each contributing similar column group contains a set of columns from the original sensing matrix, e.g., $\Gamma_i = \{\gamma_1^c, \gamma_2^c, \ldots, \gamma_{N_i}^c\}$, where $N_i$ indicates the number of columns in $\Gamma_i$.

According to Proposition 2, the true vectors spanning the correct subspace $S$, i.e., $\{b_1, b_2, \ldots, b_K\}$, are contained in the $K$ contributing similar column groups $\{\Gamma_1, \Gamma_2, \ldots, \Gamma_K\}$. On the other hand, Proposition 2 also stipulates that each contributing similar column group contains at least one true vector that spans the correct subspace $S$. Therefore, we can test each individual contributing similar column group one by one to find the true vectors spanning the correct subspace $S$. This novel algorithm is named as the contributing similar column group test algorithm, and the detailed procedures are listed in Algorithm 2.

### Algorithm 2. Contributing similar column group test algorithm.

**Input:** Initial estimate of the correct subspace, $\hat{S}_{ini}$

**Output:** Final estimate of the correct subspace, $\hat{S}$

For the $i$th contributing similar column group $\Gamma_i$:

1. Calculating the candidate subspaces corresponding to $\Gamma_i$: List all the combinations including one to $W_{ri}$ columns of $\Gamma_i$ as $F_{ri} = \{\gamma_1^c, \ldots, \gamma_{N_{ri}}^c\}$. The $k$th combination of $F_{ri}$ is denoted as $f_k^i, k = 1, \ldots, W_{ri}$, where $W_{ri}$ represents the number of combinations in $F_{ri}$. In the initial estimate of the correct subspace, $\hat{S}_{ini}$, replace the condensed column $\gamma_C^c$ (corresponding to $\Gamma_i$) with a combination from $F_{ri}$. Then each combination from $F_{ri}$, e.g., $f_k^i$, together with the $K - 1$ fixed condensed columns from $\hat{S}_{ini}$ (except $\gamma_C^c$), forms a candidate subspace, e.g., the $k$th candidate subspace is represented as $Y^k = \text{SPAN}(f_1^i, f_2^i, \ldots, f_{W_{ri}}^i), k = 1, 2, \ldots, W_{ri}$.

2. Finding the true vectors in $\Gamma_i$: For each candidate subspace, $Y^k, k = 1, 2, \ldots, W_{ri}$, calculate the candidate estimate and then the residual, via Steps 5 and 6 of Algorithm 1 in Section 2.4, respectively. Among the $R_{ri}$ candidate subspaces, find the one with the least residual, and choose its associate combination (from $F_{ri}$), which is denoted as $a_{min}^i$, as the true vectors. It should be noted that the combination $a_{min}^i$ contains one or more than one columns.

3. Obtaining a final estimate of the correct subspace: Repeat Steps 1 and 2 for $K'$ contributing similar column groups $\Gamma_i, i = 1, \ldots, K'$, and obtain $K'$ combinations (true vectors), indicated as $\{a_{min}^1, a_{min}^2, \ldots, a_{min}^{K'}\}$. Finally, we can obtain the final estimate of the correct subspace by combining these true vectors from $K'$ contributing similar column groups, as $\hat{S} = \text{SPAN}(a_{min}^1, a_{min}^2, \ldots, a_{min}^{K'})$.

### Setting of the parameter $W_{ri}$:

The parameter $W_{ri}$ is the maximin number of columns contained in a combination from $\Gamma_i$. According to Theorem 2.13 in [38], the sparsity level $K$ is less than or equal to half of the measurement number $M$, i.e., $K \leq M/2$. Furthermore, since the maximin number of columns drawn from $\Gamma_i$ is less than or equal to the sparsity level $K$, i.e., $W_{ri} \leq K$, we have $W_{ri} \leq M/2$. Therefore, we can choose the value of $W_{ri}$ according to $W_{ri} = \min (M/2, N_{ri})$.

**Computational complexity analysis:** In Algorithm 2, the number of candidate subspaces from $K'$ contributing similar column groups is $N_{cs} = \sum_{i=1}^{K} R_{ri}$, where $R_{ri}$ is the number of the total combinations in $\Gamma_i$, and $R_{ri} = C_{N_{ri}} + C_{N_{ri}}^2 + \ldots + C_{N_{ri}}^{W_{ri}}$. Assuming $G_{max}$ as the size of the largest similar column group, i.e., the number of columns contained in the largest similar column group, we have

$$N_{cs} \leq \sum_{i=1}^{K} W_{ri} \cdot C_{G_{max}} \leq K' \cdot M/2 \cdot C_{G_{max}}^{M/2}.$$  \hspace{1cm} (31)

Based on each candidate subspace, the nonzero entries of the estimated sparse vector are calculated using the least squares algorithm. The computational cost of the proposed algorithm is of the order of $N_{cs} \cdot O(\text{LS})$, where $O(\text{LS})$...
indicates the order of the computational complexity for the least squares algorithm, and \( O(\mathcal{L}) = O(KM) \). In summary, the computational complexity of the proposed contributing similar column group test algorithm is of the order of \( C_{\text{GSSMP}} C_{\text{GSSMP}} \mathcal{O}(K’KM^2) \), which is comparable to that of the original refined estimation process in the GSSMP algorithm \( (C_{\text{S}}^2 C_{\text{S}}^2 \mathcal{O}(KMKM)) \), considering that \( K’, K \) and \( M \) are much less than the length of the \( K \)-sparse vector, \( N \).

**Comparison to the original refined estimation process:** Under the condition of known sparsity level \( K \), the original refined estimation process selects all the combinations including \( K \) columns from the condensed columns of \( K’ \) contributing similar column groups at the same time, which is a global-optimal process. In Section 3, it is proved that the original refined estimation process achieves the minimum reconstruction error. However, when the sparsity level \( K \) is unknown, the original refined estimation process fails to estimate the \( K \)-sparse vector. In contrast, the proposed contributing similar column group test algorithm selects the true vectors from a contributing similar column group one time while keeping the other \( K’ – 1 \) condensed columns from \( \hat{S}_\text{ref} \) unchanged, which is a suboptimal process. It succeeds in estimating the \( K \)-sparse vector under the condition of unknown sparsity level \( K \). The simulation results from Section 6.2.4 show that the proposed contributing similar column group test algorithm can estimate the sparse vector representing the radar scene with an unknown number of targets successfully.

### 6. Simulation results and analysis

In this section, we present two simulation examples to verify the performance of the proposed GSSMP algorithm. The first example (Section 6.1) demonstrates that the proposed GSSMP algorithm outperforms the BP and SP algorithms when coping with deterministic sensing matrix with high coherence. The DOA estimation problem using compressed sensing for sensor array processing is addressed in Section 6.2. The proposed GSSMP algorithm can reconstruct the sparse DOA angle plane with high resolution and identify multiple targets more accurately, compared with the SCS, SBL, MUSIC and FOCUSS algorithms.

#### 6.1. A simple example

In this example, the proposed GSSMP algorithm is compared with the SP and BP algorithms in reconstructing \( K \)-sparse signals. In the experiments, the BP algorithm uses the default settings [39] (BP tools are from SparseLab [40]), and the SP algorithm uses the parameters given in [29].

The chirp matrix is chosen as the sensing matrix, which is deterministic in nature and composed of a family of chirps (i.e., frequency modulated sinusoids) \([41,42]\). A chirp signal \( \tau \) of length \( p \) with chirp rate \( r \) and base frequency \( m \) has the form

\[
t_{mp + r}(x) = \omega^m \omega^{mx + rx^2}, \quad x = 0, 1, \ldots, p - 1,
\]

where \( \omega \) is a primitive \( p \)th root of unity.

It is well known that when the length of the chirp signal is small, the chirp matrix will have high coherence. In this simulation example, the length of the chirp signal is set to 10, resulting in a \( 10 \times 100 \) chirp sensing matrix. A support set of size \( K (K \leq 5) \) is selected uniformly at random, and the original sparse vector is chosen as either Gaussian signal or zero–one signal [29].

#### 6.1.1. Construction of the similar sensing matrix

In the proposed algorithm, the columns of the original sensing matrix \( \Phi \) (100 columns) are mapped to a similarity space to obtain a number of column points (100 points). A hierarchical clustering method is then used to cluster the column points, resulting in 16 clusters with the minimum similarity distance set to 0.35. The columns from the same cluster constitute a similar column group. Consequently, all the columns of the sensing matrix (100 columns) are divided into 16 \( (D) \) similar column groups. A condensed column is generated from each similar column group, and a similar sensing matrix is then built based on the condensed columns. As can be seen, the size of the similar sensing matrix is \( 10 \times 16 \), far less than that of the original sensing matrix \( (10 \times 100) \), which reduces the computational complexity efficiently.

The histogram of the similarity values of the original sensing matrix and the similar sensing matrix is shown in Figs. 2 and 3, respectively. There is a shift towards the origin of the histogram from Fig. 2 (the original sensing

![Fig. 2.](image-url) Histogram of the similarity values (absolute off-diagonal entries of \( G \)) of the original sensing matrix.

![Fig. 3.](image-url) Histogram of the similarity values (absolute off-diagonal entries of \( G \)) of the similar sensing matrix.
matrix) to Fig. 3 (the similar sensing matrix). The tail representing the higher values in Fig. 2 disappears in Fig. 3. Therefore the coherence of the similar sensing matrix is far less than that of the original sensing matrix.

6.1.2. Reconstruction accuracy comparison

We perform five hundred Monte Carlo simulations for each fixed value of $K$ ($K \leq 5$). The reconstruction is considered to be exact when the $l_2$ norm of the difference between the original signal $x$ and the reconstructed signal $\hat{x}$ is smaller than a preset reconstruction accuracy level $\eta$, that is $\|x - \hat{x}\|_2 < \eta$. Here the reconstruction accuracy level $\eta$ is set to $10^{-2}$ and $10^{-3}$ respectively. Fig. 4 presents the reconstruction results for binary zero–one signal, at the reconstruction accuracy level of $10^{-2}$ and $10^{-3}$ respectively, and Fig. 5 shows the reconstruction results for Gaussian signal. Both figures (Figs. 4 and 5) show that the reconstruction performance of the proposed GSSMP algorithm is superior than that of the SP and BP algorithms, at different reconstruction accuracy levels.

6.1.3. Complexity comparison

The proposed GSSMP algorithm is compared with the SP and BP algorithms on computational complexity. The chirp sensing matrix consists of 100 columns, which are divided into 16 similar column groups. The maximum number of columns contained in a similar column group is 12 ($G_{\text{max}}$). The parameters are presented in Table 3 with its last column showing the value of $C_{H_{cc}}^K$, which is proportional to the computational complexity of the proposed GSSMP algorithm. Table 4 shows complexity comparison of the three algorithms. It can be seen from Table 4 that the computational complexity of the proposed algorithm is comparable to that of the SP algorithm and lower than BP algorithm.

6.2. Compressed sensing based DOA estimation

Recent studies have shown that the compressed sensing based DOA estimation technique always achieves better target resolution than classical methods, such as MUSIC and FOCUSS [22]. However, for most compressed sensing based DOA estimation methods, increasing the resolution of the sensor array system leads to finer gridding, which increases the correlation between the basis elements of the sensing matrix. Therefore, the sensing matrix has high coherence, which cannot guarantee a perfect reconstruction. Finer gridding also gives rise to higher computational complexity and numerical instability.

In this section, we propose the GSSMP algorithm to tackle the high coherence problem as well as high computational complexity due to finer gridding. The simulation results show that the proposed method achieves better angular resolution than both the compressed sensing based methods (SCS [22] and SBL [26]) and the classical DOA estimation methods (MUSIC [35] and FOCUSS [36]). Moreover, the proposed GSSMP algorithm is modified to cope with the unknown sparsity level problem, and the simulation results from Section 6.2.4 show that the modified GSSMP algorithm with the contributing similar column group test process can estimate the sparse vector representing the radar scene with unknown number of targets successfully.

6.2.1. Problem formulation

We consider a sensor array system, which performs spatial sampling at the Nyquist rates. It is composed of

![Fig. 4. Frequency of exact reconstruction for the zero–one signal.](image)

![Fig. 5. Frequency of exact reconstruction for the Gaussian signal.](image)

### Table 3

<table>
<thead>
<tr>
<th>Parameter setting of the GSSMP algorithm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
</tr>
<tr>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Example 1</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>Computational complexity comparison.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm</td>
</tr>
<tr>
<td>GSSMP</td>
</tr>
<tr>
<td>SP</td>
</tr>
<tr>
<td>BP</td>
</tr>
</tbody>
</table>
Since the DOA angle plane is generally sparse in practice, compressed sensing is a valid candidate for estimating the DOA angles for multiple targets. To do so, the DOA angle plane is divided into $N_{grid}$ fine grids, each grid generally with the same size $\Delta \theta$. The $i$th grid represents a DOA angle of $\theta_0 + (i-1)\Delta \theta$, where $\theta_0$ is the initial angle of the DOA plane. As the system has no prior knowledge of the locations of the targets, the information of all the grids in the DOA plane should be considered. Therefore, the steering matrix $\Phi_{\text{steering}}$ is defined as $\Phi_{\text{steering}} = [u(\theta_0) \ u(\theta_0 + \Delta \theta) \ \cdots \ u(\theta_0 + (i-1)\Delta \theta) \ \cdots \ u(\theta_0 + (N_{grid}-1)\Delta \theta)]$, where $u(\cdot)$ is a $M_{\text{antenna}} \times 1$ steering vector, e.g.,

$$u(\theta_0) = [1 \ 2\pi d \sin(\theta_0/\lambda) \ \cdots \ \rho(M_{\text{antenna}} - 1) 2\pi d \sin(\theta_0/\lambda)]^T.$$  

In (33), $d$ is the distance between the elements of the arrays and $\lambda$ denotes wavelength. Since a small number of grids are occupied by the targets, the reflection vector $x$ is an $N_{grid} \times 1$ sparse vector with the $i$th element defined as $x(i) = r(\theta_0 + (i-1)\Delta \theta)$ if the $i$th grid is occupied by a target; otherwise, $x(i) = 0$. $r(\theta_0 + (i-1)\Delta \theta)$ is the reflection coefficient of the hypothetical target at a DOA angle $\theta_0 + (i-1)\Delta \theta$. As a result, the $M_{\text{antenna}} \times 1$ received complex vector of array observations $y$ could be written as

$$y = \Phi_{\text{steering}} x + e.$$  

where $e$ is an $M_{\text{antenna}} \times 1$ complex Gaussian noise vector. Though the vectors and matrices in (34) are complex valued in contrast to the original compressed sensing environment, it is easy to transfer it to real variables according to [5], where the radar case is mapped into the standard compressed sensing framework by breaking the problem into its real and imaginary parts.

6.2.2. Comparison to other DOA estimation algorithms

The DOA estimation performance of the proposed GSSMP algorithm is compared with that of the following methods: SCS, SBL, MUSIC and single-snapshot FOCUSS. A single-snapshot scenario was considered for all tested methods (except MUSIC). MUSIC was assumed to have a known number of uncorrelated far-field sources with 200 available snapshots at the array output. The SCS and SBL use the parameters given in [22,26], respectively.

In the simulation setup, the number of transmit/receive antennas ($M_{\text{antenna}}$) is set as 20. The antennas transmit independent orthogonal quadrature phase shift keyed (QPSK) waveforms and the carrier frequency is 8.62 GHz. The range of the DOA plane is $[0^\circ, 90^\circ]$, which is divided into $90 \times (N_{grid})$ grids with the initial angle ($\theta_0$) and angle interval ($\Delta \theta$) equaling $0^\circ$ and $1^\circ$, respectively. The deterministic steering matrix $\Phi_{\text{steering}}$ is chosen as the original sensing matrix. The addressed scenarios are characterized by the signal-to-noise ratio (SNR), where the power of the noise accounts for the measurement noise.
6.2.3. Angular resolution analysis

First, the achievable resolution of the GSSMP algorithm is evaluated and compared with the SCS, SBL, MUSIC and FOCUSS algorithms in a two-target scenario. Considering an SNR of 10 dB, Figs. 6–9 show four scenarios with angular separation of 24°, 6°, 1°, and 0.5° between two targets, respectively. Figs. 6 and 7 show that the three compressed sensing based algorithms (GSSMP, SCS and SBL) outperform the MUSIC and FOCUSS algorithms, and succeed in clearly separating the two targets with different separation angles (24° and 6°). Fig. 8 shows that in scenario with very closely spaced targets (1°), both the GSSMP and SBL algorithms succeed in identifying the two targets accurately, while the other three algorithms (SCS, MUSIC and FOCUSS) cannot separate the two targets. Furthermore, it can be seen from Fig. 9 that only the GSSMP algorithm can distinguish the two targets with angular separation of 0.5°. The simulation results verify that both the GSSMP and SBL algorithm are capable of dealing with the sensing matrix with high coherence due to finer gridding, and the GSSMP algorithm achieves higher resolution than the SBL algorithm.

Next, a scenario where the spatial spectrum is induced by seven targets of different power levels with SNR=[10, 5, 10, 5, 10, 10, 10] dB is analyzed. The seven targets are at DOA angles of 33°, 34°, 50°, 61°, 62°, 77° and 85°, respectively. Note that the main challenge in this scenario stems from both variable power levels of seven targets, and the small separation between two targets at DOA angle of 33°, 34° and 61°, 62°, respectively. Fig. 10 shows that SCS, SBL, MUSIC, and FOCUSS algorithms are unsuccessful in estimating all DOAs correctly. The SBL algorithm loses track of two targets at the DOA angle of 77° and 85°,
respectively. In contrast, the GSSMP algorithm accurately estimates the DOA angles of all simulated targets. The GSSMP algorithm also provides accurate estimate of the power of the targets.

In order to further verify the performance of the proposed GSSMP algorithm, five random cases are tested. In the first case, three targets are randomly distributed at the DOA angle range of \([0^\circ, 90^\circ]\) with randomly chosen amplitudes. And similar tests are carried out for four to seven targets in the remaining four cases. The simulation results are shown in Figs. 11–15, which show that the GSSMP algorithm can estimate both the DOA angles and the power of the targets accurately in all five cases. However, the SBL algorithm fails in estimating two closely spaced targets in the fourth case of tracking six randomly distributed targets (Fig. 14). Furthermore, five hundred Monte Carlo simulations are performed for each case to calculate the average root mean square error (RMSE) of the DOA angle. The readers can refer to [44] for the calculation of RMSE. The average RMSEs of the DOA angles from Cases 1 to 5 are listed in Table 5, which shows that the proposed GSSMP algorithm achieves the lowest average RMSE among the five algorithms (GSSMP, SCS, MUSIC, FOCUSS and SBL), for each specific case.

6.2.4. Coping with the unknown sparsity level (unknown number of targets)

In order to verify the performance of the proposed contributing similar column group test algorithm, five hundred Monte Carlo simulations are performed with randomly chosen number of targets (less than eight targets), which have random locations and amplitudes. The average RMSEs of the DOA angles are listed in Table 6, which shows that the GSSMP algorithm with the contributing similar column group test process (indicated as modified GSSMP in Table 6) achieves much lower average RMSE in DOA angle than other five algorithms (the original GSSMP, SCS, MUSIC, FOCUSS and SBL).

7. Conclusions

In this paper, we formulated the GSSMP algorithm to reconstruct a K-sparse vector based on a deterministic sensing matrix with high coherence. We also proved that under appropriate conditions the GSSMP algorithm can identify the correct subspace quite well, and reconstruct the original K-sparse signal perfectly. Moreover, we enhanced the GSSMP algorithm to cope with the unknown sparsity level problem, by changing the method of finding the columns spanning the correct subspace \(S\). Though suboptimal in nature, the modified GSSMP algorithm succeeded in estimating the sparse vector representing the radar scene with an unknown number of targets successfully.

Acknowledgment

This work was supported by the Natural Science Foundations of China (Nos. 61104051 and 61473217), the CAST (China Academy of Space Technology) Foundation, National 973 project of China through Grant 2013CB329405, the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (61221063), and the Fundamental Research Funds for the Central Universities.

Appendix A. Combinatorial search procedures

Each column of the estimated subspace \(\hat{S}_{\text{ini}}\) corresponds to a similar column group of the original sensing matrix, e.g., \(\gamma_l^*\) is the condensed column from the similar column group \(\Gamma_l\). We can then obtain a set \(\Lambda\) containing the indices of \(K\) similar column groups corresponding to the columns in the estimated subspace \(\hat{S}_{\text{ini}}\) as \(\Lambda = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_l\}\). All the columns in each similar column group from \(\Lambda\) are listed and form a set of candidate columns \(\hat{f}\), where \(\hat{f} = \{f_1, f_2, \ldots, f_N\}_{\hat{f}}\). We assume that the total number of the columns in \(\hat{f}\) is \(H_{\text{cc}}\). The combinatorial search procedures to identify the vectors spanning the correct subspace are listed in the following.

1. List \(C_{\alpha}^K\) combinations based on the columns in \(\hat{f}\). Each combination of \(K\) columns spans a candidate subspace, e.g., the \(p\)th candidate subspace is represented as \(\mathbf{Y}^p = \text{SPAN}(y_1^p, y_2^p, \ldots, y_{\alpha}^p), p = 1, 2, \ldots, N_{\alpha}\), indicates the total number of candidate subspaces.
2. Based on each candidate subspace, the proposed algorithm solves a least squares problem to approximate the nonzero entries of the original K-sparse vector \(x\)

---

### Table 5
Average RMSE of DOA angles (degree).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSSMP</td>
<td>0.001</td>
<td>0.007</td>
<td>0.012</td>
<td>0.023</td>
<td>0.078</td>
</tr>
<tr>
<td>SCS</td>
<td>0.1</td>
<td>0.15</td>
<td>0.26</td>
<td>0.31</td>
<td>0.34</td>
</tr>
<tr>
<td>MUSIC</td>
<td>0.27</td>
<td>0.57</td>
<td>1.23</td>
<td>1.75</td>
<td>1.9</td>
</tr>
<tr>
<td>FOCUSS</td>
<td>0.15</td>
<td>0.19</td>
<td>0.24</td>
<td>0.45</td>
<td>0.5</td>
</tr>
<tr>
<td>SBL</td>
<td>0.03</td>
<td>0.07</td>
<td>0.08</td>
<td>0.15</td>
<td>0.18</td>
</tr>
</tbody>
</table>

### Table 6
Average RMSE of DOA angles (degree) with unknown sparsity level.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Modified GSSMP</th>
<th>Original GSSMP</th>
<th>SCS</th>
<th>MUSIC</th>
<th>FOCUSS</th>
<th>SBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average RMSE of DOA</td>
<td>0.05</td>
<td>0.45</td>
<td>0.48</td>
<td>1.76</td>
<td>0.53</td>
<td>0.23</td>
</tr>
</tbody>
</table>
\(x^p_{\cdot1_N - r}\) where \(\vdagger\) indicates the pseudo-inverse operation, \(\hat{x}^p_{\cdot}\) is composed of the entries of \(x^p\) indexed by \(i \in \mathbb{Y}^p\), and \(\hat{x}^p_{\cdot1_N - r}\) is composed of the entries of \(x^p\) indexed by \(i \in \{1, 2, \ldots, N\} - \mathbb{Y}^p\) \cite{31, 29}.

3. For each candidate estimate \(\hat{x}^p (p = 1, \ldots, N_{sc})\) calculate the residual \(r^p (p = 1, \ldots, N_{sc})\) via

\[r^p = y - \Phi \hat{x}^p\]  

(3.3)

The \(l_2\) norm of \(r^p\) is indicated as ||\(r^p||_2\). Among the residuals, \((r^p, p = 1, \ldots, N_{sc})\), find the residual with the least \(l_2\) norm, \(r_{\min}\), and its associate candidate estimate, which is denoted as \(\hat{x}_{\min}\). Set \(\hat{x}_{\min}\) and its associate subspace \(\mathbb{Y}_{\min}\) as the final estimates of the \(K\)-sparse signal \(x\) and the correct subspace, respectively.

\[\hat{x} = \hat{x}_{\min}\]  

(4.4)

\[S = \mathbb{Y}_{\min}\]  

(4.5)

References


