Block Sparse Vector Recovery via Weighted Generalized Range Space Property

Ahmed Al Hilli and Athina Petropulu

Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08854
E-mail: ahmed.alhilli@rutgers.edu, athinap@rutgers.edu

***Abstract*— In block sparse vector recovery problems we are interested in finding the vector with the least number of active blocks that best describes the observation. The convex relaxation of that problem, typically used to reduce complexity, is strictly equivalent with the original problem only when certain conditions are met, such as Restricted Isometry Property, Null Space Characterization, and Block Mutual Coherence. In practice, those conditions may not be satisfied, which implies that solving the relaxed problem may not retrieve the block sparsest solution. In this paper, we propose a weighted approach, which, in the noise free case and under certain conditions guarantees that the relaxed problem solution has the same support as the sparsest block vector. The weights can be obtained based on a low resolution estimate of the group sparse signal.**

I. Introduction

Compressed Sensing (CS) and Sparse Signal Recovery emerge in many signal processing applications, including biomedical imaging [1], [5], [13], [15], [19], and radar [3], [4], [11], [14], [20]. In sparse signal recovery, we are interested in finding the best possible representation for the observation vector using a vector with the smallest number of non-zero entries. Mathematically, this can be represented as

$$\text{(PL0)} \qquad \begin{array}{cc} \min & \|\mathbf{x}\|_0 \\ \text{Subjet to:} & \mathbf{y} = \mathbf{A}\mathbf{x} \end{array} \tag{1}$$

where the ℓ_0-norm, $\|.\|_0$, represents the number of non-zero entries in a vector, $\mathbf{x} \in R^{n\times 1}$ is the minimization variable, $\mathbf{A} \in R^{r\times n}$ is the dictionary matrix, and $\mathbf{y} \in R^{r\times 1}$ is the observation vector. It has been shown that (1) is an NP-hard problem [16]. To tackle the complexity associated with the ℓ_0-norm problem, a relaxed convex ℓ_1-norm approximation is often used to find a sparse solution. The mathematical model of the relaxed problem can be written as

$$\text{(PL1)} \qquad \begin{array}{cc} \min & \|\mathbf{x}\|_1 \\ \text{Subject to:} & \mathbf{y} = \mathbf{A}\mathbf{x} \end{array} \tag{2}$$

The problems (PL0) and (PL1) are said to be *strictly equivalent* if they both have a unique solution and the two solutions coincide [21]. When the problem (PL0) has multiple solutions, and the solution of (PL1) coincides with one of the solutions of (PL0), then we say that (PL0) and (PL1) are *equivalent* [21]. (PL0) and (PL1) are strictly equivalent when certain conditions are met, such as the Restricted Isometry Property (RIP) [6], the Null Space Property (NSP) [7], the Mutual Coherence [8], or the Range Space Property (RSP) of order K [21]. In practice however, those conditions may not be met, which means that the least ℓ_1-norm solution and the least ℓ_0-norm solution are not the same. In [1], a weighted approach is proposed for recovering the support of the sparsest solution in cases in which the dictionary matrix exhibits high coherence. Also, in [12], the optimal choice of the weights for the weighted approach, such that minimum amount of measurements is needed for exact recovery using the location of the support of the signal is discussed.

In some applications, it is known in advance that the non-zero entries of the underlying sparse vector occur in groups, a properties known as block sparsity. In block sparse signal recovery problems, we are interested in finding the vector with the least number of non-zero blocks that explains the observed vector. If we let m represent the number of groups in $\mathbf{x}$, the block sparsest vector estimation problem can be written as [10]

$$\text{(PG0)} \qquad \begin{array}{cc} \min & \sum_{i=1}^{m} I(\|\mathbf{x}_i\|_2) \\ \text{Subject to} & \mathbf{y} = \mathbf{A}\mathbf{x} \end{array} \tag{3}$$

where $I = \begin{cases} 1 & \text{if } \|\mathbf{x}_i\|_2 > 0 \\ 0 & \text{if } \|\mathbf{x}_i\|_2 = 0 \end{cases}$, $\mathbf{x}_i$ represents the i^{th} block of the vector $\mathbf{x}$.

(PG0) is hard to solve, so its convex relaxation is often considered, which consists of finding a vector with the smallest sum of the blocks ℓ_2-norm , i.e.,

$$\text{(PG1)} \qquad \begin{array}{cc} \min & \sum_{i=1}^{m} \|\mathbf{x}_i\|_2 \\ \text{Subject to} & \mathbf{y} = \mathbf{A}\mathbf{x} \end{array} \tag{4}$$

In general, (3) and (4) are not the same. Several works have provided conditions for strict equivalence between (PG0) and (PG1). Those include the generalization of the RIP condition [10], the Null Space Characterization of [18] and the generalization of the Mutual Coherence [9]. In [2], a generalized RSP (GRSP) is proposed for group sparse under-determined systems, where a set of sufficient and necessary conditions for a sparse vector $\mathbf{x}$ to be a solution to the problem of (4) are proposed.

In this paper, we propose a weighted approach to address the cases in which strict equivalence conditions may not be satisfied. We show that by multiplying the sensing matrix with a diagonal matrix $\mathbf{W}$, we transform the problem into a problem that satisfies the GRSP, and provide the conditions so that

Work supported in part by NSF under grant NSF ECCS 1408437.

ISBN 978-0-9928626-7-1 © EURASIP 2017

the transformed problem has the same support as the sparsest block vector. By multiplying by the weighting matrix $\mathbf{W}$, the range space of $\mathbf{A}$ is rotated in such as way that solving (PG1) favors the underlying block sparse vector.

The paper is organized as follow. In Section II, we discuss the background theory that related to the proposed approach. Section III introduced the proposed approach in noise free and noisy observation cases, while Section IV provides conclusion remarks.

II. Background Theory

The GRSP was proposed in [2]. The conditions for equivalence between (PG0) and (PG1) are stated in the following theorem.

Theorem 1 *[2] Let $\boldsymbol{x}_s$ represent a sparsest solution to (PG0). $\boldsymbol{x}_s$ is also a unique solution to problem (PG1) if and only if there is a vector $\boldsymbol{u}^* \in R(\boldsymbol{A}^T)$ such that*

$$\begin{cases} \|\boldsymbol{u}_i^*\|_2 = \|\frac{\boldsymbol{x}_{si}}{\|\boldsymbol{x}_{si}\|_2}\|_2 = 1 & \text{if } \|\boldsymbol{x}_{si}\|_2 > 0 \\ \|\boldsymbol{u}_i^*\|_2 < 1 & \text{if } \|\boldsymbol{x}_{si}\|_2 = 0 \end{cases} \quad (5)$$

where $R(.)$ represents the range sapce of a matrix. A sufficient condition for (6) to have a unique solution was also provided in [2].

III. The Proposed Approach

A. Noise Free Group Sparse Vectors

By substituting $t_i = \|\mathbf{x}_i\|_2$, (PG1) can be recast as Second Order Cone Program (SOCP), i.e., [10]

$$\text{(PG2)} \quad \begin{array}{cc} \min & \sum_{i=1}^m t_i \\ \text{Subject to} & t_i \geq \|\mathbf{x}_i\|_2 \\ & t_i \geq 0 \\ & \mathbf{y} = \mathbf{A}\mathbf{x} \end{array} \quad (6)$$

One can see that the problems in (4) and (6) are equivalent in the sense that the optimal solution for both problems is the same. Suppose that the *strict equivalence* conditions for (PG0) and (PG1) do not hold. Below, we show that by multiplying the sensing matrix with a diagonal matrix $\mathbf{W}$, we transform the problem into a problem that satisfies the conditions in Theorem 1. We provide a sufficient condition for the weighted problem to satisfy the GRSP. In the following, S represents the support of $\mathbf{x}_s$ (i.e., the indices of the active groups), $\bar{S} = \{1, 2, ..., m\} \mid S$ the complement of S.

Theorem 2 *If for the block sparsest solution $\mathbf{x}_s$ it holds that $\|\mathbf{A}_{\bar{S}i}^T(\mathbf{A}_S^T)^\dagger \mathbf{u}_S\|_2 < 1$, for all $i \in \bar{S}$, then $\mathbf{x}_s$ is the solution to the problem (4), where $\mathbf{A}_{\bar{S}i}$, for $i \in \bar{S}$, is the collection of columns in $\mathbf{A}$ associated with non-active blocks in $\mathbf{x}_s$, $\mathbf{A}_S$ is the concatenation of the blocks in $\mathbf{A}$ that are associated with active blocks in $\mathbf{x}_s$, $\mathbf{u}$ is a vector with $\mathbf{u}_{Si} = \frac{\mathbf{x}_{si}}{\|\mathbf{x}_{si}\|_2}$ for $i \in S$, and $\|\mathbf{u}_{\bar{S}i}\|_2 < 1$ for $i \in \bar{S}$.*

Proof: We will prove the contrapositive of this theorem, i.e., we will show that if the GRSP conditions of Theorem 1 are not satisfied, then $\|\mathbf{A}_{\bar{S}i}^T(\mathbf{A}_S^T)^\dagger \mathbf{u}_S\|_2 \geq 1$ for some $i \in \bar{S}$. Let $\mathbf{A}^T\mathbf{v} = \mathbf{u}$, and suppose that the first condition of (5) is satisfied, but the second condition is not satisfied. Then we have

$$\mathbf{u}_{Si} = \mathbf{A}_{Si}^T \mathbf{v} \quad , \forall i \in S \quad (7)$$

$$\exists i \in \bar{S} \text{ such that } \|\mathbf{A}_{\bar{S}i}^T \mathbf{v}\|_2 \geq 1 \quad (8)$$

From (7), we have $\mathbf{u}_S = \mathbf{A}_S^T\mathbf{v}$. On solving for $\mathbf{v}$, and substituting the solution in (8), we get

$$\|\mathbf{A}_{\bar{S}i}^T(\mathbf{A}_S^T)^\dagger \mathbf{u}_S\|_2 \geq 1 \quad \text{for some } i \in \bar{S} \quad (9)$$

■

Revisiting the original problem in (3), we have

$$\begin{array}{cc} \min & \sum_{i=1}^m I(\|\mathbf{x}_i\|_2) \\ \text{Subject to} & \mathbf{y} = \mathbf{A}\mathbf{x} \end{array} \quad (10)$$

Let $\mathbf{x} = \mathbf{W}\mathbf{q}$, where $\mathbf{W}$ is a diagonal weight matrix with the following structure, $\mathbf{W} = \mathbf{w} \otimes \mathbf{I}_{k\times k}$, where $\mathbf{w}$ is a diagonal matrix that contains the weights, $\otimes$ represents the Kronecker product, $\mathbf{I}_{k\times k}$ is the identity matrix with size $k \times k$, with k representing the group size. Based on its structure, W assigned the same weight to all elements of a block. Then, (10) can be rewritten as

$$\begin{array}{cc} \min & \Sigma_{i=1}^m I(\|\mathbf{w}_i\mathbf{q}_i\|_2) \\ \text{subject to} & \mathbf{y} = \mathbf{A}\mathbf{W}\mathbf{q} \end{array} \quad (11)$$

Since $\mathbf{W}$ is non-zero at the support of $\mathbf{x}$, (11) can be rewritten as

$$\begin{array}{cc} \min & \Sigma_{i=1}^m I(\|\boldsymbol{q}_i\|_2) \\ \text{subject to} & \mathbf{y} = \mathbf{A}\mathbf{W}\mathbf{q} \end{array} \quad (12)$$

The problem of (12) is NP-hard. Its convex relaxation can be solved instead by replacing the indication function with the sum of the active groups energies of $\mathbf{q}$, i.e.,

$$\begin{array}{cc} \min & \Sigma_{i=1}^m \|\boldsymbol{q}_i\|_2 \\ \text{subject to} & \mathbf{y} = \mathbf{A}\mathbf{W}\mathbf{q} \end{array} \quad (13)$$

Suppose that the block sparsest solution of (PG1) does not satisfy the GRSP conditions in Theorem 1. We can manipulate $\mathbf{W}$ such that the solution to (13) satisfies the condition on a vector that has the same support of the sparsest block sparse solution of (PG1).

To make the support of the solution of (12) coincide with that of the sparsest vector $\mathbf{x}$, we have to choose $\mathbf{W}$ such that

$$\|\mathbf{W}_{\bar{S}i}\mathbf{A}_{\bar{S}i}^T(\mathbf{A}_S^T)^\dagger \mathbf{W}_S\mathbf{u}_S\|_2 < 1 \quad \forall i \notin S \quad (14)$$

where $\mathbf{W}_{\bar{S}i}$ are the diagonal elements in $\mathbf{W}$ that associated with $i \in \bar{S}$, and $\mathbf{W}_S$ is a diagonal matrix composed of diagonal sub-matrices $\mathbf{W}_{Si}$ for $i \in S$. It is easy to show that by assigning high values to the $\mathbf{w}_i$ that corresponds to active groups, and low values to the $\mathbf{w}_i$ that corresponds to non-active groups, (14) is satisfied, and the support of the solution of (13) is the same as the support of sparsest solution.

B. The Noisy Case

In the case of noisy observations, we choose to minimize the tradeoff between the sparsity of the solution and the fitting error. The minimization problem can be written as

$$\min \quad h\Sigma_{i=1}^{m}\|\boldsymbol{x}_i\|_2 + \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2 \tag{15}$$

By setting $t_i = \|\mathbf{x}_i\|_2, \quad \forall i \in \{1,2,...,m\}$, and $\mathbf{v} = \mathbf{y} - \mathbf{A}\mathbf{x}$, we can write (15) as

$$\begin{array}{rl} \min & h\Sigma_{i=1}^{m} t_i + \|\mathbf{v}\|_\mathbf{2} \\ \text{subject to} & \mathbf{v} = \mathbf{y} - \mathbf{A}\mathbf{x} \\ & t_i \geq 0 \qquad \forall i \in \{1,2,...,m\} \\ & t_i \geq \|\mathbf{x}_i\|_2 \quad \forall i \in \{1,2,...,m\} \end{array} \tag{16}$$

The dual of (16) is

$$\begin{array}{rl} \max & \boldsymbol{\alpha}^T y \\ \text{subject to} & \|\boldsymbol{\alpha}\|_2 \leq 1 \\ & \|\mathbf{A_i^T}\boldsymbol{\alpha}\|_2 \leq \lambda_{2i} \\ & \boldsymbol{\lambda_1} + \boldsymbol{\lambda_2} = \mathbf{1}h \\ & \boldsymbol{\lambda_1} \geq \mathbf{0}, \boldsymbol{\lambda_2} \geq \mathbf{0} \end{array} \tag{17}$$

Now, we will provide the conditions for the sparsest block sparse vector to be the solution to (15). Those conditions are stated in the following theorem. In the following, we will assume that the system in (15) has a unique solution, and Slater's and strict complementary slackness conditions are satisfied.

Theorem 3 *$\mathbf{x}^*$ is a solution to the system: $h\Sigma_{i=1}^{m}\|\mathbf{x}_i\|_2 + \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2$ if and only if there is a $\mathbf{u}^* \in R(\mathbf{A}^T)$ such that*

$$\begin{array}{ll} \mathbf{u}^* = \mathbf{A}^T\boldsymbol{\alpha}^* = \mathbf{A}^T\frac{\mathbf{y}-\mathbf{A}\mathbf{x}^*}{\|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_2}, & \|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_\mathbf{2} > 0, \\ \|\mathbf{A}_i^T\boldsymbol{\alpha}^*\|_2 = h & \text{iff } \|\mathbf{x}_i\|_2 > 0 \\ \|\mathbf{A}_i^T\boldsymbol{\alpha}^*\|_2 < h & \text{iff } \|\mathbf{x}_i\|_2 = 0 \end{array} \tag{18}$$

Proof: First, we will prove the necessary condition, i.e., if $\mathbf{x}^*$ is a solution to (15), then there is $\mathbf{u} \in R(\mathbf{A}^T)$ such that

$$\begin{array}{l} \|\mathbf{A}_i^T\frac{\mathbf{y}-\mathbf{A}\mathbf{x}^*}{\|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_2}\|_2 = h \text{ when } \|\mathbf{x}_i\|_2 > 0 \\ \|\mathbf{A}_i^T\frac{\mathbf{y}-\mathbf{A}\mathbf{x}^*}{\|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_2}\|_2 < h \text{ when } \|\mathbf{x}_i\|_2 = 0 \end{array} \tag{19}$$

Now, consider the non-zero entries in (15). We have

$$h\mathbf{u^T}\mathbf{x} + \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_\mathbf{2} \tag{20}$$

Differentiating (20) with respect to $\mathbf{x}$, and equating to zero, we get

$$h\mathbf{u}_i = \mathbf{A}_i^T\frac{\mathbf{y}-\mathbf{A}\mathbf{x}^*}{\|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_2} \tag{21}$$

which implies that

$$\|\mathbf{A}_i^T\frac{\mathbf{y}-\mathbf{A}\mathbf{x}^*}{\|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_2}\|_2 = h \quad \forall i \in S \tag{22}$$

For zero blocks in $\mathbf{x}^*$, let $\mathbf{t}^*$, $\mathbf{v}^*$, and $\mathbf{x}^*$ be the solution of (16), and $\boldsymbol{\alpha}^*$, $\boldsymbol{\lambda}_1^*$ and $\boldsymbol{\lambda}_2^*$ be the solution for (17). From the strict complementary property, we should have $t_i^* + \lambda_{1i}^* > 0$, which implies $\lambda_{1i}^* > 0$ when $t_i^* = \|\mathbf{x}_i^*\|_2 = 0$. From the second and third constraints of (17), we have

$$\|\mathbf{A}_i^T\boldsymbol{\alpha}^*\|_2 < h \quad \forall i \in \bar{S} \tag{23}$$

From Slater's condition, we have

$$h\Sigma_{i=1}^{m}t_i^* + \|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_\mathbf{2} = \boldsymbol{\alpha^{*T}}(\mathbf{A}\mathbf{x}^* + \mathbf{v}^*) \tag{24}$$

We can see that $\boldsymbol{\alpha^*} = \frac{\mathbf{y}-\mathbf{A}\mathbf{x}^*}{\|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_\mathbf{2}}$ is a solution to (24), since it satisfies (24) and the constraints in (17). So, indeed there is a vector $\mathbf{u}^* \in R(\mathbf{A^T})$ such that it satisfies the conditions in (19). Now, we will provide the proof for the sufficient condition, i.e., if there is a $\mathbf{u}^*$ that satisfies the conditions in (19), then $\mathbf{x}^*$ is the solution to (15). Assume that $\hat{\mathbf{x}} \neq \mathbf{x}^*$ is the solution to (15), then we should have

$$h\Sigma_{i=1}^{m}\|\hat{\mathbf{x}}_i\|_2 + \|\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}\|_\mathbf{2} < h\Sigma_{i=1}^{m}\|\mathbf{x}_i^*\|_2 + \|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_\mathbf{2} \tag{25}$$

From the necessary condition of this theorem, if $\hat{\mathbf{x}}$ is a solution to (15), then there should be $\hat{\boldsymbol{\alpha}} = \frac{\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}}{\|\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}\|_2}$ such that

$$\begin{array}{l} \|\mathbf{A}_i^T\hat{\boldsymbol{\alpha}}\|_2 = h \text{ when } \|\hat{\mathbf{x}}_i\|_2 > 0 \\ \|\mathbf{A}_i^T\hat{\boldsymbol{\alpha}}\|_2 < h \text{ when } \|\hat{\mathbf{x}}_i\|_2 = 0 \end{array} \tag{26}$$

Assume that $\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\lambda_1}}$, and $\hat{\boldsymbol{\lambda_2}}$ are the dual solution set. The dual problem should attend its maximum at $\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\lambda_1}}$, and $\hat{\boldsymbol{\lambda_2}}$, i.e.,

$$\hat{\boldsymbol{\alpha}}^\mathbf{T}\mathbf{y} > \boldsymbol{\alpha^{*T}}\mathbf{y} \tag{27}$$

$$\hat{\boldsymbol{\alpha}}^\mathbf{T}\mathbf{A}\hat{\mathbf{x}} + \hat{\boldsymbol{\alpha}}^\mathbf{T}\hat{\mathbf{v}} > \boldsymbol{\alpha^{*T}}\mathbf{A}\mathbf{x}^* + \boldsymbol{\alpha^{*T}}\mathbf{v}^* \tag{28}$$

$$h\Sigma_{i=1}^{m}\|\hat{\mathbf{x}}_i\|_2 + \|\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}\|_\mathbf{2} > h\Sigma_{i=1}^{m}\|\mathbf{x}_i^*\|_2 + \|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_\mathbf{2} \tag{29}$$

which contradict the first assumption, i.e., $h\Sigma_{i=1}^{m}\|\hat{\mathbf{x}}_i\|_2 + \|\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}\|_\mathbf{2} > h\Sigma_{i=1}^{m}\|\mathbf{x}_i^*\|_2 + \|\mathbf{y}-\mathbf{A}\mathbf{x}^*\|_\mathbf{2}$ ■

For the weighted problem, the problem can be restated as

$$h\Sigma_{i=1}^{m}\|\mathbf{q}_i\|_2 + \|\mathbf{y}-\mathbf{A}\mathbf{W}\mathbf{q}\|_2, \tag{30}$$

where $\mathbf{W} = \mathbf{w} \otimes \mathbf{I}_{k \times x}$. The conditions of Theorem 7 can be easily modified to include the weights as follows

$$\begin{cases} \|w_i\mathbf{A}_i^T\boldsymbol{\alpha}\|_2 = h\|\frac{\mathbf{q}_i}{\|\mathbf{q}_i\|_2}\|_2 = h & \|\mathbf{q}_i\|_2 > 0 \\ \|w_i\mathbf{A}_i^T\boldsymbol{\alpha}\|_2 < h & \|\mathbf{q}_i\|_2 = 0 \end{cases} \tag{31}$$

where w_i is the i^{th} element of $\mathbf{w}$, which represents the weight associated to i^{th} block

Now, we will provide a theorem in which, if w_i is less that a specific value, the corresponding group will be non-active.

Theorem 4 *For the problem in (30), if $w_i < \frac{h}{\sqrt[2]{\lambda_{max}(\mathbf{A}_i\mathbf{A}_i^T)}}$, then q_i=0, where $\lambda_{max}(\mathbf{G})$ is the largest eigenvalue of the matrix $\mathbf{G}$.*

Proof: According to the second condition in (31), $\|w_i\mathbf{A}_i^T\boldsymbol{\alpha}\|_2 < h$ when $\mathbf{q}_i$=0, which can be rewritten as

$$\boldsymbol{\alpha}^T(\mathbf{A}_i\mathbf{A}_i^T)\boldsymbol{\alpha} < \frac{h^2}{w_i^2} \tag{32}$$

We have $\boldsymbol{\alpha}^T(\mathbf{A}_i\mathbf{A}_i^T)\boldsymbol{\alpha} \in [\lambda_{max}, \lambda_{min}]$, and the maximum achievable value is λ_{max}. On substituting λ_{max} in the above equation, we get

$$w_i < \frac{h}{\sqrt[2]{\lambda_{max}(\mathbf{A_i A_i^T})}}. \tag{33}$$

■

We can see form Theorem 4 that assigning low values to the weights that correspond to non-active blocks guarantees that these blocks will be non-active in the estimated vector. For instance, if we assign low values to w_i that are associated to the non-active blocks, such that (33) is satisfied for all non-active blocks, and assign high values to w_i that are associated with active blocks, such that (33) is not satisfied, solving (15) will retrieve a vector with the same support as the underlying sparse vector. Since we do not know the real support of the underlying block sparse vector, we propose to use a low resolution estimate to reconstruct the weighting matrix.

IV. Simulation Results

To evaluate the performance of the proposed approach, we test our approach on Synthetic Aparture Radar (SAR). To simulate the block sparsity scenario, we consider the case in which the target is composed of two adjacent pixels in the scene. We adopt the system that was used in [3] to simulate the sensing matrix $\mathbf{A}$, and following the assumption that the reflectivity of the targets does not depend on the observation angle, the system can be described as a linear system of the form $\mathbf{y} = \mathbf{A}\mathbf{x}$. The simulation parameters that were used to construct the sensing matrix $\mathbf{A}$ are as shown in Table (1).

Table 1: Parameters Used in The Simulation

Center Freq.	200MHz
Pulse Width	2.5×10^{-7}
Bandwidth	5 MHz
$\Delta\theta$	1^o

The ground patch used in the simulation is 40 m wide and 60 m long, and the scene is uniformly sampled on a grid with spacing 0.5 m. The distance between the antenna and the ground patch center is 1050 m. The diagonal weighting matrix $\mathbf{W}$ that is used in the proposed approach is constructed by assigning to the diagonal of $\mathbf{W}$ a low resolutio estimate based on Spatial Frequency Interpolation [17]. Fig. 1-(a) shows two targets inside the scene of interest, while Fig. 1-(b) shows the spatial frequency interpolation. It is clear from 1-(b) that the frequency spatial interpolation provides a rough estimate of the targets locations. Fig. 1-(c) shows the estimated targets using the estimation in Fig. 1-(b) as a weighting matrix in the proposed approach, while Fig. 1-(d) shows the estimation of the non-weighted approach. One can see from Fig. 1-(c) and 1-(d) that the non-weighted approach fails to estimate the actual targets, while the weighted approach estimates the targets locations correctly.

Next, we conduct Monte Carlo simulations to test the performance of the weighed approach as compared to the non-weighted approach for noise free and additive white Gaussian noise. 100 Monte Carlo trials are performed. In each trial, n block sources are randomly distributed around the scene, and a low resolution estimate ($\mathbf{w}$) is constructed based on the estimation result of spatial frequency interpolation to be used in the weighing approach. The performance metric is the success rate; we claim success when the indices of the n largest blocks of the estimated source coincide with the actual group indices of the actual vector.

Fig. 2 shows the performance of the proposed approach versus the non-weighted approach with the increase of the number of active blocks in the actual source, and block size of 3. One can see that the non-weighted approach degrades rapidly with the increase of the number of active blocks, while the proposed approach shows significantly better performance. Fig. 3 shows the performance of the proposed approach and that of the non-weighted approach for different SNRs, and block size of 2. One can see that the non-weighted approach performance is poor even at high SNR, while the proposed approach shows good performance for SNR above 5 dB.

V. Conclusions

In this paper, a weighted approach has been proposed to solve for the block sparsest vector in scenarios when the strict equivalence conditions may not hold. Simulation results have shown improved performance as compared to the non-weighted approach.

References

[1] A. Al Hilli, L. Najafizadeh, and A. Petropulu, "EEG sparse source localization via range space rotation," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2015 IEEE 6th International Workshop on*. IEEE, 2015, pp. 265–268.

[2] ——, "Generalized range space property for group sparsity of linear underdetermined systems," in *2016 Annual Conference on Information Science and Systems (CISS)*. IEEE, 2016, pp. 568–571.

[3] ——, "Sparse target scene reconstruction for SAR using range space rotation," in *2016 IEEE Radar Conference (RadarConf)*. IEEE, 2016, pp. 1–5.

[4] M. T. Alonso, P. López-Dekker, and J. J. Mallorquí, "A novel strategy for radar imaging based on compressive sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 12, pp. 4285–4295, 2010.

[5] S. Aviyente, "Compressed sensing framework for eeg compression," in *Proc. IEEE/SP 14th Workshop Stat. Signal Process*, 2007, pp. 181–184.

[6] E. J. Candes and T. Tao, "Decoding by linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215, 2005.

[7] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k-term approximation," *Journal of the American mathematical society*, vol. 22, no. 1, pp. 211–231, 2009.

[8] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via 1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.

[9] Y. C. Eldar, P. Kuppinger, and H. Bölcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3042–3054, 2010.

[10] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *Information Theory, IEEE Transactions on*, vol. 55, no. 11, pp. 5302–5316, 2009.

[11] J. H. Ender, "On compressive sensing applied to radar," *Signal Processing*, vol. 90, no. 5, pp. 1402–1414, 2010.

[12] A. Flinth, "Optimal choice of weights for sparse recovery with prior information," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 4276–4284, 2016.

[13] J. P. Haldar, D. Hernando, and Z.-P. Liang, "Compressed-sensing mri with random encoding," *IEEE transactions on Medical Imaging*, vol. 30, no. 4, pp. 893–903, 2011.

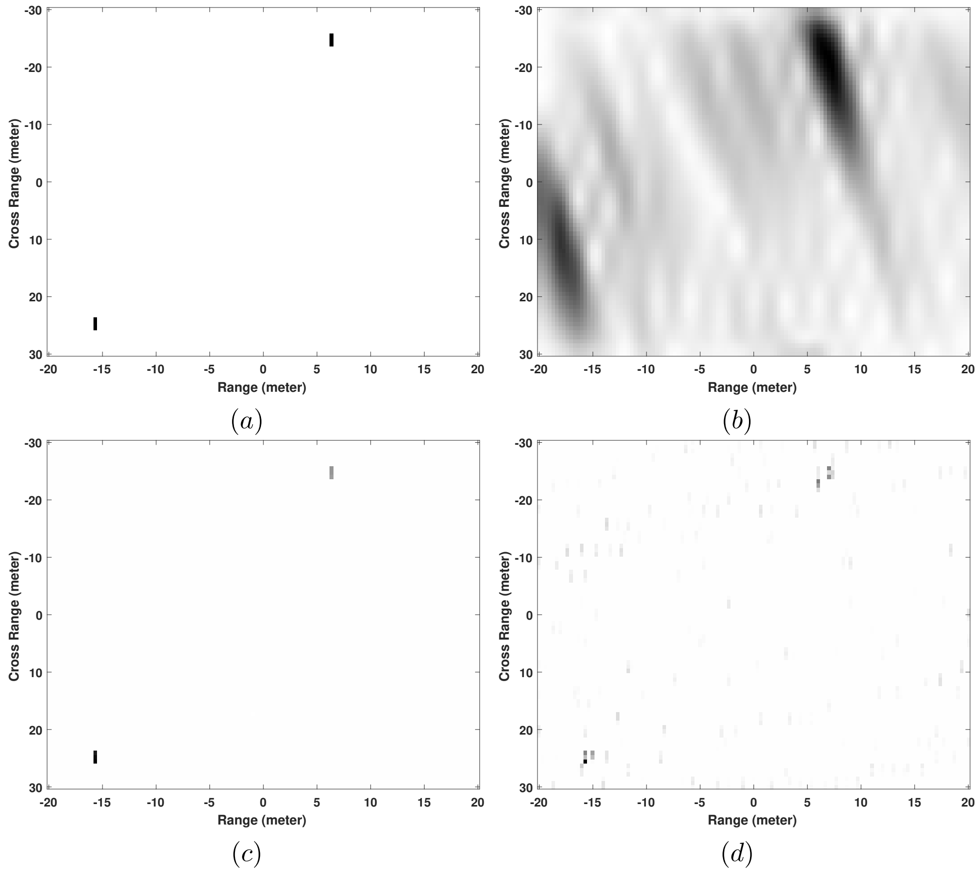

Fig. 1. a) The actual sourced for n=2, b) The estimation that is used to construct the weights, c) The estimated sources via the weighted approach, d) The estimated sources of the non-weighted approach.

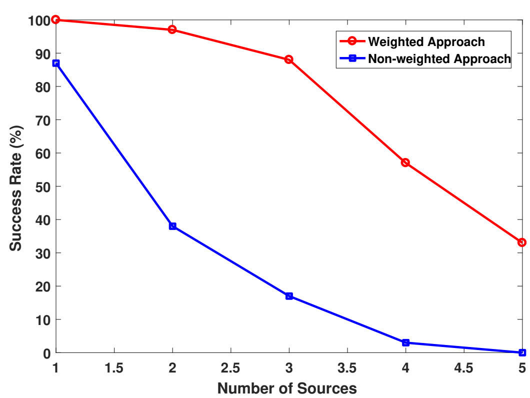

Fig. 2. Success rate of the proposed approach and the non-weighted approach with different number of active blocks sources, with block size of 3.

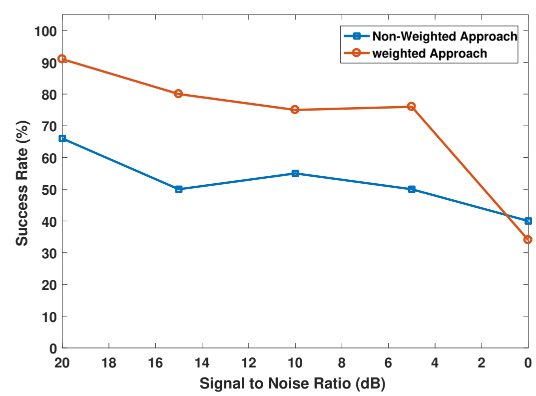

Fig. 3. Success rate of the proposed approach and the non-weighted approach with different Signal to Noise Ratios and two active blocks, with block size of 2.

[14] M. Herman and T. Strohmer, "Compressed sensing radar," in *2008 IEEE Radar Conference*. IEEE, 2008, pp. 1–6.

[15] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing mri," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, 2008.

[16] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM journal on computing*, vol. 24, no. 2, pp. 227–234, 1995.

[17] M. Soumekh, *Synthetic aperture radar signal processing*. New York: Wiley, 1999.

[18] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *Signal Processing, IEEE Transactions on*, vol. 57, no. 8, pp. 3075–3085, 2009.

[19] S. C. Wu and A. L. Swindlehurst, "Matching pursuit and source deflation for sparse EEG/MEG dipole moment estimation," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, pp. 2280–2288, 2013.

[20] Y. Yu, A. P. Petropulu, and H. V. Poor, "MIMO radar using compressive sampling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 1, pp. 146–163, 2010.

[21] Y.-B. Zhao, "Rsp-based analysis for sparsest and least-norm solutions to underdetermined linear systems," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5777–5788, 2013.