Hindawi
Mathematical Problems in Engineering
Volume 2017, Article ID 4903791, 7 pages
https://doi.org/10.1155/2017/4903791

Research Article

Support Recovery of Greedy Block Coordinate Descent Using the Near Orthogonality Property

Haifeng Li

College of Mathematics and Information Science, Henan Normal University, Xinxiang 453007, China

Correspondence should be addressed to Haifeng Li; lihaifengxx@126.com

Received 23 November 2016; Accepted 15 March 2017; Published 27 April 2017

Academic Editor: Bogdan Dumitrescu

In this paper, using the near orthogonal property, we analyze the performance of greedy block coordinate descent (GBCD) algorithm when both the measurements and the measurement matrix are perturbed by some errors. An improved sufficient condition is presented to guarantee that the support of the sparse matrix is recovered exactly. A counterexample is provided to show that GBCD fails. It improves the existing result. By experiments, we also point out that GBCD is robust under these perturbations.

1. Introduction

Greedy block coordinate descent (GBCD) algorithm was presented by [1] for direction of arrival (DOA) estimation. In the work of [1], the DOA estimation is treated as the multiple measurement vectors (MMV) model that recovers a common support shared by multiple unknown vectors from multiple measurements. The authors provided a sufficient condition, based on mutual coherence, to guarantee that GBCD exactly recover the nonzero supports with noiseless measurements.

Recently, the work of [2] discussed the following method:

$$\begin{aligned} \min_{\mathbf{X}} \quad & \|\mathbf{X}\|_{2,1} \\ \text{s.t.} \quad & \widehat{\mathbf{Y}} = \mathbf{AX} + \mathbf{N}, \\ & \widehat{\mathbf{A}} = \mathbf{A} + \mathbf{E}, \end{aligned} \tag{1}$$

with inputs $\widehat{\mathbf{Y}} \in \mathbb{R}^{m\times L}$ and $\widehat{\mathbf{A}} \in \mathbb{R}^{m\times N}$. $\mathbf{N}$ denotes the measurement noise and $\mathbf{E}$ denotes the system perturbation. The perturbations $\mathbf{E}$ and $\mathbf{N}$ are quantified with the following relative bounds:

$$\begin{aligned} \frac{\|\mathbf{N}\|_F}{\|\mathbf{Y}\|_F} &\le \varepsilon, \\ \frac{\|\mathbf{E}\|_2^{(K)}}{\|\mathbf{A}\|_2^{(K)}} &\le \varepsilon_A, \end{aligned} \tag{2}$$

where $\|\mathbf{A}\|_2^{(K)}$ and $\|\mathbf{Y}\|_F$ are nonzero. Here, $\|\mathbf{A}\|_2^{(K)}$ denotes the largest spectral norm taken over all K-column submatrices of $\mathbf{A}$. Throughout the paper, we are only interested in the case where ε and ε_A are far less than 1. In (1), $\mathbf{X}$ is a K-group sparse matrix; that is, it has no more than K nonzero rows, and $\|\mathbf{X}\|_{2,1} = \sum_{i=1}^{N} \|\mathbf{x}^i\|_2$, $\mathbf{x}^i$ is the ith row of $\mathbf{X}$. It is assumed that all columns of $\widehat{\mathbf{A}}$ are normalized to be of unit-norm [3]. Both $\mathbf{Y} = \mathbf{AX}$ and $\mathbf{A}$ are totally perturbed in (1). This case can be found in source separation [4], radar [5], remote sensing [6], and countless other problems. In addition, the total perturbations have also been discussed in [7–9].

One of the most commonly known conditions is the restricted isometry property (RIP). A matrix $\mathbf{A}$ satisfies RIP of the order K if there exists a constant $\delta \in (0, 1)$ such that

$$(1-\delta)\|\mathbf{h}\|_2^2 \le \|\mathbf{Ah}\|_2^2 \le (1+\delta)\|\mathbf{h}\|_2^2 \tag{3}$$

for all K-sparse vector $\mathbf{h}$. In particular, the minimum of all constants δ satisfying (3) is called the restricted isometry constant (RIC) δ_K.

There are many papers [8, 10–14] discussing the sufficient condition for orthogonal matching pursuit (OMP) that is one of the widely greedy algorithms for sparse recovery. In [3], using the near orthogonality property, the authors improved the sufficient condition of OMP. As cited in [3], the near orthogonality property can further develop the orthogonality characterization of columns in $\mathbf{A}$; it will play a fundamental role in the study of the signal reconstruction performance in

compressed sensing. In the noiseless case, the work of [15] analyzed the performance of GBCD using near orthogonality property and improved the results in [2].

In this paper, under the total perturbations, we use near orthogonality property to improve the theoretical guarantee for the GBCD algorithm. In [2], the authors stated that $\delta_{K+1} < 1/(\sqrt{K}+1)$ is a sufficient condition for GBCD. We improve this condition to $\delta_{K+1} < (\sqrt{4K+1}-1)/2K$. We also present a counterexample to show that GBCD fails. The example is superior to that in [2]. Under the total perturbations, the robustness of GBCD is shown by experiments.

Now we give some notations that will be used in this paper. $\mathbf{a}_i$ denotes the ith column of a matrix $\mathbf{A}$. $\mathbf{A}'$ denotes the transpose of $\mathbf{A}$. $\mathbf{I}_M$ denotes an $M \times M$ identity matrix. The symbol vec denotes the vectorization operator by stacking the columns of a matrix one underneath the other. The cardinality of a finite set Γ is denoted by $|\Gamma|$. Let $\Omega := \{1, 2, \ldots, N\}$. $\Gamma^c = \Omega \setminus \Gamma = \{i \mid i \in \Omega, \text{ and } i \notin \Gamma\}$. The support of $\mathbf{X}$ is denoted by $\operatorname{supp}(\mathbf{X})$ ($\operatorname{supp}(\mathbf{X}) = \{i \mid \mathbf{x}^i \neq \mathbf{0}\}$). $\|\mathbf{A}\|_2^{(K)}$ denotes the largest spectral norm taken over all K-column submatrices of $\mathbf{A}$. Let $\|\mathbf{A}\|_{2,\infty}$ denote the maximum ℓ_2 norm of the rows of $\mathbf{A}$. We write $\mathbf{A}_\Gamma$ for the column submatrix of $\mathbf{A}$ whose indices are listed in set of Γ and $\mathbf{X}^\Gamma$ for the row submatrix of $\mathbf{X}$ whose indices are listed in the set Γ. $\mathbf{e}_i \in \mathbb{R}^N$ denotes the ith unit standard vector.

2. Problem Formulation

Analogous to [1], (1) can be rewritten as

$$\min_{\mathbf{X}} \frac{1}{2} \left\| \widehat{\mathbf{Y}} - \widehat{\mathbf{A}}\mathbf{X} \right\|_F + \lambda \|\mathbf{X}\|_{2,1}. \tag{4}$$

Assume that $\Gamma := \operatorname{supp}(\mathbf{X})$. Obviously, $|\Gamma| = K$. The objective function in (4) can be written as

$$F(\mathbf{X}) = G(\mathbf{X}) + H(\mathbf{X}), \tag{5}$$

where $G(\mathbf{X}) = (1/2)\|\widehat{\mathbf{Y}} - \widehat{\mathbf{A}}\mathbf{X}\|_F^2 = (1/2)\|\operatorname{vec}(\widehat{\mathbf{Y}}) - \mathbf{I}_L \otimes \widehat{\mathbf{A}} \operatorname{vec}(\mathbf{X})\|_F^2$ with $\otimes$ denoting the Kronecker product and $H(\mathbf{X}) = \lambda\|\mathbf{X}\|_{2,1} = \lambda \sum_{i=1}^N \|\mathbf{x}^i\|_2$. Combining the quadratic approximation for $G(\mathbf{X})$ and standard BCD algorithm, the solution to the ith subproblem can be given by a soft-thresholding operator. The authors in [1] only update the block that yields the greatest descent distance. Now, we list GBCD algorithm (Algorithm 1).

Suppose that $\mathbf{A}$ satisfies the Kth order RIC $\delta_K \in (0,1)$. Recall that $\mathbf{X}$ has no more than K nonzero rows. According to the fact $\|\mathbf{X}\|_F^2 = \sum_{i=1}^L \|\mathbf{x}_i\|_2^2$, we can obtain

$$(1-\delta_K)\|\mathbf{X}\|_F^2 \le \|\mathbf{AX}\|_F^2 \le (1+\delta_K)\|\mathbf{X}\|_F^2 \tag{6}$$

from (3).

Combining Lemma 2.4 in [3] and (6), we have

$$\left\|\mathbf{A}_\Gamma'\mathbf{AX}\right\|_F^2 = \sum_{i=1}^L \left\|\mathbf{A}_\Gamma'\mathbf{A}\mathbf{x}_i\right\|_2^2 \tag{7}$$

$$\ge \sum_{i=1}^L \left((1-\delta_K)\|\mathbf{A}\mathbf{x}_i\|_2^2\right) \tag{8}$$

$$= (1-\delta_K)\sum_{i=1}^L \|\mathbf{A}\mathbf{x}_i\|_2^2 \tag{9}$$

$$= (1-\delta_K)\|\mathbf{AX}\|_F^2. \tag{10}$$

Lemma 1 (near orthogonality property, see [3]). *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be two orthogonal sparse vectors with supports* T_u *and* T_v *fulfilling* $|T_u \cup T_v| \le K$. *Suppose that* $\mathbf{A}$ *satisfies RIP of order* K *with RIC* δ_K. *Then we have*

$$|\cos \measuredangle(\mathbf{Au}, \mathbf{Av})| \le K, \tag{11}$$

where $\measuredangle(\mathbf{Au}, \mathbf{Av})$ *denotes the angle between* $\mathbf{Au}$ *and* $\mathbf{Av}$.

Lemma 2 (see [3]). *Under the same assumptions as in Lemma 1, we have*

$$|\langle \mathbf{Au}, \mathbf{Av}\rangle| \le \delta \|\mathbf{Au}\|_2 \|\mathbf{Av}\|_2. \tag{12}$$

Lemma 3. *For finite sets* $\overline{\Gamma}$ *and* $\widetilde{\Gamma}$, *let* $\operatorname{supp}(\overline{\mathbf{X}}) = \overline{\Gamma}$ *and* $\operatorname{supp}(\widetilde{\mathbf{X}}) = \widetilde{\Gamma}$. *Here,* $\overline{\Gamma} \cap \widetilde{\Gamma} = \emptyset$, *and* $|\overline{\Gamma} \cup \widetilde{\Gamma}| \le K$. *If* $\mathbf{A}$ *satisfies the RIP condition (3) with* $\delta_K \in (0,1)$, *then we have*

$$\left|\left\langle \mathbf{A}\overline{\mathbf{X}}, \mathbf{A}\widetilde{\mathbf{X}}\right\rangle_F\right| \le \delta_K \left\|\mathbf{A}\overline{\mathbf{X}}\right\|_F \left\|\mathbf{A}\widetilde{\mathbf{X}}\right\|_F. \tag{13}$$

Proof. Note that the Frobenius norm of $\mathbf{A}$ is derived from the Frobenius inner product.

$$\left|\left\langle \mathbf{A}\overline{\mathbf{X}}, \mathbf{A}\widetilde{\mathbf{X}}\right\rangle_F\right| = \left|\sum_{i=1}^L \left\langle \mathbf{A}\overline{\mathbf{x}}_i, \mathbf{A}\widetilde{\mathbf{x}}_i\right\rangle\right| \tag{14}$$

$$\le \delta_K \left(\sum_{i=1}^L \|\mathbf{A}\overline{\mathbf{x}}_i\|_2 \|\mathbf{A}\widetilde{\mathbf{x}}_i\|_2\right) \tag{15}$$

$$= \delta_K \sqrt{\left(\sum_{i=1}^L \|\mathbf{A}\overline{\mathbf{x}}_i\|_2 \|\mathbf{A}\widetilde{\mathbf{x}}_i\|_2\right)^2} \tag{16}$$

$$\le \delta_K \sqrt{\sum_{i=1}^L \|\mathbf{A}\overline{\mathbf{x}}_i\|_2^2 \sum_{i=1}^L \|\mathbf{A}\widetilde{\mathbf{x}}_i\|_2^2} \tag{17}$$

$$= \delta_K \left\|\mathbf{A}\overline{\mathbf{X}}\right\|_F \left\|\mathbf{A}\widetilde{\mathbf{X}}\right\|_F, \tag{18}$$

where (15) and (17) follow from Lemma 2 and Cauchy-Schwarz inequality, respectively. □

3. RIP Based Recovery Condition

In this section, we firstly present the upper bound of the noise matrix $-\mathbf{EX} + \mathbf{N}$ and provide the recovery condition for GBCD.

Lemma 4 (see [2]). *Suppose that* $\widehat{\mathbf{A}}$ *satisfies the* K*th order RIC* $\widehat{\delta}_K \in (0, 0.5)$. *Then we have*

$$\|-\mathbf{EX} + \mathbf{N}\|_F < \left(\frac{\varepsilon_A}{1/\sqrt{3} - \left(1 + 1/\sqrt{3}\right)\varepsilon_A} + \varepsilon\right) \frac{\left\|\widehat{\mathbf{Y}}\right\|_F}{1-\varepsilon}. \tag{19}$$

Input: $\widehat{\mathbf{A}}$, $\widehat{\mathbf{Y}}$, $\mathbf{X}(0) = \mathbf{0}$, $n = 1$, $\lambda > 0$, $\beta > 0$.
(1) **Repeat** until "stopping criterion" is met
(2) **for** $i = 1 : N$
(3) $\mathbf{p}^i(n-1) = \mathbf{x}^i(n-1) - \beta\widehat{\mathbf{a}}_i'(\widehat{\mathbf{A}}\mathbf{X}(n-1) - \widehat{\mathbf{Y}})$
(4) $\mathbf{x}^i(n) = (\mathbf{p}^i(n-1)/\|\mathbf{p}^i(n-1)\|_2)\max(0, \|\mathbf{p}^i(n-1)\|_2 - \lambda\beta)$
(5) $\mathrm{comp}(i) = \|\mathbf{x}^i(n) - \mathbf{x}^i(n-1)\|$
(6) **end for**
(7) Choose the index i_0 such that $\mathrm{comp}(i_0) = \max(\mathrm{comp})$
(8) $\mathbf{X}(n) \leftarrow [\mathbf{x}^1(n-1); \ldots; \mathbf{x}^{i_0-1}(n-1);$
(9) $\mathbf{x}^{i_0}(n); \mathbf{x}^{i_0+1}(n-1); \ldots; \mathbf{x}^N(n-1)]$
(10) $n \leftarrow n+1$
(11) **End Repeat**

Algorithm 1: GBCD: greedy block coordinate descent algorithm [1].

According to steps (7) and (8) of Algorithm 1, at the nth iteration, GBCD can obtain a correct index if

$$\max_{i\in\Gamma}\left\|\mathbf{X}^i(n) - \mathbf{X}^i(n-1)\right\|_2 > \max_{j\in\Gamma^c}\left\|\mathbf{X}^j(n) - \mathbf{X}^j(n-1)\right\|_2. \tag{20}$$

Theorem 5. *Consider model (4). Let $t_0 = \min_{i\in\Gamma}\|\mathbf{x}^i\|_2$. If the matrix $\widehat{\mathbf{A}}$ satisfies RIP of order $K+1$ with*

$$\widehat{\delta}_{K+1} < \frac{\sqrt{4K+1}-1}{2K}, \tag{21}$$

$$t_0 > \frac{\left(\sqrt{1+\widehat{\delta}_{K+1}} + \sqrt{K}\right)\varepsilon_0}{1-\widehat{\delta}_{K+1} - \sqrt{1-\widehat{\delta}_{K+1}}\sqrt{K}\widehat{\delta}_{K+1}}, \tag{22}$$

where $\varepsilon_0 = (\varepsilon_A/(1/\sqrt{3} - (1+1/\sqrt{3})\varepsilon_A) + \varepsilon)(\|\widehat{\mathbf{Y}}\|_F/(1-\varepsilon))$, then GBCD can exactly recover the support set Γ.

Proof. Consider $n = 1$. The initial value is $\mathbf{X}(0) = \mathbf{0}$. In order to guarantee that GBCD selects a correct index $i_0 \in \Gamma$, combining step (4) of Algorithm 1 and (20), we should verify the following inequality:

$$\max_{i\in\Gamma}\left\|\mathbf{x}^i(1) - \mathbf{x}^i(0)\right\|_2 \tag{23}$$

$$= \max_{i\in\Gamma}\left\|\frac{\mathbf{p}^i(0)}{\|\mathbf{p}^i(0)\|_2}\max\left(0, \|\mathbf{p}^i(0)\|_2 - \lambda\beta\right)\right\|_2 \tag{24}$$

$$> \max_{j\in\Gamma^c}\left\|\mathbf{x}^j(1) - \mathbf{x}^j(0)\right\|_2 \tag{25}$$

$$= \max_{j\in\Gamma^c}\left\|\frac{\mathbf{p}^j(0)}{\|\mathbf{p}^j(0)\|_2}\max\left(0, \|\mathbf{p}^j(0)\|_2 - \lambda\beta\right)\right\|_2. \tag{26}$$

If $\|\mathbf{p}^j(0)\|_2 - \lambda\beta \le 0$ ($j \in \Gamma^c$), the right-hand-side is 0. Then inequality (26) holds. Thus, we only consider $\|\mathbf{p}^j(0)\|_2 - \lambda\beta > 0$. Using Remark 1 in [2], inequality (26) is true when

$$\max_{i\in\Gamma}\left\|\mathbf{p}^i(0)\right\|_2 > \max_{j\in\Gamma^c}\left\|\mathbf{p}^j(0)\right\|_2. \tag{27}$$

Now, it is sufficient to verify (27). Let us construct an upper bound for $\max_{j\in\Gamma^c}\|\mathbf{p}^j(0)\|_2$. By step (3) of Algorithm 1, we have

$$\max_{j\in\Gamma^c}\left\|\mathbf{p}^j(0)\right\|_2 \tag{28}$$

$$= \max_{j\in\Gamma^c}\left\|\mathbf{x}^j(0) - \beta\widehat{\mathbf{a}}_j'\left(\widehat{\mathbf{A}}\mathbf{X}(0) - \widehat{\mathbf{Y}}\right)\right\|_2 \tag{29}$$

$$= \max_{j\in\Gamma^c}\left\|\beta\widehat{\mathbf{a}}_j'\widehat{\mathbf{Y}}\right\|_2 \tag{30}$$

$$= \left\|\beta\widehat{\mathbf{A}}_{\Gamma^c}'\left(\widehat{\mathbf{A}}\mathbf{X} - \mathbf{E}\mathbf{X} + \mathbf{N}\right)\right\|_{2,\infty} \tag{31}$$

$$\le \max_{j\in\Gamma^c}\left\|\beta\widehat{\mathbf{a}}_j'\widehat{\mathbf{A}}\mathbf{X}\right\|_2 + \beta\left\|\widehat{\mathbf{A}}_{\Gamma^c}'\right\|_{2,\infty}\|-\mathbf{E}\mathbf{X} + \mathbf{N}\|_F \tag{32}$$

$$= \beta\max_{j\in\Gamma^c}\left|\left\langle\widehat{\mathbf{A}}\mathbf{e}_j, \widehat{\mathbf{A}}\mathbf{X}\right\rangle_F\right| + \beta\left\|\widehat{\mathbf{A}}_{\Gamma^c}'\right\|_{2,\infty}\|-\mathbf{E}\mathbf{X} + \mathbf{N}\|_F \tag{33}$$

$$\le \beta\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F + \beta\varepsilon_0, \tag{34}$$

where (32) is from the property of norm and (34) follows from each column of $\widehat{\mathbf{A}}$ which is of unit-norm, Lemmas 3 and 4.

To prove (27), we only need to prove

$$\max_{i\in\Gamma}\left\|\mathbf{p}^i(0)\right\|_2 > \beta\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F + \beta\varepsilon_0. \tag{35}$$

We then go on to show by contradiction that (35) is true. For all $i \in \Gamma$, assume that

$$\left\|\mathbf{p}^i(0)\right\|_2 \le \beta\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F + \beta\varepsilon_0. \tag{36}$$

Then we have

$$\begin{aligned}\left\|\mathbf{P}^\Gamma(0)\right\|_F &= \sqrt{\sum_{i\in T}\|\mathbf{p}^i(0)\|_2^2} \\ &\le \beta\sqrt{K}\left(\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F + \varepsilon_0\right).\end{aligned} \tag{37}$$

Using the triangle inequality, we can get

$$\begin{aligned}\left\|\mathbf{P}^{\Gamma}(0)\right\|_F &= \left\|\mathbf{X}^{\Gamma}(0)-\beta\widehat{\mathbf{A}}_{\Gamma}'\left(\widehat{\mathbf{A}}\mathbf{X}(0)-\widehat{\mathbf{Y}}\right)\right\|_F \\ &= \left\|\beta\widehat{\mathbf{A}}_{\Gamma}'\widehat{\mathbf{Y}}\right\|_F \end{aligned} \tag{38}$$

$$\geq \beta\left\|\widehat{\mathbf{A}}_{\Gamma}'\widehat{\mathbf{A}}\mathbf{X}\right\|_F-\beta\left\|\widehat{\mathbf{A}}_{\Gamma}'(-\mathbf{E}\mathbf{X}+\mathbf{N})\right\|_F \tag{39}$$

$$\geq \beta\left(\sqrt{1-\widehat{\delta}_{K+1}}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F-\sqrt{1+\widehat{\delta}_{K+1}}\varepsilon_0\right), \tag{40}$$

where (40) is from (10) and the property of norm.

After straightforward manipulations, we have

$$\begin{aligned}&\left\|\mathbf{P}^{\Gamma}(0)\right\|_F \\ &\geq \beta\sqrt{K}\left(\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F+\varepsilon_0\right) \\ &\quad+\beta\underbrace{\left(\sqrt{1-\widehat{\delta}_{K+1}}-\sqrt{K}\widehat{\delta}_{K+1}\right)}_{\geq 0}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F \\ &\quad-\beta\left(\sqrt{1+\widehat{\delta}_{K+1}}+\sqrt{K}\right)\varepsilon_0 \end{aligned} \tag{41}$$

$$\begin{aligned}&\geq \beta\sqrt{K}\left(\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F+\varepsilon_0\right) \\ &\quad+\beta\left(\sqrt{1-\widehat{\delta}_{K+1}}-\sqrt{K}\widehat{\delta}_{K+1}\right)\sqrt{1-\widehat{\delta}_{K+1}}\left\|\mathbf{X}\right\|_F \\ &\quad-\beta\left(\sqrt{1+\widehat{\delta}_{K+1}}+\sqrt{K}\right)\varepsilon_0 \end{aligned} \tag{42}$$

$$\begin{aligned}&> \beta\sqrt{K}\left(\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F+\varepsilon_0\right)+(K-1)\sqrt{1+\widehat{\delta}_{K+1}}\varepsilon_0 \\ &\quad+\left(K-\sqrt{K}\right)\varepsilon_0 \end{aligned} \tag{43}$$

$$\geq \beta\sqrt{K}\left(\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}\mathbf{X}\right\|_F+\varepsilon_0\right), \tag{44}$$

where (41) follows from (21) and (43) follows from $\|\mathbf{X}\|_F \geq \sqrt{K}t_0$ and (22).

Obviously, (44) contradicts (37), so this fact guarantees (27).

Assume that GBCD always picks up indices from the support Γ for $n \leq k$ ($k \geq 1$ is an integer). Consider $n = k+1$. In order to prove that GBCD can choose a correct index $i_0 \in \Gamma$, analogous to [2], inequality (46) should be verified.

$$\max_{i\in\Gamma}\left\|\mathbf{p}^i(n-1)-\mathbf{x}^i(n-1)\right\|_2 \tag{45}$$

$$> \max_{j\in\Gamma^c}\left\|\mathbf{p}^j(n-1)-\mathbf{x}^j(n-1)\right\|_2. \tag{46}$$

Combining step (3) of Algorithm 1 with (46) yields

$$\max_{i\in\Gamma}\left\|\widehat{\mathbf{a}}_i'\left(\widehat{\mathbf{A}}\mathbf{X}(n-1)-\widehat{\mathbf{Y}}\right)\right\|_2 \tag{47}$$

$$> \max_{j\in\Gamma^c}\left\|\widehat{\mathbf{a}}_j'\left(\widehat{\mathbf{A}}\mathbf{X}(n-1)-\widehat{\mathbf{Y}}\right)\right\|_2. \tag{48}$$

It is sufficient to prove that (48) holds. Note that $\operatorname{supp}(\mathbf{X}(n-1)) \subseteq \Gamma$; we have

$$\max_{j\in\Gamma^c}\left\|\widehat{\mathbf{a}}_j'\left(\widehat{\mathbf{A}}\mathbf{X}(n-1)-\widehat{\mathbf{Y}}\right)\right\|_2 \tag{49}$$

$$\leq \widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}(\mathbf{X}(n-1)-\mathbf{X})\right\|_F+\varepsilon_0. \tag{50}$$

Now, we only need to prove

$$\max_{i\in\Gamma}\left\|\widehat{\mathbf{a}}_i'\left(\widehat{\mathbf{A}}\mathbf{X}(n-1)-\widehat{\mathbf{Y}}\right)\right\|_2 \tag{51}$$

$$> \widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}(\mathbf{X}(n-1)-\mathbf{X})\right\|_F+\varepsilon_0. \tag{52}$$

We then show that (52) is true by contradiction. For all $i \in \Gamma$, assume that

$$\begin{aligned}&\left\|\widehat{\mathbf{a}}_i'\left(\widehat{\mathbf{A}}\mathbf{X}(n-1)-\widehat{\mathbf{Y}}\right)\right\|_2 \\ &\quad\leq \widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}(\mathbf{X}(n-1)-\mathbf{X})\right\|_F+\varepsilon_0. \end{aligned} \tag{53}$$

Using the definition of Frobenius norm, we have

$$\begin{aligned}&\left\|\widehat{\mathbf{A}}_{\Gamma}'\left(\widehat{\mathbf{A}}\mathbf{X}(n-1)-\widehat{\mathbf{Y}}\right)\right\|_F \\ &\quad= \sqrt{\sum_{i\in\Gamma}\left\|\widehat{\mathbf{a}}_i'\left(\widehat{\mathbf{A}}\mathbf{X}(n-1)-\widehat{\mathbf{Y}}\right)\right\|_2^2} \end{aligned} \tag{54}$$

$$\leq \sqrt{K}\left(\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}(\mathbf{X}(n-1)-\mathbf{X})\right\|_F+\varepsilon_0\right). \tag{55}$$

Combining $\|\mathbf{X}(n-1)-\mathbf{X}\|_F \geq t_0$, (21), and (22), we have

$$\left\|\widehat{\mathbf{A}}_{\Gamma}'\left(\widehat{\mathbf{A}}\mathbf{X}(n-1)-\widehat{\mathbf{Y}}\right)\right\|_F \tag{56}$$

$$\geq \sqrt{1-\widehat{\delta}_{K+1}}\left\|\widehat{\mathbf{A}}(\mathbf{X}(n-1)-\mathbf{X})\right\|_F-\sqrt{1+\widehat{\delta}_{K+1}}\varepsilon_0 \tag{57}$$

$$\begin{aligned}&= \sqrt{K}\left(\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}(\mathbf{X}(n-1)-\mathbf{X})\right\|_F+\varepsilon_0\right) \\ &\quad+\left(\sqrt{1-\widehat{\delta}_{K+1}}-\sqrt{K}\widehat{\delta}_{K+1}\right)\left\|\widehat{\mathbf{A}}(\mathbf{X}(n-1)-\mathbf{X})\right\|_F \\ &\quad-\left(\sqrt{1+\widehat{\delta}_{K+1}}+\sqrt{K}\right)\varepsilon_0 \end{aligned} \tag{58}$$

$$> \sqrt{K}\left(\widehat{\delta}_{K+1}\left\|\widehat{\mathbf{A}}(\mathbf{X}(n-1)-\mathbf{X})\right\|_F+\varepsilon_0\right), \tag{59}$$

where (59) follows from

$$\left(\sqrt{1-\widehat{\delta}_{K+1}}-\sqrt{K}\widehat{\delta}_{K+1}\right)\left\|\widehat{\mathbf{A}}(\mathbf{X}(n-1)-\mathbf{X})\right\|_F \tag{60}$$

$$> \left(\sqrt{1+\widehat{\delta}_{K+1}}+\sqrt{K}\right)\varepsilon_0. \tag{61}$$

This contradicts (53). Thus, (48) is true. □

Remark 6. The weaker the RIC bound is, the less required number of measurements we need, and the improved RIC results can be used in many CS-based applications [16]. In the work of [2], the authors provided that the condition for GBCD is $\delta_{K+1} < 1/(\sqrt{K}+1)$. Obviously, it is smaller than the bound $(\sqrt{4K+1}-1)/2K$ in (21).

4. The Counterexample

Consider the measurements

$$\widehat{\mathbf{Y}} = \left(\widehat{\mathbf{A}} - \mathbf{E}\right)\mathbf{X} + \mathbf{N} = \widehat{\mathbf{A}}\mathbf{X} - \mathbf{E}\mathbf{X} + \mathbf{N}. \tag{62}$$

In this section, giving a matrix $\widehat{\mathbf{A}}$, whose RIC is a slight relaxation of $1/\sqrt{K+1}$, we will verify that GBCD can fail to recover the support of sparse matrix from (62).

Let

$$\mathbf{X} = \begin{pmatrix} \frac{\sqrt{2}}{2}t_0 & \frac{\sqrt{2}}{2}t_0 \\ \frac{\sqrt{2}}{2}t_0 & \frac{\sqrt{2}}{2}t_0 \\ \vdots & \vdots \\ \frac{\sqrt{2}}{2}t_0 & \frac{\sqrt{2}}{2}t_0 \\ 0 & 0 \end{pmatrix}_{(K+1)\times 2},$$

$$\mathbf{E} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0.5 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}_{(K+1)\times(K+1)}, \tag{63}$$

$$\mathbf{N} = \begin{pmatrix} 0 & 0 \\ \frac{\sqrt{2}}{4}t_0 & \frac{\sqrt{2}}{4}t_0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ \frac{\sqrt{2}}{2}\epsilon & \frac{\sqrt{2}}{2}\epsilon \end{pmatrix}_{(K+1)\times 2},$$

where $\operatorname{supp}(\mathbf{X}) = \{1, 2, \ldots, K\} = \Gamma$, $\Gamma^c = \{K+1\}$ and $t_0/\epsilon > 1 + 1/K$ (the value of ϵ is far less than 1; this is reasonable).

The matrix $\widehat{\mathbf{A}}$ is constructed as

$$\widehat{\mathbf{A}} = \begin{pmatrix} a & 0 & 0 & \cdots & 0 & s \\ 0 & a & 0 & \cdots & 0 & s \\ 0 & 0 & a & \cdots & 0 & s \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & s \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}_{(K+1)\times(K+1)}, \tag{64}$$

where

$$s = \frac{\delta}{\sqrt{K}}, \qquad a = \sqrt{1-\delta^2}. \tag{65}$$

Set

$$\delta = \frac{\sqrt{1-\epsilon/t_0}}{\sqrt{K+1}}. \tag{66}$$

The eigenvalues $\{\lambda_i\}_{i=1}^{K+1}$ of $\widehat{\mathbf{A}}'\widehat{\mathbf{A}}$ are

$$\begin{aligned} \lambda_i &= 1-\delta^2, \quad 1 \le i \le K-1, \\ \lambda_K &= 1-\delta, \\ \lambda_{K+1} &= 1+\delta. \end{aligned} \tag{67}$$

Thus, the RIC of $\widehat{\mathbf{A}}$ is $\overline{\delta}_{K+1}(\widehat{\mathbf{A}}) = \delta$.

Recall that condition (27) is the criterion of recovery for GBCD. Note that $\|\mathbf{p}^i(0)\|_2 = \|\beta\widehat{\mathbf{a}}_i'\widehat{\mathbf{Y}}\|_2$. One can obtain

$$\max_{i\in\Gamma}\left\|\widehat{\mathbf{a}}_i'\widehat{\mathbf{Y}}\right\|_2 = \left\|\left(\frac{\sqrt{2}}{2}t_0a^2, \frac{\sqrt{2}}{2}t_0a^2\right)\right\|_2 = a^2t_0. \tag{68}$$

On the other hand, we have

$$\begin{aligned} &\max_{j\in\Gamma^c}\left\|\widehat{\mathbf{a}}_j'\widehat{\mathbf{Y}}\right\|_2 \\ &\quad = \left\|\left(\frac{\sqrt{2}}{2}t_0Kas + \frac{\sqrt{2}}{2}\epsilon, \frac{\sqrt{2}}{2}t_0Kas + \frac{\sqrt{2}}{2}\epsilon\right)\right\|_2 \\ &\quad = Kast_0 + \epsilon. \end{aligned} \tag{69}$$

It can be derived that

$$Kast_0 + \epsilon - t_0a^2 = t_0\left(aKs - a^2\right) + \epsilon \tag{70}$$

$$= t_0\left(\sqrt{1-\delta^2}\sqrt{K}\delta - 1 + \delta^2\right) + \epsilon \tag{71}$$

$$> 0, \tag{72}$$

where (71) and (72) follow from (65) and (66).

It is obviously in contradiction to (27). Thus, GBCD fails to recover support Γ.

Remark 7. In the work of [2], the authors presented a matrix $\overline{\mathbf{A}}$ whose RIC is $\delta_{K+1}(\overline{\mathbf{A}}) = 1/\sqrt{K} - \epsilon^2/\sqrt{K}(t_0^2K + \epsilon^2)$. They showed that the GBCD algorithm fails when using $\overline{\mathbf{A}}$ as measurement matrix. After a simple calculation, we can get

$$\frac{\sqrt{1-\epsilon/t_0}}{\sqrt{K+1}} < \frac{1}{\sqrt{K}} - \frac{\epsilon^2}{\sqrt{K}\left(t_0^2K + \epsilon^2\right)}. \tag{73}$$

Thus, our result improves this existing result.

5. Experimental Results

In this section, under the total perturbations, we test the performance of the GBCD algorithm for solving the DOA estimation problem.

Consider K narrowband far-field point source signals impinging on an m-element uniform linear array. The steering vector of the matrix $\mathbf{A}$ is

$$\mathbf{a}_i = \left[1 \quad e^{-j\pi\cos\theta_{i-1}} \quad \cdots \quad e^{-j(m-1)\pi\cos\theta_{i-1}}\right]', \tag{74}$$

where $1 \le i \le N$. L is the number of snapshots.

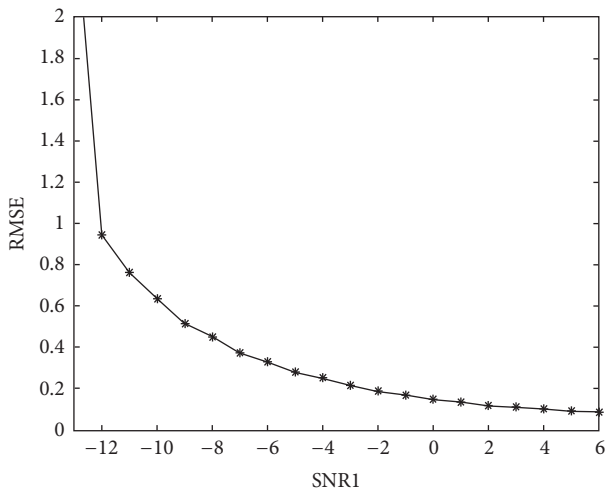

Figure 1: For SNR2 = 10. The RMSE of GBCD versus input SNR1.

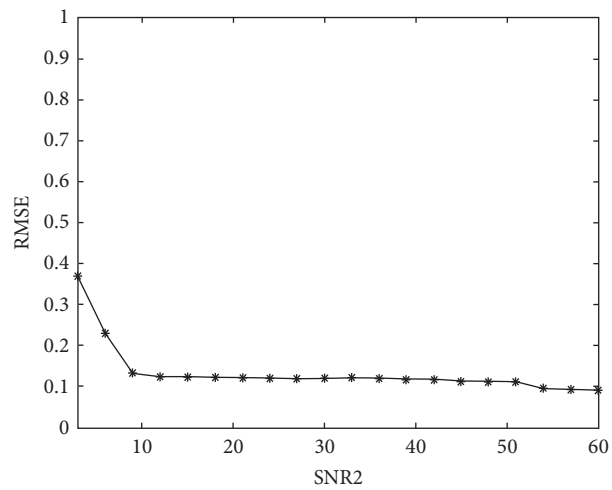

Figure 2: For SNR1 = 2. RMSE of GBCD versus input SNR2.

Using the sparse optimization approach in [1], the DOA estimation problem can be rewritten as model (1). Then the aim is hence to find out which row of the matrix $\mathbf{X}$ is nonzero, that is, the support of the matrix $\mathbf{X}$.

Analogous to the simulation of [1], we have the following assumptions:

(i) The number of the array elements is $m = 11$.

(ii) The number of snapshots is $L = 200$.

(iii) The grid spacing is 1° from 0° to 180°. Then $N = 181$.

(iv) Five ($K = 5$) uncorrelated signals impinge from $\theta_{\ell_1} = 30^\circ$, $\theta_{\ell_2} = 80^\circ$, $\theta_{\ell_3} = 100^\circ$, $\theta_{\ell_4} = 120^\circ$, and $\theta_{\ell_5} = 145^\circ$.

(v) Both the signals and the noise are white and follow Gaussian distributions. The power of nonzero entries of $\mathbf{X}$ is σ^2, and the power of each entry of $\mathbf{N}$ is σ_N^2.

(vi) Use the following SNR1 and SNR2 to measure noises $\mathbf{E}$ and $\mathbf{N}$, respectively:

$$\mathrm{SNR1} = 10\log_{10}\left(\frac{\sigma^2}{\sigma_N^2}\right). \tag{75}$$

$$\mathrm{SNR2} = \frac{\|\mathbf{A}\|_F}{\|\mathbf{E}\|_F}. \tag{76}$$

Define the root mean square error (RMSE) of 500 Monte Carlo trials as the performance index:

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{500}\sum_{k=1}^{K}\frac{\left(\widehat{\theta}_{\ell_k}(i) - \theta_{\ell_k}\right)^2}{500K}}, \tag{77}$$

where $\widehat{\theta}_{\ell_k}(i)$ is the estimate of θ_{ℓ_k} at the ith trial.

Figure 1, fixing matrix $\mathbf{E}$, describes the performance of GBCD. The results show that RMSE decreases as SNR1 increases. Figure 2, fixing matrix $\mathbf{N}$, describes the performance of GBCD. The results show that RMSE decreases as SNR2 increases. Thus, the performance of GBCD still is robust under the total perturbations.

6. Conclusion

In this paper, using the near orthogonality property, we provide a recovery condition for GBCD under the total perturbations. A counterexample is presented to show that GBCD fails. By experiments, we point out that GBCD is robust under the total perturbations.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by National Natural Science Foundation of China (nos. 11526081, 11601134, 11671122, and U1404603), the Scientific Research Foundation for Ph.D. of Henan Normal University (no. qd14142), and the Key Scientific Research Project of Colleges and Universities in Henan Province (no. 17A110008).

References

[1] X. Wei, Y. Yuan, and Q. Ling, "DOA estimation using a greedy block coordinate descent algorithm," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6382–6394, 2012.

[2] H. Li, Y. Fu, R. Hu, and R. Rong, "Perturbation analysis of greedy block coordinate descent under RIP," *IEEE Signal Processing Letters*, vol. 21, no. 5, pp. 518–522, 2014.

[3] L.-H. Chang and J.-Y. Wu, "An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit," *IEEE. Transactions on Information Theory*, vol. 60, no. 9, pp. 5702–5715, 2014.

[4] T. Blumensath and M. Davies, "Compressed sensing and source separation," in *Proceedings of the International Conference on Independent Component Analysis and Signal Separation (ICA '07)*, pp. 341–348, 2007.

[5] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2275–2284, 2009.

[6] A. C. Fannjiang, T. Strohmer, and P. Yan, "Compressed remote sensing of sparse objects," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 595–618, 2010.

[7] M. A. Herman and T. Strohmer, "General deviants: an analysis of perturbations in compressed sensing," *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 342–349, 2010.

[8] J. Ding, L. Chen, and Y. Gu, "Perturbation analysis of orthogonal matching pursuit," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 398–410, 2013.

[9] B. Li, Y. Shen, Z. Wu, and J. Li, "Sufficient conditions for generalized Orthogonal Matching Pursuit in noisy case," *Signal Processing*, vol. 108, pp. 111–123, 2015.

[10] T. Zhang, "Sparse recovery with orthogonal matching pursuit under RIP," *IEEE. Transactions on Information Theory*, vol. 57, no. 9, pp. 6215–6221, 2011.

[11] J. Wen, X. Zhu, and D. Li, "Improved bounds on restricted isometry constant for orthogonal matching pursuit," *Electronics Letters*, vol. 49, no. 23, pp. 1487–1489, 2013.

[12] R. Wu, W. Huang, and D.-R. Chen, "The exact support recovery of sparse signals with noise via orthogonal matching pursuit," *IEEE Signal Processing Letters*, vol. 20, no. 4, pp. 403–406, 2013.

[13] Q. Mo and Y. Shen, "A remark on the restricted isometry property in orthogonal matching pursuit," *IEEE. Transactions on Information Theory*, vol. 58, no. 6, pp. 3654–3656, 2012.

[14] J. Wang and B. Shim, "On the recovery limit of sparse signals using orthogonal matching pursuit," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4973–4976, 2012.

[15] H. Li, Y. Ma, W. Liu, and Y. Fu, "Improved analysis of greedy block coordinate descent under RIP," *Electronics Letters*, vol. 51, no. 6, pp. 488–490, 2015.

[16] C. B. Song, S. T. Xia, and X. J. Liu, "Improved analysis for subspace pursuit algorithm in terms of restricted isometry constant," *IEEE Signal Processing Letters*, vol. 21, no. 11, pp. 1365–1369, 2014.