Exact Support and Vector Recovery of Constrained Sparse Vectors via Constrained Matching Pursuit

Jinglai Shen* and Seyedahmad Mousavi

March 28, 2019

Abstract

Matching pursuit, especially its orthogonal version and other variations, is a greedy algorithm widely used in signal processing, compressed sensing, and sparse modeling and approximation. Inspired by constrained sparse signal recovery, this paper proposes a constrained matching pursuit algorithm and develops conditions for exact support and vector recovery on constraint sets via this algorithm. We show that exact recovery via constrained matching pursuit not only depends on a measurement matrix but also critically relies on a constraint set. We thus identify an important class of constraint sets, called coordinate projection admissible set, or simply CP admissible sets. This class of sets includes the Euclidean space, the nonnegative orthant, and many others arising from various applications; analytic and geometric properties of these sets are established. We then study exact vector recovery on convex, CP admissible cones for a fixed support. We provide sufficient exact recovery conditions for a general fixed support as well as necessary and sufficient recovery conditions for a fixed support of small size. As a byproduct of our results, we construct a nontrivial counterexample to the necessary conditions of exact vector recovery via the orthogonal matching pursuit given by Foucart, Rauhut, and Tropp, when the a given support is of size three. Moreover, by making use of cone properties and conic hull structure of CP admissible sets and constrained optimization techniques, we also establish sufficient conditions for uniform exact recovery on CP admissible sets in terms of the restricted isometry-like constant and the restricted orthogonality-like constant.

1 Introduction

Sparse models and representations find broad applications in numerous fields of contemporary interest [9], e.g., signal and image processing, high dimensional statistics, compressed sensing, and machine learning. Effective recovery of sparse signals from a few measurements poses challenging theoretical and numerical questions. A variety of sparse recovery schemes have been proposed and studied, including the basis pursuit and its extensions, greedy algorithms, and thresholding based algorithms [10, 22].

Originally introduced in signal processing and statistics, matching pursuit [15], and particularly the orthogonal matching pursuit (OMP) [20], is a greedy algorithm widely used in sparse signal recovery. At each step, the OMP uses the current target vector to select an additional "best" index via coordinate-wise optimization and adds it to the target support, and then updates the target vector over the new support via optimal fitting of a measurement vector. The deterministic and statistical performance of the OMP has been extensively studied in the literature [5, 26, 27, 33, 34]. In particular, the exact support and vector recovery via the OMP has been characterized in term of the restricted isometry constant with extensions to noisy measurements [16, 31]. Besides, many variations and extensions of the OMP have been developed in order to improve the recovery accuracy, effectiveness, and robustness under noise and

*Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21250, U.S.A. Emails: shenj@umbc.edu and smousav1@umbc.edu.

errors; representative examples of these variations and extensions include compressive sampling matching pursuit [18, 19], simultaneous OMP [28], stagewise OMP [8], subspace pursuit [7], generalized OMP [29], grouped OMP [25], and multipath matching pursuit [13], just to name a few; see [10] and the references therein for more details.

Sparse signals arising from diverse applications are subject to constraints, for example, the nonnegative constraint in nonnegative factorization in signal and image processing [4], the polyhedral constraint in index tracking problems in finance [32], and the monotone or shape constraint in order statistics and shape constrained estimation [23, 24]. Hence, constrained sparse recovery has attracted increasing interest from different areas, such as machine learning and sparse optimization [1, 2, 3, 11, 12, 14, 17, 30]. While matching pursuit, particularly the OMP and its variations or extensions, has been extensively studied on $\mathbb{R}^N$, its constrained version has received much less attention, especially the exact recovery on a general constraint set; exceptions include [4] where the uniqueness of the OMP recovery on the nonnegative orthant is considered. Inspired by the constrained sparse recovery, this paper proposes a constrained matching pursuit algorithm for a general constraint set, and develops conditions for exact support and vector recovery on constraint sets via this algorithm. Similar to the OMP, the constrained matching pursuit algorithm selects a new optimal index by solving a constrained coordinate-wise optimization problem at each step, and then updates its target vector over the updated support by solving another constrained optimization problem for the best fitting of a measurement vector. We show that exact recovery via the constrained matching pursuit not only depends on a measurement matrix but also critically relies on a constraint set. This motivates us to introduce an important class of constraint sets, called coordinate projection admissible sets, or simply CP admissible sets. This class of sets includes the Cartesian product of arbitrary copies of $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{R}_-$, and many others arising from applications. We establish analytic and geometric properties of these sets to be used for exact recovery analysis. We then study exact vector recovery on convex, CP admissible cones for a fixed support. When a fixed support has the size of two and three, we develop necessary and sufficient recovery conditions; when the support size is large, we provide sufficient exact recovery conditions. As a byproduct of our results, we construct a nontrivial counterexample to the necessary conditions of exact vector recovery via the OMP given by Foucart, Rauhut, and Tropp, when the size of a given support is three (cf. Section 5.1.2). Moreover, we establish sufficient conditions for uniform exact recovery on general CP admissible sets in terms of the restricted isometry-like constant and the restricted orthogonality-like constant, by leveraging cone properties and conic hull structure of CP admissible sets, the positive homogeneity of the aforementioned constants, as well as constrained optimization techniques. Its extensions are also discussed.

The rest of the paper is organized as follows. Section 2 presents the constrained matching pursuit algorithm and discusses underlying optimization problems in this algorithm. Section 3 studies basic properties of exact support recovery via constrained matching pursuit. In Section 4, the CP admissible sets are introduced, and their properties are established. Section 5 is concerned with the exact vector recovery of convex, CP admissible cones for a fixed support. In Section 6, sufficient conditions for uniform exact recovery on general convex, CP admissible sets are derived with conclusions made in Section 7.

Notation. Let A be an $m \times N$ real matrix. For any index set $\mathcal{S} \subseteq \{1, \ldots, N\}$, let $|\mathcal{S}|$ denote the cardinality of $\mathcal{S}$, $\mathcal{S}^c$ denote the complement of $\mathcal{S}$, and $A_{\bullet\mathcal{S}}$ be the matrix formed by the columns of A indexed by elements of $\mathcal{S}$. We write the ith column of A as $A_{\bullet i}$ instead of $A_{\bullet\{i\}}$. Further, $\mathbb{R}^N_+$ and $\mathbb{R}^N_{++}$ denote the nonnegative and positive orthants of $\mathbb{R}^N$ respectively, and $\mathbf{e}_j$ denotes the jth column of the $N \times N$ identity matrix. For $a \in \mathbb{R}$, let $a_+ := \max(a, 0) \geq 0$ and $a_- := \max(-a, 0) \geq 0$. For a given $x \in \mathbb{R}^N$, $\mathrm{supp}(x)$ denotes the support of x, i.e., $\mathrm{supp}(x) = \{i \,|\, x_i \neq 0\}$. The standard inner product on $\mathbb{R}^n$ is denoted by $\langle \cdot, \cdot \rangle$. When a minimization problem has multiple solutions, $x \in \mathrm{Argmin}$ denotes an arbitrary optimal solution; if there is a unique optimal solution, then we use $x = \arg\min$. Let $\mathrm{cone}(S)$ denote the conic hull of a set S in $\mathbb{R}^N$, i.e., the collection of nonnegative combinations of finitely many vectors in S. We always assume that a cone in $\mathbb{R}^n$ contains the zero vector. For two sets A and B, $A \subseteq B$ means that A is a subset of B and A possibly equals to B, while $A \subset B$ means that A is a proper subset of B. For $K \in \mathbb{N}$, let Σ_K be the set of all vectors $x \in \mathbb{R}^N$ satisfying $|\mathrm{supp}(x)| \le K$. For $u, v \in \mathbb{R}^n$, $u \perp v$ stands for the orthogonality of u and v, i.e., $u^T v = 0$.

2 Constrained Matching Pursuit: Algorithm and Preliminary Results

Consider the following constrained sparse recovery problem:

$$\min_{x \in \mathbb{R}^N} \|x\|_0 \qquad \text{subject to} \quad Ax = y, \quad x \in \mathcal{P}, \tag{1}$$

where $\|x\|_0 := |\mathrm{supp}(x)|$, $A \in \mathbb{R}^{m \times N}$ with $N > m$, $y \in \mathbb{R}^m$, and $\mathcal{P}$ is a closed constraint set in $\mathbb{R}^N$. Throughout this paper, we assume that $\mathcal{P}$ contains the zero vector, there is no measurement error so that y is in the range of A, and each column of A is nonzero, i.e., $\|A_{\bullet i}\|_2 > 0$ for each $i = 1, \ldots, N$. To solve the problem (1), we introduce the constrained matching pursuit scheme given below.

Algorithm 1 Constrained Matching Pursuit

1: Input: $A \in \mathbb{R}^{m \times N}$, $y \in \mathbb{R}^m$, $\mathcal{P} \subseteq \mathbb{R}^N$, and a stopping criteria
2: Initialize: $k = 0$, $x^0 = 0$, and $\mathcal{J}_0 = \emptyset$
3: **while** the stopping criteria is not met **do**
4: $\quad g_j^* = \min_{t \in \mathbb{R}} \|y - A(x^k + t\,\mathbf{e}_j)\|_2^2 \quad$ subject to $\; x^k + t\,\mathbf{e}_j \in \mathcal{P}$, $\forall\, j = 1, \ldots, N$
5: $\quad j_{k+1}^* \in \mathrm{Argmin}_{j \in \{1, \ldots, N\}}\, g_j^*$
6: $\quad \mathcal{J}_{k+1} = \mathcal{J}_k \cup \{j_{k+1}^*\}$
7: $\quad x^{k+1} \in \mathrm{Argmin}_{w \in \mathcal{P},\, \mathrm{supp}(w) \subseteq \mathcal{J}_{k+1}} \|Aw - y\|_2^2$
8: $\quad k \leftarrow k + 1$
9: **end while**
10: Output: $x^* = x^k$

At each step in the constrained matching pursuit algorithm, two constrained optimization problems are solved. The first problem, given in Line 4 of Algorithm 1, is a constrained coordinate-wise minimization problem; the second problem, given in Line 7 of Algorithm 1, is a minimization problem on the constraint set $\mathcal{P}$ subject to an additional support constraint $\mathrm{supp}(w) \subseteq \mathcal{J}_{k+1}$. In what follows, we discuss these two underlying problems and their solution properties.

For a given $x \in \mathcal{P}$ and an index $j = 1, \ldots, N$, the first minimization problem can be written as

$$(\mathrm{P}_{x,j}): \quad \min_{t \in \mathbb{R}} \|y - A(x + t\,\mathbf{e}_j)\|_2^2 \qquad \text{subject to} \quad x + t\,\mathbf{e}_j \in \mathcal{P}.$$

Since $\mathcal{P}$ is closed, it is easy to verify that the constraint set of $(\mathrm{P}_{x,j})$ given by

$$\mathbb{I}_j(x) := \big\{ t \in \mathbb{R} \,|\, x + \mathbf{e}_j t \in \mathcal{P} \big\} \tag{2}$$

is a closed set in $\mathbb{R}$. Besides, for any $x \in \mathcal{P}$ and $j = 1, \ldots, N$, we have $0 \in \mathbb{I}_j(x)$, and $(\mathrm{P}_{x,j})$ attains an optimal solution because $\|A_{\bullet j}\|_2 > 0$. Motivated by the fact that y is given by $y = Au$ for some $u \in \mathcal{P}$, we define, for any $u, v \in \mathcal{P}$ and $j = 1, \ldots, N$,

$$f_j^*(u, v) := \min_{t \in \mathbb{I}_j(v)} \|Au - A(v + t\,\mathbf{e}_j)\|_2^2 = \min_{t \in \mathbb{I}_j(v)} \|A(u - v) - tA_{\bullet j}\|_2^2.$$

A particularly interesting and important case is when $\mathcal{P}$ is closed and convex. In this case, for any $v \in \mathcal{P}$ and any index j, $\mathbb{I}_j(v)$ is also closed and convex and thus is a closed interval in $\mathbb{R}$. Letting $a_j(v) := \inf \mathbb{I}_j(v)$ and $b_j(v) := \sup \mathbb{I}_j(v)$, where $a_j(v) \in \mathbb{R}_- \cup \{-\infty\}$ and $b_j(v) \in \mathbb{R}_+ \cup \{+\infty\}$, $\mathbb{I}_j(v)$ can be written as $\mathbb{I}_j(v) = [a_j(v), b_j(v)]$. For any given $u, v \in \mathcal{P}$, since $A_{\bullet j} \neq 0$, the minimization problem $\min_{t\in[a_j(v),b_j(v)]} \|A(u-v) - t\,A_{\bullet j}\|_2^2$ attains a unique optimal solution

$$t_j^*(u,v) \;=\; \begin{cases} a_j(v), & \text{if } \widetilde{t}_j(u,v) \le a_j(v) \\ \widetilde{t}_j(u,v), & \text{if } \widetilde{t}_j(u,v) \in [a_j(v), b_j(v)] \\ b_j(v), & \text{if } \widetilde{t}_j(v) \ge b_j(v) \end{cases},$$

where

$$\widetilde{t}_j(u,v) \;:=\; \langle A(u-v), A_{\bullet j}\rangle / \|A_{\bullet j}\|_2^2. \tag{3}$$

Consequently,

$$f_j^*(u,v) \;=\; \begin{cases} \|A(u-v)\|_2^2 - \|A_{\bullet j}\|_2^2 \cdot [2a_j(v)\widetilde{t}_j(u,v) - a_j^2(v)], & \text{if } \widetilde{t}_j(u,v) \le a_j(v) \\ \|A(u-v)\|_2^2 - \|A_{\bullet j}\|_2^2 \cdot \widetilde{t}_j^{\,2}(u,v), & \text{if } \widetilde{t}_j(u,v) \in [a_j(v), b_j(v)] \\ \|A(u-v)\|_2^2 - \|A_{\bullet j}\|_2^2 \cdot [2b_j(v)\widetilde{t}_j(u,v) - b_j^2(v)], & \text{if } \widetilde{t}_j(v) \ge b_j(v) \end{cases} \tag{4}$$

For illustration, we show the expressions of $f_j^*(u,v)$ for two special cases below.

(i) $\mathbb{I}_j(v) = \mathbb{R}$, i.e., $a_j(v) = -\infty$ and $b_j(v) = +\infty$. In this case,

$$f_j^*(u,v) \;=\; \|A(u-v)\|_2^2 - \|A_{\bullet j}\|_2^2 \cdot \widetilde{t}_j^{\,2}(u,v). \tag{5}$$

(ii) $\mathbb{I}_j(v) = \mathbb{R}_+$, i.e., $a_j(v) = 0$ and $b_j(v) = +\infty$. In this case,

$$f_j^*(u,v) \;=\; \|A(u-v)\|_2^2 - \|A_{\bullet j}\|_2^2 \cdot \big([\widetilde{t}_j(u,v)]_+\big)^2. \tag{6}$$

We next study the constrained minimization problem pertaining to that in Line 7 of Algorithm 1 for a given $y \in \mathbb{R}^m$ and a given index set $\mathcal{J} \subseteq \{1, \ldots, N\}$:

$$(\mathrm{P}_{y,\mathcal{J}}): \quad \min_{w\in\mathbb{R}^N} \|Aw - y\|_2^2 \qquad \text{subject to} \quad w \in \mathcal{P} \quad \text{and} \quad \mathrm{supp}(w) \subseteq \mathcal{J}. \tag{7}$$

Since $\mathcal{P}$ contains the zero vector, $(\mathrm{P}_{y,\mathcal{J}})$ is always feasible for any index set $\mathcal{J}$, even if $\mathcal{J}$ is empty. Note that we always assume that the minimization problem in Line 7 of Algorithm 1 has a solution in each step. Moreover, certain solution existence and uniqueness results for $(\mathrm{P}_{y,\mathcal{J}})$ can be established under mild assumptions on A and $\mathcal{P}$ as shown below.

Lemma 2.1. *Let the set $\mathcal{P} \subseteq \mathbb{R}^N$ and the matrix $A \in \mathbb{R}^{m\times N}$. The following hold:*

(i) If $A\mathcal{P}$ is closed, then for any index set $\mathcal{J}$ and any $y \in \mathbb{R}^m$, $(P_{y,\mathcal{J}})$ attains an optimal solution.

(ii) If $\mathcal{P}$ is closed and an index set $\mathcal{I}$ is such that $A_{\bullet\mathcal{I}}$ has linearly independent columns, then $(P_{y,\mathcal{I}})$ has an optimal solution. If, in addition, $\mathcal{P}$ is convex, then such an optimal solution is unique.

Proof. (i) Given any $y \in \mathbb{R}^m$ and any index set $\mathcal{J}$, $(\mathrm{P}_{y,\mathcal{J}})$ is equivalent to $\min_{w\in\mathcal{P}\cap\mathcal{V}} \|Aw - y\|_2^2$, where $\mathcal{V} := \{z = (z_{\mathcal{J}}, z_{\mathcal{J}^c}) \,|\, z_{\mathcal{J}^c} = 0\}$ is a subspace of $\mathbb{R}^N$. Note that $A\mathcal{V}$ is a subspace and thus closed. Since $A(\mathcal{P}\cap\mathcal{V}) = (A\mathcal{P}) \cap (A\mathcal{V})$ and $A\mathcal{P}$ is closed, $A(\mathcal{P}\cap\mathcal{V})$ is also closed. Moreover, the function $\|\cdot\|_2^2$ is continuous, coercive, and bounded below on $\mathbb{R}^m$. By [17, Lemma 4.1], $(\mathrm{P}_{y,\mathcal{J}})$ has an optimal solution.

(ii) Suppose $\mathcal{P}$ is closed. Then the set $\mathcal{P}_{\mathcal{J}} := \mathcal{P}\cap\mathcal{V}$ is closed for any index set $\mathcal{J}$, where $\mathcal{V}$ is the subspace associated with $\mathcal{J}$ defined in the proof for (i). Since $A_{\bullet\mathcal{I}}$ has linearly independent columns, it is easy to see that $\{A_{\bullet\mathcal{I}}\, w_{\mathcal{I}} \,|\, (w_{\mathcal{I}}, 0) \in \mathcal{P}_{\mathcal{I}}\}$ is closed. By the similar argument for (i), $(\mathrm{P}_{y,\mathcal{I}})$ attains an optimal solution. If, in addition, $\mathcal{P}$ is convex, then $(\mathrm{P}_{y,\mathcal{I}})$ is a convex optimization problem with a strongly convex objective function in $w_{\mathcal{I}}$. This yields a unique optimal solution for any $y \in \mathbb{R}^m$. $\square$

Typical constraint sets $\mathcal{P}$ satisfying the closedness assumption given in statement (i) of Lemma 2.1 for an arbitrary matrix $A \in \mathbb{R}^{m\times N}$ include compact sets and polyhedral sets, e.g., $\mathbb{R}^N$ and $\mathbb{R}^N_+$. Also see Corollary 4.1 in Section 4 for a general class of sets on which $(\mathrm{P}_{y,\mathcal{J}})$ attains a solution.

When $(\mathrm{P}_{y,\mathcal{J}})$ is a convex optimization problem (whose $\mathcal{P}$ is closed and convex), well developed numerical solvers can be exploited to solve $(\mathrm{P}_{y,\mathcal{J}})$, e.g., the gradient projection method and primal-dual schemes, provided that it has a solution. In particular, the necessary and sufficient optimality condition for an optimal solution $w^* = (w^*_{\mathcal{J}}, 0) \in \mathcal{P}$ of $(\mathrm{P}_{y,\mathcal{J}})$ is given by the variational inequality (VI): $\langle A^T_{\bullet\mathcal{J}}(A_{\bullet\mathcal{J}} w^*_{\mathcal{J}} - y), w_{\mathcal{J}} - w^*_{\mathcal{J}}\rangle \geq 0$ for all $(w_{\mathcal{J}}, 0) \in \mathcal{P}$. When $\mathcal{P}$ is a closed convex cone, the above VI is equivalent to the cone complementarity problem: $\mathcal{C} \ni w^*_{\mathcal{J}} \perp A^T_{\bullet\mathcal{J}}(A_{\bullet\mathcal{J}} w^*_{\mathcal{J}} - y) \in \mathcal{C}^*$, where the closed convex cone $\mathcal{C} := \{w_{\mathcal{J}} \,|\, (w_{\mathcal{J}}, 0) \in \mathcal{P}\}$ and $\mathcal{C}^*$ denotes the dual cone of $\mathcal{C}$. Especially, when $\mathcal{P} = \mathbb{R}^N_+$, it is further equivalent to the linear complementarity problem (LCP): $0 \leq w^*_{\mathcal{J}} \perp A^T_{\bullet\mathcal{J}}(A_{\bullet\mathcal{J}} w^*_{\mathcal{J}} - y) \geq 0$. These optimality conditions will be invoked in the subsequent sections.

3 Exact Support Recovery via Constrained Matching Pursuit

Fix $K \in \mathbb{N}$ with $K < N$ throughout the rest of the paper. For a given $z \in \Sigma_K \cap \mathcal{P}$, let $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ be a sequence of triples generated by Algorithm 1 with $y = Az$ starting from $x^0 = 0$ and $\mathcal{J}_0 = \emptyset$, where $\mathcal{J}_{k+1} = \mathcal{J}_k \cup \{j^*_{k+1}\}$ such that $\mathcal{J}_0 \subseteq \mathcal{J}_1 \subseteq \cdots \subseteq \mathcal{J}_k \subseteq \cdots$. Note that there are multiple sequences in general for a given z, since the optimization problems in Lines 5 and 7 of Algorithm 1 may attain multiple solutions at each step. For example, if the underlying optimization problem (7) is a convex minimization problem with non-unique solutions for some $\mathcal{J} = \mathcal{J}_k$ and $y = Az$, then it attains infinitely many x^k's. In this case, there are infinitely many sequences $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$.

Definition 3.1. Given a matrix $A \in \mathbb{R}^{m\times N}$ and a constraint set $\mathcal{P}$, we say that *the exact support recovery* of a vector $z \in \Sigma_K \cap \mathcal{P}$ is achieved from $y = Az$ via constrained matching pursuit given by Algorithm 1, if along an *arbitrary* sequence $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ for the given z, there exists an index $s \in \mathbb{N}$ such that $\mathcal{J}_s = \mathrm{supp}(z)$. If the exact support recovery of any $z \in \Sigma_K \cap \mathcal{P}$ is achieved, then we call *the exact support recovery on* $\Sigma_K \cap \mathcal{P}$ (or simply the exact support recovery) is achieved via constrained matching pursuit.

Necessary and sufficient conditions for the exact support recovery are given as follows.

Lemma 3.1. *Given $0 \neq u \in \sum_K \cap \mathcal{P}$ and an index set $\mathcal{J} \subseteq supp(u)$, let v be an optimal solution to $\min_{w\in\mathcal{P},\ supp(w)\subseteq\mathcal{J}} \|A(u-w)\|_2^2$, where we assume that such a solution exists. Then $f^*_j(u, v) = \|A(u-v)\|_2^2$ for each $j \in \mathcal{J}$, and $f^*_j(u, v) \leq \|A(u-v)\|_2^2$ for each $j \notin \mathcal{J}$.*

Proof. Consider an arbitrary $j \notin \mathcal{J}$. Noting that $0 \in \mathbb{I}_j(v)$, we have $f^*_j(u, v) \leq \|A(u-v)\|_2^2$. We then consider an arbitrary $j \in \mathcal{J}$. For any $t \in \mathbb{I}_j(v)$, we have $v + \mathbf{e}_j t \in \mathcal{P}$ and $\mathrm{supp}(v + \mathbf{e}_j t) \subseteq \mathcal{J}$. Since v is an optimal solution to $\min_{w\in\mathcal{P},\ \mathrm{supp}(w)\subseteq\mathcal{J}} \|A(u-w)\|_2^2$, we have $\|A(u-v)\|_2^2 \leq \|Au - A(v + \mathbf{e}_j t)\|_2^2$ for all $t \in \mathbb{I}_j(v)$. This shows that $\|A(u-v)\|_2^2 \leq f^*_j(u, v)$. Furthermore, $f^*_j(u, v) \leq \|A(u-v)\|_2^2$ since $0 \in \mathbb{I}_j(v)$. Therefore, $f^*_j(u, v) = \|A(u-v)\|_2^2$ for each $j \in \mathcal{J}$. ◻

Theorem 3.1. *Given a matrix $A \in \mathbb{R}^{m\times N}$ and a constraint set $\mathcal{P}$, let $0 \neq z \in \Sigma_K \cap \mathcal{P}$ with $|supp(z)| = r$. Then the exact support recovery of z is achieved via constrained matching pursuit if and only if for any sequence $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ generated by Algorithm 1 with $y = Az$, the following holds*

$$\min_{j\in supp(z)\setminus\mathcal{J}_k} f^*_j(z, x^k) < \min_{j\in[supp(z)]^c} f^*_j(z, x^k), \qquad \forall\ k = 0, 1, \ldots, r-1. \tag{8}$$

Moreover, when the exact support recovery of z is achieved, the support of z is firstly attained at the rth step along any sequence $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$, i.e., $\mathcal{J}_r = supp(z)$ and $\mathcal{J}_k \subset supp(z)$ for each $k < r$.

Proof. "If". For the given z, suppose an arbitrary sequence $\big((x^k, j_k^*, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ generated by Algorithm 1 satisfies (8). We prove below by induction on iterative steps of Algorithm 1 that $\mathcal{J}_k \subseteq \text{supp}(z)$ with $|\mathcal{J}_k| = k$ and $j_{k+1}^* \in \text{supp}(z) \setminus \mathcal{J}_k$ for each $k = 1, \ldots, r-1$. At Step 1, since $x^0 = 0$ and $\mathcal{J}_0$ is the empty set, we deduce from (8) that $\min_{j\in\text{supp}(z)} f_j^*(z, 0) < \min_{j\in[\text{supp}(z)]^c} f_j^*(z, 0)$. It follows from Algorithm 1 that the optimal index $j_1^* \in \text{Argmin}_{j=1,\ldots,N} f_j^*(z, 0)$ satisfies $j_1^* \in \text{supp}(z)$ such that $\mathcal{J}_1 = \{j_1^*\} \subseteq \text{supp}(z)$ and $|\mathcal{J}_1| = 1$. Now suppose $\mathcal{J}_k \subseteq \text{supp}(z)$ with $|\mathcal{J}_k| = k$ and $j_k^* \in \text{supp}(z) \setminus \mathcal{J}_{k-1}$ for $1 \le k \le r-2$. Consider Step $(k+1)$. In view of Lemma 3.1, the optimal index $j_{k+1}^* \in \text{Argmin}_{j=1,\ldots,N} f_j^*(z, x^k)$ satisfies $j_{k+1}^* \notin \mathcal{J}_k$. Since $\mathcal{J}_k \subseteq \text{supp}(z)$, $j_{k+1}^* \in [\text{supp}(z) \setminus \mathcal{J}_k] \cup [\text{supp}(z)]^c$. Further, it follows from (8) that $j_{k+1}^* \in \text{supp}(z) \setminus \mathcal{J}_k$. Therefore, $\mathcal{J}_{k+1} := \mathcal{J}_k \cup \{j_{k+1}^*\}$ satisfies $\mathcal{J}_{k+1} \subseteq \text{supp}(z)$ and $|\mathcal{J}_{k+1}| = k+1$. By the induction principle, we see that $\mathcal{J}_r \subseteq \text{supp}(z)$ and $|\mathcal{J}_r| = r = |\text{supp}(z)|$. This implies that $\mathcal{J}_r = \text{supp}(z)$ and $\mathcal{J}_k \subset \text{supp}(z)$ for each $k < r$.

"Only if". Suppose the exact support recovery of z is achieved via Algorithm 1. By Definition 3.1, we claim that for any given sequence $\big((x^k, j_k^*, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ generated by Algorithm 1 with $y = Az$ starting from $x^0 = 0$ and $\mathcal{J}_0 = \emptyset$, the following must hold:

$$\min_{j\in\text{supp}(z)} f_j^*(z, x^k) < \min_{j\in[\text{supp}(z)]^c} f_j^*(z, x^k), \qquad \forall\, k = 0, 1, \ldots, r-1,$$

This is because otherwise, $\min_{j\in\text{supp}(z)} f_j^*(z, x^\ell) \ge \min_{j\in[\text{supp}(z)]^c} f_j^*(z, x^\ell)$ for some $\ell = 0, 1, \ldots, r-1$. Hence, there exists an optimal index $j_{\ell+1}^* \notin \text{supp}(z)$ such that $\mathcal{J}_{\ell+1} \neq \text{supp}(z)$ (along a possibly different sequence), leading to $\mathcal{J}_s \neq \text{supp}(z)$ for all $s \ge \ell$. Note that $\mathcal{J}_k \neq \text{supp}(z)$ for each $k = 1, \ldots, \ell$ since each $|\mathcal{J}_k| < r$. Therefore, there exists a sequence so that $\mathcal{J}_k \neq \text{supp}(z)$ for all $k \in \mathbb{N}$, yielding a contradiction. Finally, since each x^k is a minimizer of $\min_{w\in\mathcal{P}, \text{supp}(w)\subseteq\mathcal{J}_k} \|A(z-w)\|_2^2$, we deduce via Lemma 3.1 that $\min_{j\in\text{supp}(z)} f_j^*(z, x^k) = \min_{j\in\text{supp}(z)\setminus\mathcal{J}_k} f_j^*(z, x^k)$. This leads to (8). □

In what follows, we show the implications of the exact support recovery.

Proposition 3.1. *Given a matrix A and a constraint set $\mathcal{P}$, let $0 \neq z \in \Sigma_K \cap \mathcal{P}$ with $|supp(z)| = r$ be such that the exact support recovery of z is achieved. Then for any sequence $\big((x^k, j_k^*, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ generated by Algorithm 1 with $y = Az$, the following hold:*

(i) $\|A(z - x^{k+1})\|_2^2 \le f_{j_{k+1}^*}^*(z, x^k) < \|A(z - x^k)\|_2^2$ for each $k = 0, 1, \ldots, r-1$;

(ii) For each $k = 1, \ldots, r$, $(x^k)_{j_k^*} \neq 0$, and $x^k_{\mathcal{J}_{k-1}} \neq 0$ when $k > 1$. Hence, $supp(x^k) = \mathcal{J}_k$ for $k = 1, 2$.

Proof. (i) Fix $k \in \{0, 1, \ldots, r-1\}$. Since x^k is an optimal solution to $\min_{w\in\mathcal{P}, \text{supp}(w)\subseteq\mathcal{J}_k} \|A(z-w)\|_2^2$, it follows from Lemma 3.1 that $f_j^*(z, x^k) \le \|A(z - x^k)\|_2^2$ for all $j = 1, \ldots, N$. In light of the inequality given by (8), we have $\min_{j\in\text{supp}(z)\setminus\mathcal{J}_k} f_j^*(z, x^k) < \min_{j\in[\text{supp}(z)]^c} f_j^*(z, x^k) \le \|A(z - x^k)\|_2^2$. Since $j_{k+1}^* \in \text{Argmin}_{j\in\text{supp}(z)\setminus\mathcal{J}_k} f_j^*(z, x^k)$, we have $f_{j_{k+1}^*}^*(z, x^k) < \|A(z - x^k)\|_2^2$. Besides, by virtue of the definition of $f_j^*(\cdot, \cdot)$, we deduce that there exists $0 \neq t_* \in \mathbb{I}_{j_{k+1}^*}$ such that

$$f_{j_{k+1}^*}^*(z, x^k) \;=\; \big\|Az - A\big(x^k + t_* \mathbf{e}_{j_{k+1}^*}\big)\big\|_2^2.$$

Note that $x^k + t_* \mathbf{e}_{j_{k+1}^*} \in \mathcal{P}$ and $\text{supp}(x^k + t_* \mathbf{e}_{j_{k+1}^*}) = \mathcal{J}_k \cup \{j_{k+1}^*\} = \mathcal{J}_{k+1}$. Since x^{k+1} is an optimal solution to $\min_{w\in\mathcal{P}, \text{supp}(w)\subseteq\mathcal{J}_{k+1}} \|A(z-w)\|_2^2$, we have $\|A(z - x^{k+1})\|_2^2 \le \|Az - A(x^k + t_* \mathbf{e}_{j_{k+1}^*})\|_2^2 = f_{j_{k+1}^*}^*(z, x^k)$.

(ii) Fix $k \in \{1, \ldots, r\}$. We first show the following claim: $x^k_{\mathcal{J}_k\setminus\mathcal{J}_s} \neq 0$ for each $s \in \{0, 1, \ldots, k-1\}$. Suppose, in contrast, that $(x^k)_{\mathcal{J}_k\setminus\mathcal{J}_s} = 0$ for some $s \in \{0, 1, \ldots, k-1\}$. In light of $\mathcal{J}_s \subset \mathcal{J}_k$, we have $\text{supp}(x^k) \subseteq \mathcal{J}_s$. Since $x^k \in \mathcal{P}$ and x^s is an optimal solution to $\min_{w\in\mathcal{P}, \text{supp}(w)\subseteq\mathcal{J}_s} \|A(z-w)\|_2^2$, we deduce that $\|A(z - x^s)\|_2^2 \le \|A(z - x^k)\|_2^2$. Since $s < k$, this yields a contradiction to statement (i). Hence, the claim holds. In view of $\mathcal{J}_k \setminus \mathcal{J}_{k-1} = \{j_k^*\}$, we obtain $(x^k)_{j_k^*} \neq 0$.

We then show that $x^k_{\mathcal{J}_{k-1}} \neq 0$ when $k > 1$. Suppose, in contrast, that $x^k_{\mathcal{J}_{k-1}} = 0$. Then $\text{supp}(x^k) = \{j^*_k\}$ since $(x^k)_{j^*_k} \neq 0$. By the definition of $f^*_j(\cdot,\cdot)$, we have that $f^*_{j^*_k}(z, 0) \le \|A(z - x^k)\|_2^2$. Furthermore, we deduce via $x^0 = 0$ that $f^*_{j^*_1}(z, x^0) \le f^*_{j^*_k}(z, 0)$. Therefore, $f^*_{j^*_1}(z, x^0) \le \|A(z-x^k)\|_2^2$. On the other hand, it follows from statement (i) that $\|A(z - x^1)\|_2^2 \le f^*_{j^*_1}(z, x^0)$. This leads to $\|A(z - x^1)\|_2^2 \le \|A(z - x^k)\|_2^2$. Since $k > 1$, we attain a contradiction to statement (i). Consequently, $x^k_{\mathcal{J}_{k-1}} \neq 0$ when $k > 1$. $\square$

We specify particular conditions for the exact support recovery on $\mathbb{R}^N$ and $\mathbb{R}^N_+$, respectively.

Corollary 3.1. *Given a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns (i.e., $\|A_{\bullet i}\|_2 = 1$ for all i) and a constraint set $\mathcal{P}$, let $0 \neq z \in \Sigma_K \cap \mathcal{P}$ with $|supp(z)| = r$. The following hold:*

(i) When $\mathcal{P} = \mathbb{R}^N$, the exact support recovery of z is achieved if and only if for any sequence $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ generated by Algorithm 1 with $y = Az$,

$$\max_{j\in supp(z)\setminus\mathcal{J}_k} |A^T_{\bullet j}A(z - x^k)| > \max_{j\in[supp(z)]^c} |A^T_{\bullet j}A(z - x^k)|, \qquad \forall\ k = 0, 1, \dots, r - 1;$$

(ii) When $\mathcal{P} = \mathbb{R}^N_+$, the exact support recovery of z is achieved if and only if for any sequence $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ generated by Algorithm 1 with $y = Az$,

$$\max_{j\in supp(z)\setminus\mathcal{J}_k} [A^T_{\bullet j}A(z - x^k)]_+ > \max_{j\in[supp(z)]^c} [A^T_{\bullet j}A(z - x^k)]_+, \qquad \forall\ k = 0, 1, \dots, r - 1.$$

Proof. (i) Let $\mathcal{P} = \mathbb{R}^N$. Then for any $v \in \mathbb{R}^N$ and any index j, $\mathbb{I}_j(v) = \mathbb{R}$. It follows from the definition of $\widetilde{t}_j(u, v)$ given by (3) and $f^*_j(u, v)$ given by (5) that (8) holds if and only if for each $k = 0, 1, \dots, r - 1$,

$$\max_{j\in\text{supp}(z)\setminus\mathcal{J}_k} \langle A(z - x^k), A_{\bullet j}\rangle^2 > \max_{j\in[\text{supp}(z)]^c} \langle A(z - x^k), A_{\bullet j}\rangle^2.$$

The latter is equivalent to $\max_{j\in\text{supp}(z)\setminus\mathcal{J}_k} |A^T_{\bullet j}A(z - x^k)| > \max_{j\in[\text{supp}(z)]^c} |A^T_{\bullet j}A(z - x^k)|$.

(ii) Let $\mathcal{P} = \mathbb{R}^N_+$. Consider the pair $(x^k, \mathcal{J}_k)$ for any fixed $k \in \{0, 1, \dots, r - 1\}$. For each $j \in \text{supp}(z) \setminus \mathcal{J}_k$, we have $(x^k)_j = 0$ such that $\mathbb{I}_j(x^k) = \mathbb{R}_+$. Further, since $\text{supp}(x^k) \subset \text{supp}(z)$ as shown in Theorem 3.1, we see that for any $j \in [\text{supp}(z)]^c$, $j \notin \text{supp}(x^k)$ such that $(x^k)_j = 0$ and $\mathbb{I}_j(x^k) = \mathbb{R}_+$. Hence, in view of $f^*_j(\cdot,\cdot)$ given by (6), we see that $\min_{j\in\text{supp}(z)\setminus\mathcal{J}_k} f^*_j(z, x^k) < \min_{j\in[\text{supp}(z)]^c} f^*_j(z, x^k)$ if and only if $\max_{j\in\text{supp}(z)\setminus\mathcal{J}_k}([A^T_{\bullet j}A(z - x^k)]_+)^2 > \max_{j\in[\text{supp}(z)]^c}([A^T_{\bullet j}A(z - x^k)]_+)^2$, which is equivalent to $\max_{j\in\text{supp}(z)\setminus\mathcal{J}_k}[A^T_{\bullet j}A(z-x^k)]_+ > \max_{j\in[\text{supp}(z)]^c}[A^T_{\bullet j}A(z-x^k)]_+$. This yields the desired result. $\square$

Inspired by Theorem 3.1, we introduce the following condition for a matrix A and a constraint set $\mathcal{P}$:

(**H**) : For any $0 \neq u \in \Sigma_K \cap \mathcal{P}$, any index set $\mathcal{J} \subset \text{supp}(u)$ (where $\mathcal{J}$ is possibly the empty set), and an arbitrary optimal solution v of $\min_{w\in\mathcal{P},\ \text{supp}(w)\subseteq\mathcal{J}} \|A(u - w)\|_2^2$, the following holds:

$$\min_{j\in\text{supp}(u)\setminus\mathcal{J}} f^*_j(u, v) < \min_{j\in[\text{supp}(u)]^c} f^*_j(u, v). \tag{9}$$

The next proposition states that (**H**) is a sufficient condition for the exact support recovery. We omit its proof since it follows directly from the fact that the inequality in (9) implies (8) given in Theorem 3.1.

Proposition 3.2. *Given a matrix $A \in \mathbb{R}^{m\times N}$ and a constraint set $\mathcal{P}$, suppose condition* (**H**) *holds. Then the exact support recovery is achieved on $\Sigma_K \cap \mathcal{P}$.*

Remark 3.1. In general, condition (**H**) is *not* necessary for the exact support recovery. This is because the exact support recovery of a vector $z \in \Sigma_K \cap \mathcal{P}$ requires that the inequality (8) hold for $\mathcal{J}_k$'s only along a sequence $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ for z, while condition (**H**) says that the inequality (9) hold for *all* proper subsets $\mathcal{J} \subset \text{supp}(z)$. Nevertheless, condition (**H**) is necessary for the exact support recovery when K is small; see Corollary 5.1 for $\Sigma_2 \cap \mathbb{R}^N$ and Corollary 5.3 for $\Sigma_2 \cap \mathbb{R}^N_+$, respectively.

Before ending this section, we give an example of a closed convex set $\mathcal{P}$, on which no matrix A can achieve the exact support recovery. It demonstrates that the exact support recovery and condition (**H**) not only depend on the measurement matrix A but also critically rely on the constraint set $\mathcal{P}$.

Example 3.1. Let $d = (d_1, \ldots, d_N)^T \in \mathbb{R}^N$ be such that $d_i \neq 0$ for each i. Consider the hyperplane $\mathcal{P} := \{x \in \mathbb{R}^N \,|\, d^T x = 0\}$. Clearly, $\mathcal{P}$ is closed and convex, and it contains the zero vector and other sparse vectors. Since each $d_i \neq 0$, it is easy to verify that for any $v \in \mathcal{P}$ and any index j, the set $\mathbb{I}_j(v) = \{0\}$. This shows that for any $u, v \in \mathcal{P}$ and any index j, $f_j^*(u, v) = \|A(u - v)\|_2^2$ for any matrix A. Hence, for any $z \in \Sigma_K \cap \mathcal{P}$, we deduce that at Step 1 of Algorithm 1, $\mathrm{Argmin}_{j \in \{1,\ldots,N\}} f^*(z, 0) = \{1, \ldots, N\}$. Thus j_1^* can be chosen as $j_1^* \notin \mathrm{supp}(z)$. This means that no matrix A achieves the exact support recovery of any $z \in \Sigma_K \cap \mathcal{P}$. It also implies that no matrix A satisfies condition (**H**) on $\mathcal{P}$.

4 Coordinate Projection Admissible Sets

Since the exact recovery via constrained matching pursuit critically relies on a constraint set, it is essential to find a class of constraint sets to which the constrained matching pursuit can be successfully applied for exact recovery. An ideal class of constraint sets is expected to satisfy some crucial conditions, including but not limited to: (i) each set in this class contains sufficiently many sparse vectors; (ii) this class of sets is broad enough to include important sets arising from applications, such as $\mathbb{R}^N$ and $\mathbb{R}^N_+$; and (iii) (relatively) easily verifiable sufficient recovery conditions can be established using general properties of this class of sets. Motivated by these requirements, we identify an important class of constraint sets in this section and study their analytic properties to be used for the exact recovery.

We introduce some notation first. Let $\mathcal{U}$ be a nonempty set in $\mathbb{R}$, and $\mathcal{I}$ be an index subset of $\{1, \ldots, N\}$. We let $\mathcal{U}^{\mathcal{I}} := \{x = (x_1, \ldots, x_N)^T \in \mathbb{R}^N \,|\, x_i \in \mathcal{U}, \forall\, i \in \mathcal{I}, \text{ and } \ x_{\mathcal{I}^c} = 0\}$, and $\mathcal{U}_{\mathcal{I}} := \{u \in \mathbb{R}^{|\mathcal{I}|} \,|\, u_i \in \mathcal{U}, \forall\, i \in \mathcal{I}\}$. For each $x \in \mathbb{R}^N$ and an index set $\mathcal{I}$, define the coordinate projection operator $\pi_{\mathcal{I}} : \mathbb{R}^N \to \mathbb{R}^N$ as $\pi_{\mathcal{I}}(x) := z$, where $z_i = x_i, \forall\, i \in \mathcal{I}$ and $z_{\mathcal{I}^c} = 0$. If $\mathcal{I}$ is the empty set, then $\pi_{\mathcal{I}}(x) = 0, \forall\, x$. We often write $\pi_{\mathcal{I}}(x) = (x_{\mathcal{I}}, 0)$ with $x_{\mathcal{I}^c} = 0$ for notational simplicity. We also write $\pi_{\{i\}}$ as π_i for $i = 1, \ldots, N$ when the context is clear. For each index set $\mathcal{I}$, $\pi_{\mathcal{I}}$ is obviously a linear operator on $\mathbb{R}^N$ given by $\pi_{\mathcal{I}}(x) = Wx$ for $x = (x_{\mathcal{I}}, x_{\mathcal{I}^c}) \in \mathbb{R}^N$, where the matrix $W = \begin{bmatrix} W_{\mathcal{I}\mathcal{I}} & W_{\mathcal{I}\mathcal{I}^c} \\ W_{\mathcal{I}^c\mathcal{I}} & W_{\mathcal{I}^c\mathcal{I}^c} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{N \times N}$. For any index sets $\mathcal{I}, \mathcal{J} \subseteq \{1, \ldots, N\}$, the following results can be easily established:

$$\pi_{\mathcal{I}} \circ \pi_{\mathcal{J}} \; = \; \pi_{\mathcal{I} \cap \mathcal{J}} \; = \; \pi_{\mathcal{J}} \circ \pi_{\mathcal{I}}, \tag{10}$$

where $\circ$ denotes the composition of two functions.

Definition 4.1. We call a nonempty set $\mathcal{P} \subseteq \mathbb{R}^N$ *coordinate projection admissible* or simply *CP admissible* if for any $x \in \mathcal{P}$ and any index set $\mathcal{J} \subseteq \mathrm{supp}(x)$, $\pi_{\mathcal{J}}(x) = (x_{\mathcal{J}}, 0) \in \mathcal{P}$, where $\mathcal{J}$ may be the empty set.

Clearly, $\mathcal{P}$ must contain the zero vector (by setting $\mathcal{J} = \emptyset$). An equivalent geometric condition for a CP admissible set is shown in the following lemma.

Lemma 4.1. *$\mathcal{P}$ is CP admissible if and only if $\pi_{\mathcal{I}}(\mathcal{P}) \subseteq \mathcal{P}$ for any index set $\mathcal{I} \subseteq \{1, \ldots, N\}$.*

Proof. "If". Since $\pi_{\mathcal{I}}(\mathcal{P}) \subseteq \mathcal{P}$ for any index set $\mathcal{I}$, we have $\pi_{\mathcal{I}}(x) \in \mathcal{P}$ for any $x \in \mathcal{P}$ and any $\mathcal{I}$. Hence, for any $x \in \mathcal{P}$ and any index set $\mathcal{J} \subseteq \mathrm{supp}(x)$, we have $\pi_{\mathcal{J}}(x) \in \mathcal{P}$. This shows that $\mathcal{P}$ is CP admissible.

"Only If". Suppose $\mathcal{P}$ is CP admissible, and let $\mathcal{I}$ be an arbitrary index set. It suffices to show that $\pi_{\mathcal{I}}(x) \in \mathcal{P}$ for any given $x \in \mathcal{P}$. Toward this end, in view of $\mathcal{I} = (\mathcal{I} \cap \mathrm{supp}(x)) \cup (\mathcal{I} \setminus \mathrm{supp}(x))$ and $x_{\mathcal{I} \setminus \mathrm{supp}(x)} = 0$, we have $\pi_{\mathcal{I}}(x) = (x_{\mathcal{I} \cap \mathrm{supp}(x)}, x_{\mathcal{I} \setminus \mathrm{supp}(x)}, x_{\mathcal{I}^c}) = (x_{\mathcal{I} \cap \mathrm{supp}(x)}, 0, 0) = \pi_{\mathcal{I} \cap \mathrm{supp}(x)}(x) \in \mathcal{P}$, where the last membership is due to the facts that $\mathcal{I} \cap \mathrm{supp}(x) \subseteq \mathrm{supp}(x)$ and that $\mathcal{P}$ is CP admissible. $\square$

Examples of bounded CP admissible sets include $\{x \in \mathbb{R}^N \mid a^T x \leq 1, \text{ and } x \geq 0\}$ for a vector $a \in \mathbb{R}^N_{++}$, and any ℓ_p-ball $\{x \in \mathbb{R}^N \mid \|\|\|x\|_p \leq \varepsilon\}$ with $p > 0$ and $\varepsilon > 0$, and $\mathcal{P} = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_N, b_N]$ where $a_i \leq 0 \leq b_i$ for each i. Examples of unbounded CP admissible sets include $\mathbb{R}^N$, $\mathbb{R}^N_+$, and $\Sigma_K = \{x \in \mathbb{R}^N \mid \|x\|_0 \leq K\}$ for some $K \in \mathbb{N}$, Note that Σ_K and the ℓ_p-ball with $0 < p < 1$ are non-convex. Another example of non-convex CP admissible set is $\mathcal{P} = \mathbb{R}^N_+ \cup \mathbb{R}^N_-$. Further, a CP admissible set may be neither open nor closed, e.g., $\mathcal{P} = [0, 1) \times (-1, 2]$ in $\mathbb{R}^2$.

The following proposition provides a list of important properties of CP admissible sets.

Proposition 4.1. *The following hold:*

(i) The set $\mathcal{P}$ is CP admissible if and only if $\lambda\mathcal{P}$ is CP admissible for any real number $\lambda \neq 0$, and the intersection and union of CP admissible sets are CP admissible;

(ii) The algebraic sum of two CP admissible sets is CP admissible;

(iii) If $\mathcal{P}$ is CP admissible, then for any index set $\mathcal{I}$, $\pi_{\mathcal{I}}(\mathcal{P})$ is also CP admissible;

(iv) If $\mathcal{P}$ is a convex and CP admissible set, then $\dim(\mathcal{P}) = \max\{|\mathit{supp}(x)| : x \in \mathcal{P}\}$.

Proof. (i) This is a direct consequence of the definition of a CP admissible set.

(ii) Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be two CP admissible sets, and z be an arbitrary vector in $\mathcal{P}_1 + \mathcal{P}_2$. Hence, $z = x + y$, where $x \in \mathcal{P}_1$ and $y \in \mathcal{P}_2$. For any index set $\mathcal{I}$, it follows from Lemma 4.1 that $\pi_{\mathcal{I}}(x) \in \mathcal{P}_1$ and $\pi_{\mathcal{I}}(y) \in \mathcal{P}_2$. Therefore, $\pi_{\mathcal{I}}(z) = \pi_{\mathcal{I}}(x) + \pi_{\mathcal{I}}(y) \in \mathcal{P}_1 + \mathcal{P}_2$. By Lemma 4.1 again, we deduce that $\mathcal{P}_1 + \mathcal{P}_2$ is CP admissible.

(iii) Let $\mathcal{P}$ be CP admissible, and $\mathcal{I}$ be an arbitrary but fixed index set. Then for any index set $\mathcal{J}$, we deduce via equation (10) that $\pi_{\mathcal{J}}(\pi_{\mathcal{I}}(\mathcal{P})) = \pi_{\mathcal{I}}(\pi_{\mathcal{J}}(\mathcal{P}))$. Since $\mathcal{P}$ is CP admissible, $\pi_{\mathcal{J}}(\mathcal{P}) \subseteq \mathcal{P}$. Hence, by Lemma 4.1, we have $\pi_{\mathcal{I}}(\pi_{\mathcal{J}}(\mathcal{P})) \subseteq \pi_{\mathcal{I}}(\mathcal{P})$. This shows that $\pi_{\mathcal{I}}(\mathcal{P})$ is CP admissible.

(iv) Suppose $\mathcal{P}$ is a convex and CP admissible set. Let $\widehat{x} \in \mathcal{P}$ be such that $|\text{supp}(\widehat{x})| \geq |\text{supp}(x)$ for all $x \in \mathcal{P}$. We claim that for any $x \in \mathcal{P}$, $\text{supp}(x) \subseteq \text{supp}(\widehat{x})$. Suppose not. Then there exist a point $x' \in \mathcal{P}$ and an index $i \in \text{supp}(x')$ such that $i \notin \text{supp}(\widehat{x})$. Since $\mathcal{P}$ is convex, $z(\lambda) := \lambda x' + (1 - \lambda)\widehat{x} \in \mathcal{P}$ for all $\lambda \in [0, 1]$. However, for all $\lambda > 0$ sufficiently small, $(\text{supp}(\widehat{x}) \cup \{i\}) \subseteq \text{supp}(z(\lambda))$. This shows that $|\text{supp}(z(\lambda))| > |\text{supp}(\widehat{x})|$, leading to a contradiction. Therefore, $\text{supp}(x) \subseteq \text{supp}(\widehat{x})$ for all $x \in \mathcal{P}$. Furthermore, it is known that $\dim(\mathcal{P}) = \dim(\text{aff}(\mathcal{P}))$, where $\text{aff}(\cdot)$ denotes the affine hull of a set. Since $\mathcal{P}$ contains the zero vector, $\text{aff}(\mathcal{P}) = \text{span}(\mathcal{P})$. In view of the claim that $\text{supp}(x) \subseteq \text{supp}(\widehat{x})$ for any $x \in \mathcal{P}$, we deduce that $\dim(\mathcal{P}) = \dim(\text{span}(\mathcal{P})) \leq |\text{supp}(\widehat{x})|$. Letting $p := |\text{supp}(\widehat{x})|$, we assume without loss of generality that $\text{supp}(\widehat{x}) = \{1, \ldots, p\}$. For each $s \in \{1, \ldots, p\}$, let $\widehat{\mathcal{J}}_s := \{1, 2, \ldots, s\}$ and $z^s := (\widehat{x}_{\widehat{\mathcal{J}}_s}, 0)$. Therefore, $z^p = \widehat{x}$. Since $\mathcal{P}$ is CP admissible, each $z^s \in \mathcal{P}$. Besides, $\{z^1, z^2, \ldots, z^p\}$ is linearly independent. Since $\mathcal{P}$ is convex and $\{0, z^1, z^2, \ldots, z^p\}$ is affinely independent, the convex hull of $\{0, z^1, z^2, \ldots, z^p\}$ is a simplex of dimension p and is contained in $\mathcal{P}$. Therefore, it follows from [21, Theorem 2.4] that $\dim(\mathcal{P}) \geq p = |\text{supp}(\widehat{x})|$. Consequently, $\dim(\mathcal{P}) = |\text{supp}(\widehat{x})|$. ☐

Using (iv) of Proposition 4.1, we see that the hyperplane $\mathcal{P} = \{x \in \mathbb{R}^N \mid d^T x = 0\}$ with each $d_i \neq 0$ given in Example 3.1 is *not* CP admissible, since $\dim(\mathcal{P}) = N - 1$ but $\max\{|\text{supp}(x)| : x \in \mathcal{P}\} = N$.

Lemma 4.2. *Let $\mathcal{P}$ be a closed and CP admissible set. Then for any index set $\mathcal{J}$, $\pi_{\mathcal{J}}(\mathcal{P})$ is closed.*

Proof. Fix an index set $\mathcal{J}$. Let (z^k) be a convergent sequence in $\pi_{\mathcal{J}}(\mathcal{P})$ such that $(z^k) \to z^*$. Hence, for each k, $z^k = (z^k_{\mathcal{J}}, z^k_{\mathcal{J}^c}) \in \pi_{\mathcal{J}}(\mathcal{P})$ with $z^k_{\mathcal{J}^c} = 0$. Since (z^k) converges to z^*, we have $z^* = (z^*_{\mathcal{J}}, 0)$ and $(z^k_{\mathcal{J}}) \to z^*_{\mathcal{J}}$. Since $\mathcal{P}$ is CP admissible, $\pi_{\mathcal{J}}(\mathcal{P}) \subseteq \mathcal{P}$ such that $z^k \in \mathcal{P}$ for each k. Further, since $\mathcal{P}$ is closed, we have $z^* \in \mathcal{P}$. Clearly, $\pi_{\mathcal{J}}(z^*) = z^* \in \mathcal{P}$. Hence, $z^* \in \pi_{\mathcal{J}}(\mathcal{P})$. This shows that $\pi_{\mathcal{J}}(\mathcal{P})$ is closed. ☐

Note that the above result may fail when $\mathcal{P}$ is not CP admissible, even if it is closed and convex. For example, consider $\mathcal{P} = \{x = (x_1, x_2) \,|\, x_2 \geq \frac{1}{x_1},\ x_1 > 0\} \subset \mathbb{R}^2$. Clearly, $\mathcal{P}$ is closed and convex but not CP admissible. Letting $\mathcal{J} = \{1\}$, we see that $\pi_{\mathcal{J}}(\mathcal{P}) = \{(x_1, 0) \,|\, x_1 \in (0, \infty)\}$ and thus is not closed.

The following result gives a complete characterization of a closed, convex and CP admissible cone. Particularly, it shows that a closed, convex and CP admissible cone is a Cartesian product of Euclidean spaces and nonnegative or nonpositive orthants.

Proposition 4.2. *Let $\mathcal{C}$ be a closed convex cone in $\mathbb{R}^N$. Then $\mathcal{C}$ is CP admissible if and only if there exist four disjoint index subsets $\mathcal{I}_1$, $\mathcal{I}_+$, $\mathcal{I}_-$, and $\mathcal{I}_0$ (some of which can be empty) whose union is $\{1, \ldots, N\}$ such that $\mathcal{C} = \mathbb{R}^{\mathcal{I}_1} + (\mathbb{R}_+)^{\mathcal{I}_+} + (\mathbb{R}_-)^{\mathcal{I}_-} + \{0\}^{\mathcal{I}_0}$ or equivalently $\mathcal{C} = \mathbb{R}_{\mathcal{I}_1} \times (\mathbb{R}_+)_{\mathcal{I}_+} \times (\mathbb{R}_-)_{\mathcal{I}_-} \times \{0\}_{\mathcal{I}_0}$.*

Proof. "If". Suppose $\mathcal{C} = \mathbb{R}^{\mathcal{I}_1} + \mathbb{R}_+^{\mathcal{I}_+} + \mathbb{R}_-^{\mathcal{I}_-} + \{0\}^{\mathcal{I}_0}$, where the four index sets $\mathcal{I}_1$, $\mathcal{I}_+$, $\mathcal{I}_-$, and $\mathcal{I}_0$ form a disjoint union of $\{1, \ldots, N\}$. It is easy to see that $\mathcal{C}$ is closed and convex and that $\mathbb{R}^{\mathcal{I}_1}$, $\mathbb{R}_+^{\mathcal{I}_+}$, $\mathbb{R}_-^{\mathcal{I}_-}$ and $\{0\}^{\mathcal{I}_0}$ are all CP admissible. By (ii) of Proposition 4.1, $\mathcal{C}$ is also CP admissible.

"Only If". Let $\mathcal{C}$ be a closed convex cone which is CP admissible. For an arbitrary index $i \in \{1, \ldots, N\}$, let $\pi_i(\mathcal{C}) := \{\pi_i(x) \,|\, x \in \mathcal{C}\} \subseteq \mathbb{R}^N$ and $[\pi_i(\mathcal{C})]_i := \{\big(\pi_i(x)\big)_i \,|\, x \in \mathcal{C}\} \subseteq \mathbb{R}$. Since $\mathcal{C}$ is a closed convex cone, it is easy to show via a similar argument for Lemma 4.2 that $[\pi_i(\mathcal{C})]_i$ is a closed convex cone in $\mathbb{R}$. This implies that $[\pi_i(\mathcal{C})]_i$ equals either one of the following (polyhedral) cones in $\mathbb{R}$: $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_-$, or $\{0\}$. Define the index sets $\mathcal{I}_1 := \{i \,|\, [\pi_i(\mathcal{C})]_i = \mathbb{R}\}$, $\mathcal{I}_+ := \{i \,|\, [\pi_i(\mathcal{C})]_i = \mathbb{R}_+\}$, $\mathcal{I}_- := \{i \,|\, [\pi_i(\mathcal{C})]_i = \mathbb{R}_-\}$, and $\mathcal{I}_0 := \{i \,|\, [\pi_i(\mathcal{C})]_i = \{0\}\}$. Clearly, these index sets form a disjoint union of $\{1, \ldots, N\}$. Furthermore, since $\mathcal{C}$ is CP admissible, we have $\mathbb{R}^{\mathcal{I}_1} \subseteq \mathcal{C}$, $\mathbb{R}_+^{\mathcal{I}_+} \subseteq \mathcal{C}$, $\mathbb{R}_-^{\mathcal{I}_-} \subseteq \mathcal{C}$, and $\{0\}^{\mathcal{I}_0} \subseteq \mathcal{C}$. Since $\mathcal{C}$ is a convex cone, $\mathbb{R}^{\mathcal{I}_1} + (\mathbb{R}_+)^{\mathcal{I}_+} + (\mathbb{R}_-)^{\mathcal{I}_-} + \{0\}^{\mathcal{I}_0} \subseteq \mathcal{C}$. Conversely, for any $x \in \mathcal{C}$, it follows from the definition of $[\pi_i(\mathcal{C})]_i$ and the disjoint property of the index sets $\mathcal{I}_1, \mathcal{I}_+, \mathcal{I}_-$ and $\mathcal{I}_0$ that $x \in \mathbb{R}^{\mathcal{I}_1} + (\mathbb{R}_+)^{\mathcal{I}_+} + (\mathbb{R}_-)^{\mathcal{I}_-} + \{0\}^{\mathcal{I}_0}$. This shows that $\mathcal{C} = \mathbb{R}^{\mathcal{I}_1} + (\mathbb{R}_+)^{\mathcal{I}_+} + (\mathbb{R}_-)^{\mathcal{I}_-} + \{0\}^{\mathcal{I}_0}$. $\square$

The next proposition presents a decomposition of a closed, convex and CP admissible set.

Proposition 4.3. *Let $\mathcal{P} \subseteq \mathbb{R}^N$ be closed, convex and CP admissible. Then $\mathcal{P} = \mathcal{W} + \mathcal{K}$, where $\mathcal{W} \subseteq \mathcal{P}$ is a compact, convex and CP admissible set, and $\mathcal{K} \subseteq \mathcal{P}$ is a closed, convex and CP admissible cone.*

Proof. For a given closed, convex and CP admissible set $\mathcal{P}$, we first construct a compact, convex and CP admissible set $\mathcal{W}$ contained in $\mathcal{P}$. It follows from the similar argument for Lemma 4.2 and Proposition 4.2 that for each $i \in \{1, \ldots, N\}$, $[\pi_i(\mathcal{P})]_i$ is a closed convex set in $\mathbb{R}$ which contains 0 . Hence, each $[\pi_i(\mathcal{P})]_i$ must be in one of the following forms: $\mathbb{R}$, $[a_i, \infty)$ with $a_i \leq 0$, $(-\infty, b_i]$ with $b_i \geq 0$, and $[a_i, b_i]$ with $a_i \leq 0 \leq b_i$, where in the last case, $a_i = b_i = 0$ if $a_i = b_i$. These four forms respectively correspond to an unbounded set without lower and upper bounds, an unbounded set that is bounded from below, an unbounded set that is bounded frow above, and a bounded set. Define the following disjoint index sets whose union is $\{1, \ldots, N\}$:

$$\begin{aligned} \mathcal{I}_1 &:= \{i \,|\, [\pi_i(\mathcal{P})]_i = \mathbb{R}\}, & \mathcal{I}_+ &:= \{i \,|\, [\pi_i(\mathcal{P})]_i \text{ is unbounded but bounded from below } \}, \\ \mathcal{I}_0 &:= \{i \,|\, [\pi_i(\mathcal{P})]_i \text{ is bounded}\}, & \mathcal{I}_- &:= \{i \,|\, [\pi_i(\mathcal{P})]_i \text{ is unbounded but bounded from above } \}. \end{aligned}$$

Define the closed convex cone $\mathcal{K} := \mathbb{R}^{\mathcal{I}_1} + (\mathbb{R}_+)^{\mathcal{I}_+} + (\mathbb{R}_-)^{\mathcal{I}_-} + \{0\}^{\mathcal{I}_0}$. Since $\mathcal{P}$ is CP admissible and convex, we have $\mathcal{K} \subseteq \mathcal{P}$. Further, $\mathcal{K}$ is CP admissible in view of Proposition 4.2. Moreover, define the set

$$\mathcal{W} := \mathcal{P} \cap \underbrace{\Big\{ x = (x_{\mathcal{I}_1}, x_{\mathcal{I}_+}, x_{\mathcal{I}_-}, x_{\mathcal{I}_0}) \,|\, x_{\mathcal{I}_1} = 0,\ x_{\mathcal{I}_+} \leq 0,\ x_{\mathcal{I}_-} \geq 0 \Big\}}_{:= \mathcal{C}}. \tag{11}$$

Clearly, $\mathcal{W} \subseteq \mathcal{P}$. Since the set $\mathcal{C}$ defined in (11) is closed and convex, $\mathcal{W}$ is also closed and convex. We show next that $\mathcal{W}$ is bounded and CP admissible. To proved the boundedness of $\mathcal{W}$, recall that (i) for each $i \in \mathcal{I}_+$, $[\pi_i(\mathcal{P})]_i = [a_i, \infty)$ for some $a_i \leq 0$; (ii) for each $i \in \mathcal{I}_-$, $[\pi_i(\mathcal{P})]_i = (-\infty, b_i]$ for some $b_i \geq 0$; and (iii) for each $i \in \mathcal{I}_0$, $[\pi_i(\mathcal{P})]_i = [a_i, b_i]$ for some $a_i \le 0 \le b_i$. Hence, $\pi_i(\mathcal{W}) = \{0\}$ for each $i \in \mathcal{I}_1$, $\pi_i(\mathcal{W}) \in [a_i, 0]$ for each $i \in \mathcal{I}_+$, $\pi_i(\mathcal{W}) \in [0, b_i]$ for each $i \in \mathcal{I}_-$, and $\pi_i(\mathcal{W}) \in [a_i, b_i]$ for each $i \in \mathcal{I}_0$. Therefore, for each $x \in \mathcal{W}$, we have $\|x\|_1 = \|x_{\mathcal{I}_+}\|_1 + \|x_{\mathcal{I}_-}\|_1 + \|x_{\mathcal{I}_0}\|_1 \le \sum_{i\in\mathcal{I}_+} |a_i| + \sum_{i\in\mathcal{I}_-} |b_i| + \sum_{i\in\mathcal{I}_0} \max(|a_i|, b_i)$. This shows that $\mathcal{W}$ is bounded and thus compact. Lastly, it is easy to see that the set $\mathcal{C}$ defined in (11) is CP admissible. Since $\mathcal{P}$ is CP admissible, by statement (i) of Proposition 4.1, $\mathcal{W}$ is also CP admissible.

We show that $\mathcal{P} = \mathcal{W} + \mathcal{K}$ as follows. We first show that $\mathcal{W} + \mathcal{K} \subseteq \mathcal{P}$. Consider an arbitrary $z \in \mathcal{W} + \mathcal{K}$, i.e., $z = x + y$ with $x \in \mathcal{W}$ and $y \in \mathcal{K}$. Since $\mathcal{W}$ and $\mathcal{K}$ are both contained in the convex set $\mathcal{P}$ and since $\mathcal{K}$ is a cone, we see that for any $\lambda \in [0, 1)$,

$$\lambda x + y = \lambda x + (1-\lambda)\frac{y}{1-\lambda} \in \mathcal{P}.$$

Furthermore, since $\mathcal{P}$ is closed, $x + y = \lim_{\lambda\uparrow 1} \big(\lambda x + y\big) \in \mathcal{P}$. This shows that $z \in \mathcal{P}$ and thus $\mathcal{W} + \mathcal{K} \subseteq \mathcal{P}$. We finally show that $\mathcal{P} \subseteq \mathcal{W} + \mathcal{K}$. Toward this end, consider an arbitrary $z = (z_1, \dots, z_N)^T \in \mathcal{P}$, and define the vectors $x = (x_1, \dots, x_N)^T$ and $y = (y_1, \dots, y_N)^T$ as follows:

$$x_i := \begin{cases} 0 & \text{if} \quad i \in \mathcal{I}_1 \\ -(z_i)_- & \text{if} \quad i \in \mathcal{I}_+ \\ (z_i)_+ & \text{if} \quad i \in \mathcal{I}_- \\ z_i & \text{if} \quad i \in \mathcal{I}_0 \end{cases}, \qquad y_i := \begin{cases} z_i & \text{if} \quad i \in \mathcal{I}_1 \\ (z_i)_+ & \text{if} \quad i \in \mathcal{I}_+ \\ -(z_i)_- & \text{if} \quad i \in \mathcal{I}_- \\ 0 & \text{if} \quad i \in \mathcal{I}_0 \end{cases}.$$

Clearly, $z = x + y$, $y \in \mathcal{K}$, and $x \in \mathcal{C}$, where $\mathcal{C}$ is defined in (11). Moreover, letting the index set $\mathcal{J} := \{i \in \mathcal{I}_+ \,|\, z_i < 0\} \cup \{i \in \mathcal{I}_- \,|\, z_i > 0\} \cup \mathcal{I}_0$, we have $x = \pi_{\mathcal{J}}(z)$. Since $\mathcal{P}$ is CP admissible, it follows from Lemma 4.1 that $x \in \mathcal{P}$, leading to $x \in \mathcal{W}$. This shows that $z \in \mathcal{W} + \mathcal{K}$, and thus $\mathcal{P} \subseteq \mathcal{W} + \mathcal{K}$. $\square$

The above proposition shows that $\mathcal{K}$ is the asymptotic cone (or recession cone) of $\mathcal{P}$. Furthermore, by using this proposition, we show the existence of an optimal solution of the underlying minimization problem given in Line 7 of Algorithm 1 for an arbitrary index set $\mathcal{J}$ as follows.

Corollary 4.1. *Let $\mathcal{P} \subseteq \mathbb{R}^N$ be a closed, convex and CP admissible set. Then for any matrix $A \in \mathbb{R}^{m\times N}$, any index set $\mathcal{J} \subseteq \{1, \dots, N\}$, and any $y \in \mathbb{R}^m$, $\min_{w\in\mathcal{P}, supp(w)\subseteq\mathcal{J}} \|Aw - y\|_2^2$ attains an optimal solution.*

Proof. We first show that $A\mathcal{P}$ is a closed set for any matrix $A \in \mathbb{R}^{m\times N}$. It follows from Proposition 4.3 that $A\mathcal{P} = A\mathcal{W} + A\mathcal{K}$, where $\mathcal{W}$ is compact and $\mathcal{K}$ is a polyhedral cone. Note that $A\mathcal{W}$ is compact, and $A\mathcal{K}$ is a polyhedral cone and thus is closed. This implies that $A\mathcal{P}$ is closed. The desired result thus follows readily from statement (i) of Lemma 2.1. $\square$

In what follows, we let cone($\mathcal{U}$) denote the conic hull of a nonempty set $\mathcal{U}$ in $\mathbb{R}^N$, i.e., cone($\mathcal{U}$) is the collection of all nonnegative combinations of finitely many vectors in $\mathcal{U}$.

Proposition 4.4. *Let $\mathcal{P}$ be a closed, convex and CP admissible set in $\mathbb{R}^N$. Then cone($\mathcal{P}$) $= \{\lambda x \,|\, \lambda \ge 0, x \in \mathcal{P}\}$, and cone($\mathcal{P}$) is a closed, convex and CP admissible cone.*

Proof. Since $\mathcal{P}$ is a convex set, it follows from a standard argument in convex analysis, e.g., [21, Corollary 2.6.3], that cone($\mathcal{P}$) $= \{\lambda x \,|\, \lambda \ge 0, x \in \mathcal{P}\}$. Define the disjoint index sets whose union is $\{1, \dots, N\}$:

$$\begin{aligned} \mathcal{L}_1 &:= \{i \,|\, 0 \text{ is in the interior of } [\pi_i(\mathcal{P})]_i\}, \qquad \mathcal{L}_0 := \{i \,|\, [\pi_i(\mathcal{P})]_i = \{0\}\}, \\ \mathcal{L}_+ &:= \{i \,|\, \inf[\pi_i(\mathcal{P})]_i = 0, \ \text{and } [\pi_i(\mathcal{P})]_i \text{ contains a positive number }\}, \\ \mathcal{L}_- &:= \{i \,|\, \sup[\pi_i(\mathcal{P})]_i = 0, \ \text{and } [\pi_i(\mathcal{P})]_i \text{ contains a negative number }\}. \end{aligned} \tag{12}$$

Let $\mathcal{C} := \mathbb{R}^{\mathcal{L}_1} + (\mathbb{R}_+)^{\mathcal{L}_+} + (\mathbb{R}_-)^{\mathcal{L}_-} + \{0\}^{\mathcal{L}_0}$. In view of Proposition 4.2, $\mathcal{C}$ is a closed, convex and CP admissible cone. In what follows, we show that $\mathcal{C} = $ cone($\mathcal{P}$) in two steps.

(i) We first show that $\text{cone}(\mathcal{P}) \subseteq \mathcal{C}$. For a given $x \in \mathcal{P}$, we write it as $x = (x_{\mathcal{L}_1}, x_{\mathcal{L}_0}, x_{\mathcal{L}_+}, x_{\mathcal{L}_-})$. Hence, $x = \pi_{\mathcal{L}_1}(x) + \pi_{\mathcal{L}_+}(x) + \pi_{\mathcal{L}_-}(x) + \pi_{\mathcal{L}_0}(x)$, where $\pi_{\mathcal{L}_1}(x) \in \mathbb{R}^{\mathcal{L}_1}$, $\pi_{\mathcal{L}_+}(x) \in (\mathbb{R}_+)^{\mathcal{L}_+}$, $\pi_{\mathcal{L}_-}(x) \in (\mathbb{R}_-)^{\mathcal{L}_-}$, and $\pi_{\mathcal{L}_0}(x) = 0 \in \{0\}^{\mathcal{L}_0}$. By the definition of $\mathcal{C}$, we have that $x \in \mathcal{C}$. Therefore, $\mathcal{P} \subseteq \mathcal{C}$. Since $\text{cone}(\mathcal{P})$ is the smallest convex cone containing $\mathcal{P}$, we have $\text{cone}(\mathcal{P}) \subseteq \mathcal{C}$.

(ii) We next show that $\mathcal{C} \subseteq \text{cone}(\mathcal{P})$. Consider a vector $x \in \mathbb{R}^{\mathcal{L}_1}$, where $x = (x_{\mathcal{L}_1}, x_{\mathcal{L}_1^c}) = (x_{\mathcal{L}_1}, 0)$. By the definition of the index set $\mathcal{L}_1$ given in (12), we see that there exists a sufficiently small positive number λ such that $\lambda x_i \in [\pi_i(\mathcal{P})]_i$ for each $i \in \mathcal{L}_1$. Let $v = (v_{\mathcal{L}_1}, v_{\mathcal{L}_1^c})$ with $v_{\mathcal{L}_1} := \lambda x_{\mathcal{L}_1}$ and $v_{\mathcal{L}_1^c} := 0$. Hence, $v \in \pi_{\mathcal{L}_1}(\mathcal{P})$. Since $\mathcal{P}$ is CP admissible, $\pi_{\mathcal{L}_1}(\mathcal{P}) \subseteq \mathcal{P}$ such that $v \in \mathcal{P}$. In view of $x = (1/\lambda)v$ and $\text{cone}(\mathcal{P}) = \{\lambda x \,|\, \lambda \geq 0, x \in \mathcal{P}\}$, we deduce that $x \in \text{cone}(\mathcal{P})$. Therefore, $\mathbb{R}^{\mathcal{L}_1} \subseteq \text{cone}(\mathcal{P})$. It follows from a similar argument that $\mathbb{R}_+^{\mathcal{L}_+} \subseteq \text{cone}(\mathcal{P})$, $\mathbb{R}_-^{\mathcal{L}_-} \subseteq \text{cone}(\mathcal{P})$, and $\{0\}^{\mathcal{L}_0} \subseteq \text{cone}(\mathcal{P})$. Since $\text{cone}(\mathcal{P})$ is convex, we see that $\mathbb{R}^{\mathcal{L}_1} + \mathbb{R}_+^{\mathcal{L}_+} + \mathbb{R}_-^{\mathcal{L}_-} + \{0\}^{\mathcal{L}_0} \subseteq \text{cone}(\mathcal{P})$. Hence, $\mathcal{C} \subseteq \text{cone}(\mathcal{P})$.

Consequently, $\mathcal{C} = \text{cone}(\mathcal{P})$. Finally, since $\mathcal{C}$ is closed and CP admissible, so is $\text{cone}(\mathcal{P})$. □

Note that if $\mathcal{P}$ is not CP admissible (even though closed and convex), its conic hull may *not* be closed in general. An example is the closed unit ℓ_2-ball in $\mathbb{R}^N$ centered at $\mathbf{e}_1 \in \mathbb{R}^N$.

Definition 4.2. A closed, convex and CP admissible set $\mathcal{P}$ is *irreducible* if the index set $\{i \,|\, [\pi_i(\mathcal{P})]_i = \{0\}\}$ is the empty set.

In light of Proposition 4.4, it is easy to see that a closed, convex and CP admissible set $\mathcal{P}$ is irreducible if and only if $\text{cone}(\mathcal{P})$ is irreducible.

The above development shows that the class of CP admissible sets enjoy favorable properties indicated at the beginning of this section. For example, each CP admissible set contains sufficiently many sparse vectors due to the CP admissible property. Moreover, $\mathbb{R}^N$, $\mathbb{R}^N_+$ and their alikes belong to the class of CP admissible sets. In what follows, we show an additional important implication of CP admissible sets in Proposition 4.5, which is crucial to the development of sufficient conditions for uniform exact recovery in Section 6. To this end, we first present a technical result on the support of vectors.

Lemma 4.3. *Let $u, v \in \mathbb{R}^N$ and $\mathcal{J} \subseteq \{1, \ldots, N\}$ be such that $supp(v) \subseteq \mathcal{J} \subseteq supp(u)$. Then $supp(u - v) \setminus \mathcal{J} = supp(u) \setminus \mathcal{J}$.*

Proof. We show $\text{supp}(u - v) \subseteq \text{supp}(u)$ first. Let $i \in \text{supp}(u - v)$. Hence, $u_i - v_i \neq 0$. We claim that $u_i \neq 0$, because otherwise, $u_i = 0$ and $v_i \neq 0$, which implies $i \in \text{supp}(v) \subseteq \text{supp}(u)$, yielding a contradiction. Hence, $\text{supp}(u - v) \subseteq \text{supp}(u)$. This leads to $\text{supp}(u - v) \setminus \mathcal{J} \subseteq \text{supp}(u) \setminus \mathcal{J}$. Conversely, for any $i \in \text{supp}(u) \setminus \mathcal{J}$, we have $v_i = 0$ (due to $\text{supp}(v) \subseteq \mathcal{J}$) so that $(u - v)_i = u_i \neq 0$. Hence, $i \in \text{supp}(u - v)$. Since $i \notin \mathcal{J}$, we have $i \in \text{supp}(u - v) \setminus \mathcal{J}$. Therefore, $\text{supp}(u) \setminus \mathcal{J} \subseteq \text{supp}(u - v) \setminus \mathcal{J}$. As a result, $\text{supp}(u - v) \setminus \mathcal{J} = \text{supp}(u) \setminus \mathcal{J}$. □

Proposition 4.5. *Let $\mathcal{P}$ be a closed, convex and CP admissible set in $\mathbb{R}^N$. Given a matrix $A \in \mathbb{R}^{m \times N}$, a vector $0 \neq u \in \Sigma_K \cap \mathcal{P}$, and any index set $\mathcal{J} \subset supp(u)$, let v be an arbitrary solution to $\mathbf{Q}$: $\min_{w \in \mathcal{P},\, supp(w) \subseteq \mathcal{J}} \|A(w - u)\|_2^2$. Then the following hold:*

$$\sum_{j \in supp(u-v) \cap \mathcal{J}} \langle A(u - v), A_{\bullet j} \rangle \cdot (u - v)_j \;\leq\; 0,$$

and

$$\|A(u - v)\|_2^2 \;\leq \sum_{j \in supp(u) \setminus \mathcal{J}} \langle A(u - v), A_{\bullet j} \rangle \cdot (u - v)_j.$$

Proof. Note that such an optimal solution v exists due to Corollary 4.1. Define the convex function $g(z) := \|A_{\bullet \mathcal{J}} z - Au\|_2^2$ with $z \in \mathbb{R}^{|\mathcal{J}|}$, and the constraint set $\mathcal{W} := \{z \,|\, (z, 0) \in \pi_{\mathcal{J}}(\mathcal{P})\}$. It follows from

Lemma 4.2 that $\pi_{\mathcal{J}}(\mathcal{P})$ is closed. Since $\mathcal{P}$ is convex, so is $\pi_{\mathcal{J}}(\mathcal{P})$. Hence, $\pi_{\mathcal{J}}(\mathcal{P})$ is closed and convex. This shows that $\mathcal{W}$ is also a closed convex set. Moreover, the underlying optimization problem **Q** can be equivalently formulated as the convex optimization problem: $\min_{z\in\mathcal{W}} g(z)$. Therefore, the optimal solution $v = (v_{\mathcal{J}}, 0)$ satisfies the necessary and sufficient optimality condition given by the following variational inequality: $\langle \nabla g(v_{\mathcal{J}}), z - v_{\mathcal{J}}\rangle \geq 0$ for all $z \in \mathcal{W}$. Since $\mathcal{P}$ is CP admissible, we have $(u_{\mathcal{J}}, 0) \in \mathcal{P}$ so that $u_{\mathcal{J}} \in \mathcal{W}$. In view of $\nabla g(v_{\mathcal{J}}) = A^T_{\bullet\mathcal{J}}(A_{\bullet\mathcal{J}} v_{\mathcal{J}} - Au) = A^T_{\bullet\mathcal{J}}(Av - Au)$, we have

$$0 \leq \langle A^T_{\bullet\mathcal{J}}(A_{\bullet\mathcal{J}}\, v_{\mathcal{J}} - Au), u_{\mathcal{J}} - v_{\mathcal{J}}\rangle = \langle Av - Au, A_{\bullet\mathcal{J}}(u-v)_{\mathcal{J}}\rangle.$$

This implies that $\langle A(u-v), A_{\bullet\mathcal{J}}(u-v)_{\mathcal{J}}\rangle \leq 0$. Consequently, we obtain

$$\begin{aligned}
&\sum_{j\in \mathrm{supp}(u-v)\cap\mathcal{J}} \langle A(u-v), A_{\bullet j}\rangle \cdot (u-v)_j \\
&= \sum_{j\in \mathrm{supp}(u-v)\cap\mathcal{J}} \langle A(u-v), A_{\bullet j}\rangle \cdot (u-v)_j + \sum_{j\in [\mathrm{supp}(u-v)]^c\cap\mathcal{J}} \langle A(u-v), A_{\bullet j}\rangle \cdot (u-v)_j \\
&= \sum_{j\in\mathcal{J}} \langle A(u-v), A_{\bullet j}\rangle \cdot (u-v)_j \;=\; \langle A(u-v), A_{\bullet\mathcal{J}}(u-v)_{\mathcal{J}}\rangle \\
&\leq \; 0.
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
\|A(u-v)\|_2^2 &= \sum_{j=1}^{N}\langle A(u-v), A_{\bullet j}(u-v)_j\rangle = \sum_{j\in\mathrm{supp}(u-v)} \langle A(u-v), A_{\bullet j}\rangle\cdot(u-v)_j \\
&= \sum_{j\in\mathrm{supp}(u-v)\setminus\mathcal{J}} \langle A(u-v), A_{\bullet j}\rangle\cdot(u-v)_j + \sum_{j\in\mathrm{supp}(u-v)\cap\mathcal{J}} \langle A(u-v), A_{\bullet j}\rangle\cdot(u-v)_j \\
&\leq \sum_{j\in\mathrm{supp}(u-v)\setminus\mathcal{J}} \langle A(u-v), A_{\bullet j}\rangle\cdot(u-v)_j \\
&= \sum_{j\in\mathrm{supp}(u)\setminus\mathcal{J}} \langle A(u-v), A_{\bullet j}\rangle\cdot(u-v)_j,
\end{aligned}$$

where the last equation follows from Lemma 4.3. □

5 Exact Vector Recovery on Closed, Convex, CP Admissible Cones for a Fixed Support via Constrained Matching Pursuit

This section is focused on the exact vector recovery on closed, convex and CP admissible cones for a fixed support. By Proposition 4.2, such a cone is a Cartesian product of copies of $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{R}_-$, which includes $\mathbb{R}^N$ and $\mathbb{R}^N_+$. It is shown in Section 6 that closed, convex and CP admissible cones play an important role in characterizing exact recovery, even for general closed, convex and CP admissible sets (cf. Section 6.2). We first introduce the definition of exact vector recovery.

Definition 5.1. Let a matrix $A \in \mathbb{R}^{m\times N}$ and a constraint set $\mathcal{P}$ be given. For a fixed vector $z \in \Sigma_K \cap \mathcal{P}$, we say that *the exact vector recovery* of z is achieved from $y = Az$ via Algorithm 1 if (i) the exact support recovery of z is achieved, and (ii) along *any* sequence $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ for the given z, once $\mathcal{J}_s = \mathrm{supp}(z)$ is reached, then the minimization problem given by Line 7 of Algorithm 1 has a *unique* solution $x^s = z$. If the exact vector recovery of each $z \in \Sigma_K \cap \mathcal{P}$ is achieved, then we call *the exact vector recovery on* $\Sigma_K \cap \mathcal{P}$ (or simply the exact vector recovery) is achieved via Algorithm 1. We also say that a matrix A

achieves the exact vector (resp. support) recovery on $\mathcal{P}$ if the exact vector (resp. support) recovery on $\Sigma_K \cap \mathcal{P}$ is achieved using A. Besides, for a fixed index set $\mathcal{S}$, we say that *the exact vector recovery on* $\mathcal{P}$ *for* $\mathcal{S}$ is achieved if the exact vector recovery of any $z \in \mathcal{P}$ with $\text{supp}(z) = \mathcal{S}$ is achieved.

5.1 Revisit of Exact Vector Recovery on $\mathbb{R}^N$ for a Fixed Support via OMP: A Counterexample to a Necessary Exact Recovery Condition in the Literature

When the sparse recovery problem (1) is constraint free, i.e., $\mathcal{P} = \mathbb{R}^N$, the constrained matching pursuit scheme given by Algorithm 1 reduces to the OMP [20]. The OMP has been extensively studied in the signal processing and compressed sensing literature, and many results have been developed for support or vector recovery using the OMP [10, 16]. In particular, "necessary" and sufficient conditions are established in [10, Proposition 3.5] for exact vector recovery via the OMP for a fixed support; the same "necessary" and sufficient conditions are also given by Tropp [26, Theorems 3.1 and 3.10]. For the sake of completeness and the ease of the subsequent discussions, we present the real version of [10, Proposition 3.5] as follows, i.e., $A \in \mathbb{R}^{m\times N}$, $y \in \mathbb{R}^m$, and $x \in \mathbb{R}^N$, using slightly modified wording.

Proposition 5.1. *[10, Proposition 3.5] Given a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns, every nonzero vector $x \in \mathbb{R}^N$ supported on a given index set $\mathcal{S}$ of size s (i.e., $supp(x) = \mathcal{S}$ and $|supp(x)| = s$) is recovered from $y = Ax$ after at most s iterations of OMP if and only if the following two conditions hold:*

(i) The matrix $A_{\bullet\mathcal{S}}$ is injective (i.e., $A_{\bullet\mathcal{S}}$ has full column rank), and

(ii)

$$\max_{j\in\mathcal{S}} \big|(A^T Az)_j\big| > \max_{j\in\mathcal{S}^c} \big|(A^T Az)_j\big|, \qquad \forall\, 0 \neq z \in \mathbb{R}^N \ \ \textit{with} \ \ supp(z) \subseteq \mathcal{S}. \tag{13}$$

Further, under condition (i), condition (13) holds if and only if

$$\big\|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\big\|_1 < 1, \tag{14}$$

where $\|\cdot\|_1$ denotes the matrix 1-norm.

The "proof" of this proposition can be found on page 68 of the well received monograph [10] by Foucart and Rauhut, and its equivalent condition (14) in term of the matrix 1-norm follows from [10, Remark 3.6]. Also see a similar sufficiency proof in [26, Theorem 3.1] and a "necessity" proof in [26, Theorem 3.10], where condition (14) is referred to as the *exact recovery condition* coined by Tropp in [26]. Clearly, conditions (i) and (ii) are sufficient for the exact vector recovery. Further, condition (i) is necessary for the exact vector recovery. However, we find that condition (ii) only *partially* holds for the necessity of the exact vector recovery. Specifically, condition (ii) is necessary when the index set $\mathcal{S}$ satisfies $|\mathcal{S}| = 1$ or $|\mathcal{S}| = 2$; when $|\mathcal{S}| = 3$, we construct a nontrivial counterexample (i.e., a matrix A) such that any nonzero vector $x \in \mathbb{R}^N$ with $\text{supp}(x) = \mathcal{S}$ is exactly recovered via the OMP using the matrix A but this A does not satisfy (13) or its equivalence (14).

The construction of our counterexample is motivated by an unsuccessful attempt to justify the following implication, which is the last key step given in the necessity proof for [10, Proposition 3.5]:

$$\left[\max_{j\in\mathcal{S}} \big|(A^T Az)_j\big| > \max_{j\in\mathcal{S}^c} \big|(A^T Az)_j\big|, \quad \forall\, 0 \neq z \in \mathbb{R}^N \ \text{ with } \ \text{supp}(z) = \mathcal{S}\right] \Longrightarrow$$
$$\left[\max_{j\in\mathcal{S}} \big|(A^T Az)_j\big| > \max_{j\in\mathcal{S}^c} \big|(A^T Az)_j\big|, \quad \forall\, 0 \neq z \in \mathbb{R}^N \ \text{ with } \ \text{supp}(z) \subseteq \mathcal{S}\right], \tag{15}$$

where we assume that the exact vector recovery is achieved and $A_{\bullet\mathcal{S}}$ has full column rank. Note that the hypothesis of the implication given by (15) holds since it follows from the first step of the OMP using A. To elaborate an underlying reason for the failure of this implication, we define the function $q(z) :=$

$\max_{j\in\mathcal{S}} |(A^T Az)_j| - \max_{j\in\mathcal{S}^c} |(A^T Az)_j|$ for $z \in \mathbb{R}^N$ and the set $\mathcal{R} := \{z \in \mathbb{R}^N \,|\, z \neq 0,\ \mathrm{supp}(z) = \mathcal{S}\}$. Clearly, $q(\cdot)$ is continuous. Further, any nonzero $\widetilde{z} \in \mathbb{R}^N$ with $\mathrm{supp}(\widetilde{z}) \subset \mathcal{S}$ is on the boundary of $\mathcal{R}$ such that there exists a sequence (z_k) in $\mathcal{R}$ converging to $\widetilde{z}$. Hence, the sequence $(q(z_k))$ converges to $q(\widetilde{z})$, where each $q(z_k) > 0$ in view of the hypothesis of the implication (15). However, one can only conclude that $q(\widetilde{z}) \geq 0$ instead $q(\widetilde{z}) > 0$. The counterexample we construct shows that when $|\mathcal{S}| = 3$, there exists a matrix A achieving the exact vector recovery via the OMP but the corresponding $q(\widetilde{z}) = 0$ for some $0 \neq \widetilde{z} \in \mathbb{R}^N$ with $\mathrm{supp}(\widetilde{z}) \subset \mathcal{S}$; see Remark 5.1 for details. This example invalidates the implication (15).

A similar argument also explains the failure of Tropp's necessity proof in [26, Theorem 3.10]. In fact, the (nonzero) signal $\mathbf{s}_{bad}$ constructed in that proof is shown to satisfy $\rho(\mathbf{s}_{bad}) \geq 1$, which is equivalent to $q(\mathbf{s}_{bad}) \leq 0$. However, if $\mathrm{supp}(\mathbf{s}_{bad})$ is a proper subset of the index set Λ_{opt}, which is equivalent to the index set $\mathcal{S}$ defined above, then the argument based on the first step of the OMP used in the proof for [26, Theorem 3.10] becomes invalid. In fact, the counterexample we construct shows that when $|\mathcal{S}| = 3$, there exists a matrix A achieving the exact vector recovery via the OMP but a nonzero $\widetilde{z}$ with $\mathrm{supp}(\widetilde{z}) \subset \mathcal{S}$ exists such that the corresponding $q(\widetilde{z}) = 0$ or equivalently $\rho(\widetilde{z}) = 1$. See Remark 5.1 for details. [1]

We introduce more assumptions and notation through the rest of the development in this section. Consider a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns, i.e., $\|A_{\bullet i}\|_2 = 1$ for each $i = 1, \ldots, N$. Define $\vartheta_{ij} := \langle A_{\bullet i}, A_{\bullet j}\rangle$ for $i, j \in \{1, \ldots, N\}$, and for each i, define the function

$$g_i(z) := \big|\langle A_{\bullet i}, Az\rangle\big| = \Big|\sum_{j=1}^{N} \vartheta_{ij} z_j\Big|, \qquad \forall\, z = (z_1, \ldots, z_N)^T \in \mathbb{R}^N. \tag{16}$$

5.1.1 Positive Necessity Results and Their Implications

This subsection presents certain cases where condition (14) (or equivalently (13)) is indeed necessary for the exact vector recovery for a given support $\mathcal{S}$. The first result shows that [10, Proposition 3.5] (or Proposition 5.1 of the present paper) holds when the index set $\mathcal{S}$ is of size 1 or 2.

Theorem 5.1. *For a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns and an index set $\mathcal{S}$ with $|\mathcal{S}| = 1$ or $|\mathcal{S}| = 2$, the exact vector recovery of every nonzero vector $x \in \mathbb{R}^N$ with supp$(x) = \mathcal{S}$ is achieved from $y = Ax$ via the OMP if and only if the conditions (i) and (ii) in Proposition 5.1 hold.*

Proof. In light of the prior discussions and the argument for [10, Proposition 3.5], we only need to show that the implication (15) holds when A achieves the exact vector recovery via the OMP and $A_{\bullet\mathcal{S}}$ has full column rank. The case of $|\mathcal{S}| = 1$ is trivial, and we focus on the case of $|\mathcal{S}| = 2$ as follows. Without loss of generality, let $\mathcal{S} = \{1, 2\}$. In view of g_i's defined in (16), it suffices to show that if $\max(g_1(z), g_2(z)) > \max_{i\geq 3} g_i(z), \forall\, z$ with $\mathrm{supp}(z) = \{1, 2\}$, then $\max(g_1(z), g_2(z)) > \max_{i\geq 3} g_i(z), \forall\, z$ with $\mathrm{supp}(z) = \{1\}$ or $\mathrm{supp}(z) = \{2\}$. Since $A_{\bullet\mathcal{S}}$ has full column rank, the 2×2 matrix $A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}} = \begin{bmatrix} 1 & \vartheta_{12} \\ \vartheta_{12} & 1 \end{bmatrix}$ is positive definite. Hence, $|\vartheta_{12}| < 1$. For any z with $\mathrm{supp}(z) = \{1\}$, we have $\max(g_1(z), g_2(z)) = \max(|z_1|, |\vartheta_{12} z_1|) = g_1(z) > g_2(z)$ because $z_1 \neq 0$ and $|\vartheta_{12}| < 1$. Similarly, $\max(g_1(z), g_2(z)) = g_2(z) > g_1(z)$ when $\mathrm{supp}(z) = \{2\}$.

In what follows, we consider an arbitrary z^* with $\mathrm{supp}(z^*) = \{1\}$ first. Note that $g_j(z^*) = |\vartheta_{j1} z_1^*|$ for each j, where $z_1^* \neq 0$. Since z^* is on the boundary of $\mathcal{R} := \{z \in \mathbb{R}^N \,|\, \mathrm{supp}(z) = \{1, 2\}\}$ on which $\max(g_1(z), g_2(z)) > \max_{i\geq 3} g_i(z)$, we deduce via the continuity of g_i's that $g_1(z^*) = \max(g_1(z^*), g_2(z^*)) \geq g_i(z^*)$ for each $i \geq 3$. We show next that $g_1(z^*) > g_i(z^*)$ for all $i \geq 3$ by contradiction. Suppose, in contrast, $g_1(z^*) = g_i(z^*)$ for some $i \geq 3$, i.e., $|z_1^*| = |\vartheta_{i1} z_1^*| = \gamma$. For any $v \in \mathbb{R}^N$ with $\mathrm{supp}(v) = \{1, 2\}$ and $\|v\|_2 > 0$ sufficiently small, $\max(g_1(z^* + v), g_2(z^* + v)) = g_1(z^* + v)$ due to $g_1(z^*) > g_2(z^*)$, and $z^* + v \in \mathcal{R}$ so that $g_1(z^* + v) > g_i(z^* + v)$. Therefore, we have

$$|z_1^* + p^T v_{\mathcal{S}}| > |\vartheta_{i1} z_1^* + q^T v_{\mathcal{S}}|, \tag{17}$$

[1]In a private communication, Dr. Joel A. Tropp pointed out to the authors that this issue may be related to the borderline case indicated in Footnote 2 in his paper [26].

where $p = (1, \vartheta_{12})^T$, $q = (\vartheta_{i1}, \vartheta_{i2})^T$, and $v_{\mathcal{S}} = (v_1, v_2)^T \in \mathbb{R}^2$. Letting $\gamma := |z_1^*| > 0$, we obtain four possible cases from $|z_1^*| = |\vartheta_{i1} z_1^*|$: (i) $(z_1^*, \vartheta_{i1} z_1^*) = (\gamma, \gamma)$; (ii) $(z_1^*, \vartheta_{i1} z_1^*) = (\gamma, -\gamma)$; (iii) $(z_1^*, \vartheta_{i1} z_1^*) = (-\gamma, \gamma)$; and (iv) $(z_1^*, \vartheta_{i1} z_1^*) = (-\gamma, -\gamma)$. In each of these cases, it follows from (17) that $(\mathrm{sgn}(z_1^*) \cdot p - \mathrm{sgn}(\vartheta_{i1} z_1^*) \cdot q)^T v_{\mathcal{S}} > 0$ for all $\|v_{\mathcal{S}}\| > 0$ sufficiently small, where $\mathrm{sgn}(\cdot)$ is the signum function. In view of $\mathrm{supp}(v_{\mathcal{S}}) = \mathrm{supp}(-v_{\mathcal{S}})$, we have $(\mathrm{sgn}(z_1^*) \cdot p - \mathrm{sgn}(\vartheta_{i1} z_1^*) \cdot q)^T v_{\mathcal{S}} > 0$ and $(\mathrm{sgn}(z_1^*) \cdot p - \mathrm{sgn}(\vartheta_{i1} z_1^*) \cdot q)^T (-v_{\mathcal{S}}) > 0$ for all $\|v_{\mathcal{S}}\|_2 > 0$ sufficiently small. This yields a contradiction. Hence, $\max(g_1(z^*), g_2(z^*)) > g_i(z^*)$ for all $i \geq 3$ when $\mathrm{supp}(z^*) = \{1\}$. The case of $\mathrm{supp}(z^*) = \{2\}$ also follows by interchanging the roles of g_1 and g_2. Consequently, the implication (15) holds, which leads to condition (ii) in Proposition 5.1. $\square$

By leveraging the necessary and sufficient recovery conditions in Theorem 5.1 for a fixed support of size 2, we show that condition (**H**) is necessary for the exact vector or support recovery on Σ_2.

Corollary 5.1. *Let $A \in \mathbb{R}^{m \times N}$ have unit columns. Then A achieves the exact vector recovery on Σ_2 if and only if (i) condition* (**H**) *holds on Σ_2, and (ii) any two distinct columns of A are linearly independent.*

Proof. "If". In view of Proposition 3.2, condition (**H**) yields the exact support recovery on Σ_2. Besides, condition (ii) guarantees that the unique x^2 equals z for any $z \in \Sigma_2$ with $|\mathrm{supp}(z)| = 2$. It also ensures that the unique $x^1 = z$ for any $z \in \Sigma_2$ with $|\mathrm{supp}(z)| = 1$. This yields the exact vector recovery on Σ_2.

"Only if". Suppose A achieves the exact vector recovery on Σ_2. Clearly, condition (ii) is necessary as shown before. To show that condition (i) is also necessary, consider a vector $z \in \Sigma_2$ with $|\mathrm{supp}(z)| = 2$. Without loss of generality, we assume that $\mathrm{supp}(z) = \{1, 2\}$. Since A achieves the exact vector recovery on Σ_2, it must achieve the exact support recovery for the fixed support $\mathcal{S} = \{1, 2\}$. Hence it follows from Theorem 5.1 that $\|(A_{\bullet \mathcal{S}}^T A_{\bullet \mathcal{S}})^{-1} A_{\bullet \mathcal{S}}^T A_{\bullet \mathcal{S}^c}\|_1 < 1$, which is equivalent to

$$1 - \vartheta_{12}^2 > \max_{j \in \mathcal{S}^c} \big(|\vartheta_{j1} - \vartheta_{j2} \vartheta_{12}| + |\vartheta_{j2} - \vartheta_{j1} \vartheta_{12}| \big). \tag{18}$$

Consider the three proper subsets of $\mathrm{supp}(z) = \{1, 2\}$, i.e., $\mathcal{J} = \emptyset$, $\mathcal{J} = \{1\}$, and $\mathcal{J} = \{2\}$. When $\mathcal{J} = \emptyset$, the inequality (9) holds for $u = z$ and $v = 0$ in light of $\max_{j \in \mathcal{S}} \big|(A^T A z)_j\big| > \max_{j \in \mathcal{S}^c} \big|(A^T A z)_j\big|$ obtained from the first step of the OMP. Moreover, we have either $|z_1 + \vartheta_{12} z_2| \geq |\vartheta_{12} z_1 + z_2|$ or $|z_1 + \vartheta_{12} z_2| \leq |\vartheta_{12} z_1 + z_2|$. For the former case, we deduce from the exact support recovery of z via the OMP that $j_1^* = 1$ and $\mathcal{J}_1 = \{1\}$ such that $x^1 = (A_{\bullet 1}^T A z) \mathbf{e}_1$ is the unique optimal solution to $\min_{\mathrm{supp}(w) \subseteq \mathcal{J}_1} \|A(z - w)\|_2^2$. Hence, by Corollary 3.1, the exact support recovery shows that $f_2^*(z, x^1) < \min_{j \in \mathcal{S}^c} f_j^*(z, x^1)$, leading to the inequality (9) for $u = z$ and $v = x^1$ when $\mathcal{J} = \{1\}$. We then consider $\mathcal{J} = \{2\}$. In this case, the unique optimal solution v^* to $\min_{\mathrm{supp}(w) \subseteq \mathcal{J}} \|A(z - w)\|_2^2$ is given by $v^* = (A_{\bullet 2}^T A z) \mathbf{e}_2 = (\vartheta_{12} z_1 + z_2) \mathbf{e}_2$. Therefore, $A_{\bullet j}^T A(z - v^*) = (\vartheta_{j1} - \vartheta_{j2} \vartheta_{12}) z_1$ for any j. We thus have $|A_{\bullet 1}^T A(z - v^*)| = |1 - \vartheta_{12}^2| \cdot |z_1|$ and $|A_{\bullet j}^T A(z - v^*)| = |\vartheta_{j1} - \vartheta_{12} \vartheta_{j2}| \cdot |z_1|$, where $z_1 \neq 0$. Noting that $f_1^*(z, v^*) < \min_{j \in \mathcal{S}^c} f_j^*(z, v^*)$ if and only if $|A_{\bullet 1}^T A(z - v^*)| > \max_{j \in \mathcal{S}^c} |A_{\bullet j}^T A(z - v^*)|$, we deduce via the above results and (18) that $f_1^*(z, v^*) < \min_{j \in \mathcal{S}^c} f_j^*(z, v^*)$, leading to the inequality (9) for $u = z$ and $v = v^*$ when $\mathcal{J} = \{2\}$. The other case where $|z_1 + \vartheta_{12} z_2| \leq |\vartheta_{12} z_1 + z_2|$ can be established in a similar way. Further, for any $u \in \Sigma_2$ with $|\mathrm{supp}(u)| = 1$ and $\mathcal{J} = \emptyset$, the inequality (9) also holds. Thus condition (**H**) holds on Σ_2. $\square$

The next result shows that even though condition (14) (or equivalently (13)) may fail to be necessary, it is necessary for *almost all* the matrices achieving the exact vector recovery associated with a fixed support $\mathcal{S}$. This result also illustrates the challenge of constructing a counterexample. Toward this end, let $\mathcal{U}$ be the set of all matrices in $\mathbb{R}^{m \times N}$ with unit columns, i.e., $\mathcal{U} := \big\{ A \in \mathbb{R}^{m \times N} \mid \|A_{\bullet i}\|_2 = 1, \ \forall\, i = 1, \ldots, N \big\}$. Note that $\mathcal{U}$ is the Cartesian product of N copies of unit ℓ_2-spheres in $\mathbb{R}^m$. Hence, $\mathcal{U}$ is a compact manifold of dimension $(m-1)N$, and it attains a (finite) positive measure with a Lebesgue measure μ on $\mathcal{U}$. For a fixed index set $\mathcal{S}$, define the set

$$\mathcal{D} := \big\{ A \in \mathcal{U} \mid \ A \text{ achieves the exact vector recovery for the given support } \mathcal{S} \ \big\}.$$

Recall that for any $A \in \mathcal{D}$, $A_{\bullet \mathcal{S}}$ has full column rank.

Proposition 5.2. *Let the set $\mathcal{D}' := \{A \in \mathcal{D} \,|\, \|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\|_1 = 1\}$, and μ be a Lebesgue measure on $\mathcal{U}$. Then $\mu(\mathcal{D}) > 0$ and $\mu(\mathcal{D}') = 0$.*

Proof. For a given matrix $A \in \mathcal{D}$, we recall the function $q(z) := \max_{j\in\mathcal{S}} |(A^T Az)_j| - \max_{j\in\mathcal{S}^c} |(A^T Az)_j|$ for $z \in \mathbb{R}^N$ and the set $\mathcal{R} := \{z \in \mathbb{R}^N \,|\, z \neq 0,\ \text{supp}(z) = \mathcal{S}\}$ given below (15). Since A achieves the exact vector recovery for the given support $\mathcal{S}$, we have $q(z) > 0$ for all $z \in \mathcal{R}$. Moreover, it follows from the discussioins below (15) that $q(\widetilde{z}) \geq 0$ for any nonzero $\widetilde{z} \in \mathbb{R}^N$ with $\text{supp}(\widetilde{z}) \subset \mathcal{S}$. By a similar argument for [10, Remark 3.6], we have $\|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\|_1 \leq 1$ for any $A \in \mathcal{D}$.

Define the set $\mathcal{D}'' := \{A \in \mathcal{D} \,|\, \|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\|_1 < 1\}$. In view of the above argument, we see that $\mathcal{D}$ is the disjoint union of $\mathcal{D}'$ and $\mathcal{D}''$. Since $\mathcal{D}''$ is a (relatively) open subset in $\mathcal{U}$, we deduce that $\mu(\mathcal{D}'') > 0$. Therefore, $\mu(\mathcal{D}) \geq \mu(\mathcal{D}'') > 0$. Moreover, define

$$\widetilde{\mathcal{D}} := \Big\{ A \in \mathcal{U} \,\big|\ A_{\bullet\mathcal{S}} \text{ has full column rank, and } \|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\|_1 = 1 \Big\},$$
$$\mathcal{W}_j := \Big\{ A \in \mathcal{U} \,\big|\ A_{\bullet\mathcal{S}} \text{ has full column rank, and } \big\| \big[(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\big]_{\bullet j} \big\|_1 = 1 \Big\}, \quad j = 1, \ldots, |\mathcal{S}^c|.$$

Hence, $\mathcal{D}' \subseteq \widetilde{\mathcal{D}} \subset \bigcup_{j=1}^{|\mathcal{S}^c|} \mathcal{W}_j$. Let $\mathbf{a} \in \mathbb{R}^{mN}$ be the vectorization of $A \in \mathbb{R}^{m\times N}$, i.e., $\mathbf{a}$ is generated by stacking the columns of A on top of one another. For each fixed j, $\big\| \big[(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\big]_{\bullet j} \big\|_1 = 1$ holds if and only if the piecewise polynomial function $G_j(\mathbf{a}) = 0$, where $G_j(\mathbf{a}) := \sum_{k=1}^{|\mathcal{S}|} |G_{j,k}(\mathbf{a})| - G_{j,k+1}(\mathbf{a})$, and each $G_{j,k}(\cdot) : \mathbb{R}^{mN} \to \mathbb{R}$ is a polynomial function. In view of this result, it is easy to verify that $\mathcal{W}_j$ is a subset of a finite union of the sets of the following form: $\big\{ A \in \mathcal{U} \,|\ A_{\bullet\mathcal{S}}$ has full column rank, and $H_s(\mathbf{a}) = 0 \big\}$, where $H_s(\cdot)$ is a (nonzero) polynomial function. Clearly, each set of this form is a lower dimensional submanifold of $\mathcal{U}$ and thus is of zero measure. Thus $\mu(\mathcal{W}_j) = 0$ for each j, and we thus have $\mu(\mathcal{D}') = 0$. ☐

5.1.2 Construction of a Counterexample for a Fixed Support of Size 3

In this subsection, we construct a nontrivial counterexample to show that condition (14) (or equivalently (13)) fails to be necessary. The main result is given by the following theorem.

Theorem 5.2. *For an index set $\mathcal{S}$ with $|\mathcal{S}| = 3$, there exists an $A \in \mathbb{R}^{4\times 4}$ with unit columns such that A achieves exact vector recovery for the fixed support $\mathcal{S}$ via the OMP, $A_{\bullet\mathcal{S}}$ has full column rank, and $\big\|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\big\|_1 = 1$.*

To construct such a matrix A indicated in the above theorem, we first present some preliminary results. Without loss of generality, let $\mathcal{S} = \{1, 2, 3\}$ and $\mathcal{S}^c = \{4\}$. In view of the function g_i's defined in (16), we introduce the following functions for $i = 1, \ldots, 4$:

$$\widehat{g}_i(v) \ := \ \big|h_i^T v\big|, \ \forall\, v \in \mathbb{R}^3, \qquad \text{where} \qquad h_i := \big(\vartheta_{i1}, \vartheta_{i2}, \vartheta_{i3}\big)^T \in \mathbb{R}^3,$$

where we recall that $\vartheta_{ij} = \langle A_{\bullet i}, A_{\bullet j}\rangle$ for $i, j \in \{1, \ldots, 4\}$. Hence, $\max_{j\in\mathcal{S}} \big|(A^T Az)_j\big| > \max_{j\in\mathcal{S}^c} \big|(A^T Az)_j\big|$ for all $0 \neq z \in \mathbb{R}^N$ with $\text{supp}(z) = \mathcal{S}$ if and only if the following holds:

$$(\mathbf{P}): \quad \max_{i=1,2,3} \widehat{g}_i(v) \ > \ \widehat{g}_4(v), \quad \forall\, v = (v_1, v_2, v_3)^T \in \mathbb{R}^3 \ \text{ with } v_1 \cdot v_2 \cdot v_3 \neq 0.$$

Since each $\widehat{g}_i$ and $\max_{i=1,2,3} \widehat{g}_i(v)$ are convex piecewise affine functions [17], it is not surprising that the feasibility of $(\mathbf{P})$ can be characterized by that of certain linear inequalities. The following lemma gives a necessary and sufficient condition for $(\mathbf{P})$ in term of the feasibility of some linear inequalities.

Lemma 5.1. *Let the matrix $H := \begin{bmatrix} h_4 + h_1 & h_4 - h_1 & h_4 + h_2 & h_4 - h_2 & h_4 + h_3 & h_4 - h_3 \end{bmatrix} \in \mathbb{R}^{3\times 6}$. Then $\max_{i=1,2,3} \widehat{g}_i(v) > \widehat{g}_4(v)$ for all $v = (v_1, v_2, v_3)^T \in \mathbb{R}^3$ with $v_1 \cdot v_2 \cdot v_3 \neq 0$ holds if and only if for each $\sigma := (\sigma_1, \sigma_2, \sigma_3) \in \{(\pm 1, \pm 1, \pm 1)\}$, there exist vectors $0 \neq u \geq 0$ and $w \geq 0$ such that $u + D_\sigma H w = 0$, where the diagonal matrix $D_\sigma := diag(\sigma_1, \sigma_2, \sigma_3) \in \mathbb{R}^{3\times 3}$.*

Proof. Note that (**P**) fails if and only if there exists $\widehat{v} \in \mathbb{R}^3$ with $\widehat{v}_1 \cdot \widehat{v}_2 \cdot \widehat{v}_3 \neq 0$ such that $\widehat{g}_i(\widehat{v}) \leq \widehat{g}_4(\widehat{v})$ for each $i = 1,2,3$. We claim that the latter statement holds if and only if there exists $v^* \in \mathbb{R}^3$ with $v_1^* \cdot v_2^* \cdot v_3^* \neq 0$ such that $|h_i^T v^*| \leq h_4^T v^*$ for each $i = 1,2,3$. The "if" part is obvious since $h_4^T v^* \leq |h_4^T v^*| = \widehat{g}_4(v^*)$. To show the "only if" part, we let $v^* = \operatorname{sgn}(h_4^T \widehat{v}) \cdot \widehat{v}$, where $\widehat{v}$ satisfies the specified conditions. In view of $\widehat{g}_i(v^*) = |h_i^T v^*| = |h_i^T \widehat{v}|$ for $i = 1,2,3$, $h_4^T v^* = |h_4^T \widehat{v}| = \widehat{g}_4(\widehat{v})$, and each $v_i^* \neq 0$, we conclude that the desired result holds. This completes the proof of the claim.

By using the above claim, we see that (**P**) fails if and only if there exists $v^* \in \mathbb{R}^3$ with $v_1^* \cdot v_2^* \cdot v_3^* \neq 0$ such that $|h_i^T v^*| \leq h_4^T v^*$ for each $i = 1,2,3$, where the latter is further equivalent to $\pm h_i^T v^* \leq h_4^T v^*$ for each $i = 1,2,3$ or equivalently $H^T v^* \geq 0$. Therefore, (**P**) fails if and only if there exist $\sigma \in \{(\pm1, \pm1, \pm1)\}$ and $\widetilde{v} \in \mathbb{R}^3_{++}$ (i.e., $\widetilde{v}_i > 0$ for each $i = 1,2,3$) such that $H^T D_\sigma \widetilde{v} \geq 0$. By virtue of the Motzkin's Transposition Theorem, we see that for a fixed σ, the linear inequality system $(D_\sigma H)^T \widetilde{v} \geq 0, \widetilde{v} > 0$ has a solution $\widetilde{v}$ if and only if the linear inequality system $u + D_\sigma H w = 0, 0 \neq u \geq 0$ and $w \geq 0$ has no solution (u, w). In other words, (**P**) fails if and only if there exist $\sigma \in \{(\pm1, \pm1, \pm1)\}$ such that the linear inequality system $u + D_\sigma H w = 0$, $0 \neq u \geq 0$, and $w \geq 0$ has no solution. As a result, (**P**) holds if and only if for any $\sigma \in \{(\pm1, \pm1, \pm1)\}$, there exist vectors $0 \neq u \geq 0$ and $w \geq 0$ such that $u + D_\sigma H w = 0$. □

By making use of the above preliminary results, we prove Theorem 5.2 as follows.

Proof of Theorem 5.2. Consider the matrix

$$A = \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ 0 & \frac{2\sqrt{2}}{3} & -\frac{\sqrt{2}}{3} & \frac{\sqrt{2}}{3} \\ 0 & 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{6}}{12} \\ 0 & 0 & 0 & \frac{\sqrt{10}}{4} \end{bmatrix} \in \mathbb{R}^{4\times4}. \tag{19}$$

Recall that $\mathcal{S} = \{1,2,3\}$ and $\mathcal{S}^c = \{4\}$. It is easy to verify that A is invertible with unit columns (i.e., $\|A_{\bullet i}\|_2 = 1$ for each i), $A_{\bullet\mathcal{S}}$ has full column rank, and

$$\vartheta_{12} = \vartheta_{21} = \vartheta_{13} = \vartheta_{31} = \vartheta_{23} = \vartheta_{32} = -\frac{1}{3}, \quad \vartheta_{41} = \vartheta_{42} = \frac{1}{3}, \quad \vartheta_{43} = -\frac{1}{2}. \tag{20}$$

Hence, $A_{\bullet\mathcal{S}}^T A_{\bullet\mathcal{S}} = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} \in \mathbb{R}^{3\times3}$ and $A_{\bullet\mathcal{S}}^T A_{\bullet\mathcal{S}^c} = h_4$, where

$$h_1 = \begin{bmatrix} 1 \\ -\frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}, \quad h_2 = \begin{bmatrix} -\frac{1}{3} \\ 1 \\ -\frac{1}{3} \end{bmatrix}, \quad h_3 = \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ 1 \end{bmatrix}, \quad h_4 = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ -\frac{1}{2} \end{bmatrix}.$$

Furthermore,

$$\left(A_{\bullet\mathcal{S}}^T A_{\bullet\mathcal{S}}\right)^{-1} = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & 1 & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & 1 \end{bmatrix}^{-1} = \frac{3}{4} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

such that $\left\|(A_{\bullet\mathcal{S}}^T A_{\bullet\mathcal{S}})^{-1} A_{\bullet\mathcal{S}}^T A_{\bullet\mathcal{S}^c}\right\|_1 = \frac{3}{8} + \frac{3}{8} + \frac{1}{4} = 1$. The rest of the proof consists of two parts: the first part shows that $\max_{j\in\mathcal{S}} \left|(A^T A z)_j\right| > \max_{j\in\mathcal{S}^c} \left|(A^T A z)_j\right|$ for all z with $\operatorname{supp}(z) = \mathcal{S}$, and the second part shows that A achieves the exact vector recovery for the index set $\mathcal{S}$.

We first show the following claim:

$$\text{Claim I:} \quad \max_{j\in\mathcal{S}} \left|(A^T A z)_j\right| > \max_{j\in\mathcal{S}^c} \left|(A^T A z)_j\right| \text{ for all } z \text{ with } \operatorname{supp}(z) = \mathcal{S}. \tag{21}$$

In view of Lemma 5.1, we only need to show that for each $\sigma \in \{(\pm1, \pm1, \pm1)\}$, there exist vectors $0 \neq u \geq 0$ and $w \geq 0$ such that $u + D_\sigma H w = 0$, where the matrix

$$H = \begin{bmatrix} h_4 + h_1 & h_4 - h_1 & h_4 + h_2 & h_4 - h_2 & h_4 + h_3 & h_4 - h_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} & -\frac{2}{3} & 0 & \frac{2}{3} & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & \frac{4}{3} & -\frac{2}{3} & 0 & \frac{2}{3} \\ -\frac{5}{6} & -\frac{1}{6} & -\frac{5}{6} & -\frac{1}{6} & \frac{1}{2} & -\frac{3}{2} \end{bmatrix}.$$

Toward this end, we give a specific solution (u, w) to the above linear inequality system for each σ:

(1) $\sigma = (1, 1, 1)$. A solution is given by $u = -(H_{\bullet 2} + H_{\bullet 4}) = (0, 0, \frac{1}{3})^T$ and $w = \mathbf{e}_2 + \mathbf{e}_4$;

(2) $\sigma = (1, 1, -1)$. A solution is given by $u = H_{\bullet 5} = (0, 0, \frac{1}{2})^T$ and $w = \mathbf{e}_5$;

(3) $\sigma = (1, -1, 1)$. A solution is given by $u = (\frac{2}{3}, \frac{2}{3}, \frac{1}{6})^T$ and $w = \mathbf{e}_2$;

(4) $\sigma = (-1, 1, 1)$. A solution is given by $u = (\frac{2}{3}, \frac{2}{3}, \frac{1}{6})^T$ and $w = \mathbf{e}_4$;

(5) $\sigma = (1, -1, -1)$. A solution is given by $u = H_{\bullet 5} = (0, 0, \frac{1}{2})^T$ and $w = \mathbf{e}_5$;

(6) $\sigma = (-1, 1, -1)$. A solution is given by $u = H_{\bullet 5} = (0, 0, \frac{1}{2})^T$ and $w = \mathbf{e}_5$;

(7) $\sigma = (-1, -1, 1)$. A solution is given by $u = (\frac{3}{4}, 0, \frac{5}{6})^T$ and $w = \mathbf{e}_1$;

(8) $\sigma = (-1, 1, -1)$. A solution is given by $u = H_{\bullet 5} = (0, 0, \frac{1}{2})^T$ and $w = \mathbf{e}_5$.

Hence, Claim I holds in light of Lemma 5.1.

We show next that the matrix A achieves the exact vector recovery via the OMP for the given index set $\mathcal{S} = \{1, 2, 3\}$. Let z be an arbitrary vector in $\mathbb{R}^4$ with $\text{supp}(z) = \mathcal{S}$, and $y = Az = A_{\bullet \mathcal{S}} z_{\mathcal{S}}$. Consider the following three steps of the OMP:

• Step 1: Since $x^0 = 0$ and $y = Az$, it follows from (21) that $\max_{i=1,2,3} |A_{\bullet i}^T A(z - x^0)| > |A_{\bullet 4}^T A(z - x^0)|$. Hence, by Corollary 3.1, $j_1^* \in \mathcal{S} = \{1, 2, 3\}$ and $\mathcal{J}_1 = \{j_1^*\}$. Also, $x^1 := \arg\min_{\text{supp}(w) \subseteq \mathcal{J}_1} \|y - Aw\|_2^2$ is given by $x^1 = (A_{\bullet j_1^*}^T A_{\bullet \mathcal{S}} z_{\mathcal{S}}) \cdot \mathbf{e}_{j_1^*}$. Note that $x_{j_1^*}^1 \neq 0$ in view of Proposition 3.1.

• Step 2: We first prove the following claim: for any $j_1 \in \mathcal{S} = \{1, 2, 3\}$ and $u = (A_{\bullet j_1}^T Az) \cdot \mathbf{e}_{j_1} \in \mathbb{R}^4$, $\max_{i=1,2,3} |A_{\bullet i}^T A(z - u)| > |A_{\bullet 4}^T A(z - u)|$.

Proof of the above claim. For any $j_1 \in \{1, 2, 3\}$ and its corresponding u, let $v := z - u$. Note that $v_i = z_i \neq 0$ for each $i \in \mathcal{S} \setminus \{j_1\}$. Therefore, if $z_{j_1} \neq A_{\bullet j_1}^T Az$, then $\text{supp}(v) = \mathcal{S}$ so that the claim holds by virtue of (21). To handle the case where $z_{j_1} = A_{\bullet j_1}^T Az$, we consider $j_1 = 1$ first. Since $A_{\bullet 1}^T Az = A_{\bullet 1}^T A_{\bullet \mathcal{S}} z_{\mathcal{S}} = h_1^T z_{\mathcal{S}} = z_1 - \frac{1}{3} z_2 - \frac{1}{3} z_3$, we must have $z_3 = -z_2 \neq 0$. Therefore, $v_{\mathcal{S}} = z_{\mathcal{S}} - u_{\mathcal{S}} = z_2 \cdot (0, 1, -1)^T$. It follows from $\widehat{g}_i(v_{\mathcal{S}}) = |h_i^T v_{\mathcal{S}}|$ and h_i's given before that $\widehat{g}_1(v_{\mathcal{S}}) = 0$, $\widehat{g}_2(v_{\mathcal{S}}) = \widehat{g}_3(v_{\mathcal{S}}) = \frac{4}{3}|z_2|$, and $\widehat{g}_4(v_{\mathcal{S}}) = \frac{5}{6}|z_2|$. Consequently, $\max_{i=1,2,3} |A_{\bullet i}^T A(z - u)| = \max_{i=1,2,3} \widehat{g}_i(v_{\mathcal{S}}) > \widehat{g}_4(v_{\mathcal{S}}) = |A_{\bullet 4}^T A(z - u)|$. Due to the symmetry of the matrix A, it can be shown via a similar argument that the above result also holds for $z_{j_1} = A_{\bullet j_1}^T Az$ with $j_1 = 2$ or $j_1 = 3$. This completes the proof of the claim. $\square$

By the above claim, we see that $\max_{i=1,2,3} |A_{\bullet i}^T A(z - x^1)| > |A_{\bullet 4}^T A(z - x^1)|$ for the vector x^1 obtained from Step 1. Therefore, $j_2^* \in \mathcal{S}$ and $j_2^* \neq j_1^*$ in view of Lemma 3.1. Hence, $\mathcal{J}_2 = \{j_1^*, j_2^*\}$, and $x^2 := \arg\min_{\text{supp}(w) \subseteq \mathcal{J}_2} \|y - Aw\|_2^2$ is given by $x_{\mathcal{J}_2}^2 = \big(A_{\bullet \mathcal{J}_2}^T A_{\bullet \mathcal{J}_2}\big)^{-1} A_{\bullet \mathcal{J}_2}^T A_{\bullet \mathcal{S}} z_{\mathcal{S}}$, and $x_i^2 = 0$ for $i \notin \mathcal{J}_2$.

• Step 3: Note that for any index set $\mathcal{I} \in \big\{\{1, 2\}, \{1, 3\}, \{2, 3\}\big\}$, it follows from a direct calculation on the matrix A that

$$w = \big(A_{\bullet \mathcal{I}}^T A_{\bullet \mathcal{I}}\big)^{-1} A_{\bullet \mathcal{I}}^T A_{\bullet \mathcal{S}} z_{\mathcal{S}} = \begin{bmatrix} z_s - \frac{1}{2} z_p \\ z_t - \frac{1}{2} z_p \end{bmatrix},$$

where $s, t \in \mathcal{I}$ with $s < t$, and $p \in \mathcal{S} \setminus \mathcal{I}$. Hence, (i) if $\mathcal{J}_2 = \{1, 2\}$, then $(z - x^2)_{\mathcal{S}} = z_3 \cdot (\frac{1}{2}, \frac{1}{2}, 1)^T$; (ii) if $\mathcal{J}_2 = \{1, 3\}$, then $(z - x^2)_{\mathcal{S}} = z_2 \cdot (\frac{1}{2}, 1, \frac{1}{2})^T$; and (iii) if $\mathcal{J}_3 = \{2, 3\}$, then $(z - x^2)_{\mathcal{S}} = z_1 \cdot (1, \frac{1}{2}, \frac{1}{2})^T$. Therefore, for the vector x^2 obtained from Step 2, we have $\text{supp}(z - x^2) = \mathcal{S}$. It follows from (21) that $\max_{i=1,2,3} |A_{\bullet i}^T A(z - x^2)| > |A_{\bullet 4}^T A(z - x^2)|$. This shows that $j_3^* \in \mathcal{S}$ with $j_3^* \notin \mathcal{J}_2$. Hence, $\mathcal{J}_3 = \mathcal{J}_2 \cup \{j_3^*\} = \mathcal{S}$. Since $A_{\bullet \mathcal{S}}$ has full column rank, we see that $x^3 := \arg\min_{\text{supp}(w) \subseteq \mathcal{J}_3} \|y - Aw\|_2^2$ satisfies $x^3 = z$. This shows that z is uniquely recovered via the OMP using the matrix A. $\square$

Remark 5.1. We make a few remarks about the counterexample constructed above.

(a) It is easy to verify that for the given matrix A in (19), when $v = \alpha \cdot (1,1,0)^T$ for any $0 \neq \alpha \in \mathbb{R}$, $\widehat{g}_1(v) = \widehat{g}_2(v) = \widehat{g}_3(v) = \widehat{g}_4(v) = \frac{2}{3}|\alpha|$. Hence, $\max_{i=1,2,3} \widehat{g}_i(v) = \widehat{g}_4(v)$. Letting $z = (z_{\mathcal{S}}, z_{\mathcal{S}^c}) \in \mathbb{R}^4$ with $z_{\mathcal{S}} = v$ and $z_{\mathcal{S}^c} = 0$, we have $\max_{i=1,2,3} g_i(z) = g_4(z)$, leading to a counterexample to the implication (15) used in the necessity proof for [10, Proposition 3.5]. Besides, letting $\widetilde{m} = |\mathcal{S}| = 3$, since A is invertible, all the $\widetilde{m}$-term representations are unique, and the condition $\|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\|_1 = 1$ implies the failure of the "Exact Recovery Condition" defined in Tropp's paper [26] (i.e., $\|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}})^{-1} A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}^c}\|_1 < 1$). However, any z with $\text{supp}(z) = \mathcal{S}$ can be exactly recovered via the OMP, yielding a counterexample to [26, Theorem 3.10].

(b) There are multiple 4×4 real matrices satisfying the conditions specified in Theorem 5.2 as long as their columns are unit and the inner products of their distinct columns defined by ϑ_{ij} equal to the values given in (20). In particular, for the matrix A given in (19) and any orthogonal matrix $P \in \mathbb{R}^{4\times 4}$, PA also satisfies the conditions in Theorem 5.2.

The counterexample constructed in the previous theorem can be extended to one with a larger size.

Corollary 5.2. *Suppose an index set $\mathcal{S} \subseteq \{1, \ldots, N\}$ is of size 3, i.e., $|\mathcal{S}| = 3$. Then for any $m \geq 4$ and $N \geq 4$, there exists a matrix $\widehat{A} \in \mathbb{R}^{m\times N}$ with unit columns such that $\widehat{A}$ achieves the exact vector recovery for the fixed support $\mathcal{S}$ via the OMP, $\widehat{A}_{\bullet\mathcal{S}}$ has full column rank, and $\big\|(\widehat{A}^T_{\bullet\mathcal{S}} \widehat{A}_{\bullet\mathcal{S}})^{-1} \widehat{A}^T_{\bullet\mathcal{S}} \widehat{A}_{\bullet\mathcal{S}^c}\big\|_1 = 1$.*

Proof. Without loss of generality, let $\mathcal{S} = \{1,2,3\}$. For any $N \geq 4$, define the matrix $B \in \mathbb{R}^{4\times N}$ as $B := \begin{bmatrix} A & B_{\bullet 5} & \cdots & \cdots & B_{\bullet N} \end{bmatrix}$, where the matrix A is given in (19), and $B_{\bullet k} = \pm A_{\bullet 4}$ for each $k \geq 5$. Then let $\widehat{A} := \begin{bmatrix} B \\ 0_{(m-4)\times N} \end{bmatrix} \in \mathbb{R}^{m\times N}$. Straightforward calculations show that $\widehat{A}$ satisfies the desired properties by observing that almost all the required properties of $\widehat{A}$ rely on $\langle \widehat{A}_{\bullet i}, \widehat{A}_{\bullet j}\rangle$'s, which are defined by ϑ_{ij}'s or h_i's of the matrix A. $\square$

5.2 Exact Vector Recovery on the Nonnegative Orthant $\mathbb{R}^N_+$ for a Fixed Support

We consider the exact vector recovery on the nonnegative orthant $\mathbb{R}^N_+$ for a fixed support $\mathcal{S}$ using constrained matching pursuit. Without loss of generality, we assume that the matrix $A \in \mathbb{R}^{m\times N}$ has unit columns, i.e., $\|A_{\bullet i}\|_2 = 1$ for each $i = 1, \ldots, N$. A necessary condition is given as follows.

Lemma 5.2. *Given a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns and an index set $\mathcal{S}$ of size s, the exact vector recovery of every nonzero vector $x \in \mathbb{R}^N_+$ with $supp(x) = \mathcal{S}$ is achieved via constrained matching pursuit only if $A_{\bullet\mathcal{S}}$ has full column rank.*

Proof. Assume, in contrast, that $A_{\bullet\mathcal{S}}$ does not have full column rank. Let $r := |\mathcal{S}|$. Then there exist a nonzero vector $v \in \mathbb{R}^r$ such that $A_{\bullet\mathcal{S}} v = 0$. For a given nonzero $x \geq 0$ with $\text{supp}(x) = \mathcal{S}$, suppose at the rth step, the exact support of x is recovered from $y = Ax$ via constrained matching pursuit. It follows from Algorithm 1 that one need to solve the constrained minimization problem $\mathbf{Q}:\ \min_{w\in\mathbb{R}^r_+} \|A_{\bullet\mathcal{S}} w - y\|_2^2$, where $y = A_{\bullet\mathcal{S}} x_{\mathcal{S}}$, to recover $x_{\mathcal{S}}$. Since $x_{\mathcal{S}} > 0$ and $v \neq 0$, there exists a small positive constant ε such that $x_{\mathcal{S}} + \varepsilon v > 0$. Noting that $A_{\mathcal{S}}(x_{\mathcal{S}} + \varepsilon v) = A_{\mathcal{S}} x_{\mathcal{S}} = y$, we see that $x_{\mathcal{S}} + \varepsilon v$ is a solution to the minimization problem $\mathbf{Q}$. Hence, $\mathbf{Q}$ has multiple optimal solutions which can be different from the desired solution $x_{\mathcal{S}}$. This leads to a contradiction. Consequently, $A_{\bullet\mathcal{S}}$ has full column rank. $\square$

In light of statement (ii) of Corollary 3.1 for $x^0 = 0$, we easily obtain another necessary condition for the exact support recovery (and thus exact vector recovery) of any $z \in \mathbb{R}^N_+$ with $\text{supp}(z) = \mathcal{S}$:

$$\max_{j\in \text{supp}(z)} (A^T_{\bullet j} A z)_+ > \max_{j\in [\text{supp}(z)]^c} (A^T_{\bullet j} A z)_+, \qquad \forall\, z \in \mathbb{R}^N_+ \text{ with } \text{supp}(z) = \mathcal{S},$$

which is equivalent to $\|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}} v)_+\|_\infty > \|(A^T_{\bullet\mathcal{S}^c} A_{\bullet\mathcal{S}} v)_+\|_\infty$ for all $v \in \mathbb{R}^{|\mathcal{S}|}_{++}$.

5.2.1 Necessary and Sufficient Conditions for Exact Vector Recovery for a Fixed Support of Size 2

We derive necessary and sufficient conditions for exact vector recovery on $\mathbb{R}^N_+$ for a given support $\mathcal{S}$ with $|\mathcal{S}| = 2$. Recall that $\vartheta_{ij} := \langle A_{\bullet i}, A_{\bullet j} \rangle$ for $i, j \in \{1, \ldots, N\}$. Besides, the following lemma is needed.

Lemma 5.3. *Let $M \in \mathbb{R}^{m\times m}$ be a positive definite matrix. Then for any $z \in \mathbb{R}^m$ with $z > 0$, there exists $i \in \{1, \ldots, m\}$ such that $(Mz)_i > 0$.*

Proof. Suppose, in contrast, that there exists $z > 0$ such that $Mz \le 0$. Since $z > 0$, we have $z^T Mz \le 0$. As M is positive definite, we deduce that $z^T Mz = 0$ so that $z = 0$. This yields a contradiction. $\square$

Theorem 5.3. *Given a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns and the index set $\mathcal{S} = \{1, 2\}$, every nonzero vector $x \in \mathbb{R}^N_+$ with supp$(x) = \mathcal{S}$ is recovered from $y = Ax$ via constrained matching pursuit if and only if the following conditions hold:*

(i) $A_{\bullet\mathcal{S}}$ *has full column rank or equivalently* $|\vartheta_{12}| < 1$*;*

(ii) $\max\big((z_1 + \vartheta_{12}z_2)_+, \, (\vartheta_{12}z_1 + z_2)_+\big) > \max_{j\in\mathcal{S}^c}\big(\vartheta_{j1}z_1 + \vartheta_{j2}z_2\big)_+, \ \forall\, (z_1, z_2)^T \in \mathbb{R}^2_{++}$*;*

(iii) $1 - \vartheta_{12}^2 > \max_{j\in\mathcal{S}^c}\big((\vartheta_{j2} - \vartheta_{12}\vartheta_{j1})_+, \ (\vartheta_{j1} - \vartheta_{12}\vartheta_{j2})_+ \big)$.

Proof. "Only if". Clearly, the condition that $A_{\bullet\mathcal{S}}$ has full column rank is necessary for the exact vector recovery in view of Lemma 5.2. Since $A_{\bullet\mathcal{S}}$ has full column rank if and only if $A^T_{\bullet\mathcal{S}}A_{\bullet\mathcal{S}} = \begin{bmatrix} 1 & \vartheta_{12} \\ \vartheta_{12} & 1 \end{bmatrix}$ is positive definite, we see that $A_{\bullet\mathcal{S}}$ has full column rank if and only if $|\vartheta_{12}| < 1$. For an arbitrary $z \in \mathbb{R}^N$ with $z_{\mathcal{S}} = (z_1, z_2) > 0$, let $y = Az = A_{\bullet\mathcal{S}}z_{\mathcal{S}}$. At Step 1, since $x^0 = 0$, it follows from statement (ii) of Corollary 3.1 that any $j^*_1 \in \mathcal{S}$ if and only if $\max_{j\in\mathcal{S}}\langle A_{\bullet j}, Az\rangle_+ > \max_{j\in\mathcal{S}^c}\langle A_{\bullet j}, Az\rangle_+$. This leads to condition (ii), in light of $\langle A_{\bullet 1}, Az\rangle_+ = (z_1 + \vartheta_{12}z_2)_+$, $\langle A_{\bullet 2}, Az\rangle_+ = (\vartheta_{12}z_1 + z_2)_+$, and $\langle A_{\bullet j}, Az\rangle_+ = (\vartheta_{j1}z_1 + \vartheta_{j2}z_2)_+$. Since $\begin{bmatrix} 1 & \vartheta_{12} \\ \vartheta_{12} & 1 \end{bmatrix}$ is positive definite, it follows from Lemma 5.3 that for any $(z_1, z_2) > 0$, at least one of $\vartheta_{12}z_1 + z_2$ and $\vartheta_{j1}z_1 + \vartheta_{j2}z_2$ is positive. Further, in view of $|\vartheta_{12}| < 1$ and the fact that for $a, b \in \mathbb{R}$, $b_+ > a_+$ if and only if $b > 0$ and $b > a$, it is easy to verify that for any $(z_1, z_2) > 0$, (a) $(z_1 + \vartheta_{12}z_2)_+ > (\vartheta_{12}z_1 + z_2)_+$ if and only if $z_1 > z_2$; (b) $(z_1 + \vartheta_{12}z_2)_+ < (\vartheta_{12}z_1 + z_2)_+$ if and only if $z_1 < z_2$; and (c) $(z_1 + \vartheta_{12}z_2)_+ = (\vartheta_{12}z_1 + z_2)_+ > 0$ if and only if $z_1 = z_2$. Hence, we have that $j^*_1 = 1$ if $z_1 > z_2 > 0$, $j^*_1 = 2$ if $z_2 > z_1 > 0$, and $j^*_1 \in \{1, 2\}$ if $z_1 = z_2 > 0$. Moreover, $\mathcal{J}_1 = \{j^*_1\}$, and $x^1 := \arg\min_{w\ge 0, \mathrm{supp}(w)\subseteq\mathcal{J}_1} \|A_{\bullet\mathcal{S}}z_{\mathcal{S}} - Aw\|_2^2$ is given by $x^1 = \langle A_{\bullet\mathcal{S}}z_{\mathcal{S}}, A_{\bullet j^*_1}\rangle_+ \cdot \mathbf{e}_{j^*_1}$, where $\langle A_{\bullet\mathcal{S}}z_{\mathcal{S}}, A_{\bullet j^*_1}\rangle_+ > 0$ by Proposition 3.1. In what follows, we consider $j^*_1 = 1$ corresponding to $z_1 \ge z_2 > 0$ first. In this case, $x^1 = (z_1 + \vartheta_{12}z_2) \cdot \mathbf{e}_1$. Hence, $(z - x^1)_{\mathcal{S}} = (-\vartheta_{12}, 1)^T \cdot z_2$. It follows from statement (ii) of Corollary 3.1 that a necessary and sufficient condition to select $j^*_2 = 2$ at Step 2 is

$$\langle A(z - x^1), A_{\bullet 2}\rangle_+ > \max_{j\in\mathcal{S}^c}\langle A(z - x^1), A_{\bullet j}\rangle_+, \tag{22}$$

where $\langle A(z - x^1), A_{\bullet 2}\rangle_+ = (1 - \vartheta_{12}^2) \cdot z_2$ and $\langle A(z - x^1), A_{\bullet j}\rangle_+ = (\vartheta_{j2} - \vartheta_{12}\vartheta_{j1})_+ \cdot z_2$ for each $j \in \mathcal{S}^c$. Hence, when $z_1 \ge z_2 > 0$, an equivalent condition for (22) is $1 - \vartheta_{12}^2 > \max_{j\in\mathcal{S}^c}(\vartheta_{j2} - \vartheta_{12}\vartheta_{j1})_+$. When $j^*_1 = 2$ corresponding to $z_2 \ge z_1 > 0$, we deduce via a similar argument that a necessary and sufficient condition for $j^*_2 = 1$ at Step 2 is $1 - \vartheta_{12}^2 > \max_{j\in\mathcal{S}^c}(\vartheta_{j1} - \vartheta_{12}\vartheta_{j2})_+$. This gives rise to condition (iii).

"If". As indicated in the "only if" part, condition (ii) is sufficient for $j^*_1 \in \mathcal{S}$ at Step 1, and condition (iii) is sufficient for $j^*_2 \in \mathcal{S} \setminus \{j^*_1\}$ at Step 2. Hence, under conditions (ii) and (iii), the exact support $\mathcal{S}$ is recovered from $y = Az$ in two steps for any $z \in \mathbb{R}^N$ with $z_{\mathcal{S}} > 0$, i.e., $\mathcal{J}_2 = \mathcal{S}$. Note that the optimality condition for $x^2 := \arg\min_{w\ge 0, \mathrm{supp}(w)\subseteq\mathcal{J}_2} \|A_{\bullet\mathcal{S}}z_{\mathcal{S}} - Aw\|_2^2$ is given by the linear complementarity problem (LCP): $0 \le x^2_{\mathcal{S}} \perp A^T_{\bullet\mathcal{S}}A_{\bullet\mathcal{S}}(x^2_{\mathcal{S}} - z_{\mathcal{S}}) \ge 0$. Since $A_{\bullet\mathcal{S}}$ has full column rank, $A^T_{\bullet\mathcal{S}}A_{\bullet\mathcal{S}}$ is positive definite such that the LCP has a unique solution $x^2_{\mathcal{S}} = z_{\mathcal{S}}$ or equivalently $x^2 = z$. This shows that the exact vector recovery is achieved for any $z \in \mathbb{R}^N$ with $z_{\mathcal{S}} > 0$ under conditions (i)-(iii). $\square$

Applying the necessary and sufficient conditions given in Theorem 5.3, it is shown in the next corollary that condition (**H**) is necessary for the exact vector or support recovery on $\Sigma_2 \cap \mathbb{R}^N_+$.

Corollary 5.3. *Let $A \in \mathbb{R}^{m\times N}$ be a matrix with unit columns. Then the exact vector recovery on $\Sigma_2 \cap \mathbb{R}^N_+$ is achieved if and only if (i) condition (**H**) holds on $\Sigma_2 \cap \mathbb{R}^N_+$, and (ii) any two distinct columns of A are linearly independent.*

Proof. The "if" part is similar to that given in the proof of Corollary 5.1. For the "only if" part, let A achieve the exact vector recovery on $\Sigma_2 \cap \mathbb{R}^N_+$. Clearly, condition (ii) is necessary in light of Lemma 5.2. To show that condition (i) is necessary, we consider an arbitrary $z \in \Sigma_2 \cap \mathbb{R}^N_+$ with $\mathrm{supp}(z) = \{1, 2\} := \mathcal{S}$. Hence, A achieves the exact support recovery for the fixed support $\mathcal{S}$. Therefore, conditions (ii) and (iii) of Theorem 5.3 hold. Consider the three proper subsets of $\mathcal{S}$, i.e., $\mathcal{J} = \emptyset$, $\mathcal{J} = \{1\}$, and $\mathcal{J} = \{2\}$. When $\mathcal{J} = \emptyset$, we see that the inequality (9) holds for $u = z$ and $v = 0$ in light of statement (ii) of Corollary 3.1 and conditions (ii) of Theorem 5.3. Furthermore, we have either $(z_1 + \vartheta_{12}z_2)_+ \geq (\vartheta_{12}z_1 + z_2)_+$ or $(z_1 + \vartheta_{12}z_2)_+ \leq (\vartheta_{12}z_1 + z_2)_+$. For the former case, we deduce from Algorithm 1 that $j^*_1 = 1$ and $\mathcal{J}_1 = \{1\}$ such that $x^1 = (A^T_{\bullet 1}Az)_+\mathbf{e}_1$ is the unique optimal solution to $\min_{w\geq 0, \mathrm{supp}(w)\subseteq \mathcal{J}_1} \|A(z - w)\|^2_2$. Hence, the exact support recovery of z shows that $f^*_2(z, x^1) < \min_{j\in\mathcal{S}^c} f^*_j(z, x^1)$, yielding (9) for $u = z$ and $v = x^1$ when $\mathcal{J} = \{1\}$. We then consider $\mathcal{J} = \{2\}$. Similarly, the unique optimal solution v^* to $\min_{w\geq 0, \mathrm{supp}(w)\subseteq \mathcal{J}} \|A(z - w)\|^2_2$ is given by $v^* = (A^T_{\bullet 2}Az)_+\mathbf{e}_2 = (\vartheta_{12}z_1 + z_2)_+\mathbf{e}_2$. Consider two sub-cases:

(a) $(\vartheta_{12}z_1 + z_2)_+ \leq 0$. In this case, $v^* = 0$ such that $z - v^* = z$. Hence, $(A^T_{\bullet 1}A(z - v^*))_+ = (z_1 + \vartheta_{12}z_2)_+$ and $(A^T_{\bullet j}A(z - v^*))_+ = (\vartheta_{j1}z_1 + \vartheta_{j2}z_2)_+$ for $j \in \mathcal{S}^c$. Since $\max\big((z_1 + \vartheta_{12}z_2)_+, (\vartheta_{12}z_1 + z_2)_+\big) = (z_1 + \vartheta_{12}z_2)_+$, we deduce via condition (ii) of Theorem 5.3 that $f^*_1(z, v^*) < \min_{j\in\mathcal{S}^c} f^*_j(z, v^*)$, yielding the inequality (9) for $u = z$ and $v = v^*$ when $\mathcal{J} = \{2\}$.

(b) $(\vartheta_{12}z_1 + z_2)_+ \geq 0$. In this case, $z - v^* = (1, -\vartheta_{12})z_1$ such that $(A^T_{\bullet 1}A(z - v^*))_+ = (1 - \vartheta^2_{12}) \cdot z_1$ and $(A^T_{\bullet j}A(z - v^*))_+ = (\vartheta_{j1} - \vartheta_{12}\vartheta_{j2})_+ \cdot z_1$ for $j \in \mathcal{S}^c$, where $z_1 > 0$. By condition (iii) of Theorem 5.3 that $f^*_1(z, v^*) < \min_{j\in\mathcal{S}^c} f^*_j(z, v^*)$, yielding (9) for $u = z$ and $v = v^*$ when $\mathcal{J} = \{2\}$.

The other case where $(z_1 + \vartheta_{12}z_2)_+ \leq (\vartheta_{12}z_1 + z_2)_+$ can be established similarly. In addition, for any $u \in \Sigma_2 \cap \mathbb{R}^N_+$ with $|\mathrm{supp}(u)| = 1$ and $\mathcal{J} = \emptyset$, (9) also holds. Hence, condition (**H**) holds on $\Sigma_2 \cap \mathbb{R}^N_+$. $\square$

5.2.2 Necessary and Sufficient Conditions for Exact Vector Recovery for a Fixed Support of Size 3

We first present some preliminary results. Given a (possibly non-square) matrix

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix},$$

where M_{ij}'s are submatrices of M with M_{11} being invertible, the Schur complement of M_{11} in M, denoted by M/M_{11}, is given by $M/M_{11} := M_{22} - M_{21}M^{-1}_{11}M_{12}$. When M is square, the Schur determinant formula says that $\det(M/M_{11}) = \det M/ \det M_{11}$ [6, Proposition 2.3.5]. Particularly, when M is positive definite, any of its Schur complement is also positive definite.

Lemma 5.4. *Given a matrix $A \in \mathbb{R}^{m\times N}$ and an index set $\mathcal{S}$ such that $A_{\bullet\mathcal{S}}$ has full column rank, let the matrix $M := A^T_{\bullet\mathcal{S}}A_{\bullet\mathcal{S}}$. For a nonempty index set $\mathcal{J} \subset \mathcal{S}$ and $z \in \mathbb{R}^N_+$ with supp$(z) = \mathcal{S}$, let x^* be the unique solution to $\min_{w\geq 0, supp(w)\subseteq\mathcal{J}} \|A(z - w)\|^2_2$ whose support is given by $\mathcal{J}^*$, i.e., supp$(x^*) = \mathcal{J}^*$. Define the index set $\mathcal{I} := \mathcal{S} \setminus \mathcal{J}^*$. Then $A^T_{\bullet\mathcal{J}^*}A(z - x^*) = 0$, and*

$$A^T_{\bullet\mathcal{I}}A(z - x^*) = \big(M/M_{\mathcal{J}^*\mathcal{J}^*}\big) \cdot z_{\mathcal{I}}, \quad A^T_{\bullet\mathcal{S}^c}A(z - x^*) = A^T_{\bullet\mathcal{S}^c}\big[I - A_{\bullet\mathcal{J}^*}\big(A^T_{\bullet\mathcal{J}^*}A_{\bullet\mathcal{J}^*}\big)^{-1}A^T_{\bullet\mathcal{J}^*}\big]A_{\bullet\mathcal{I}} \cdot z_{\mathcal{I}}.$$

Moreover, $\max_{j\in\mathcal{I}}[A^T_{\bullet j}A(z - x^*)]_+ = \max_{j\in\mathcal{S}\setminus\mathcal{J}}[A^T_{\bullet j}A(z - x^*)]_+ > 0.$

Proof. Since x^* is the unique optimal solution to $\min_{w\geq 0, \text{supp}(w)\subseteq \mathcal{J}} \|A(z-w)\|_2^2$, we have $x^* = (x^*_{\mathcal{J}}, 0)$, where $x^*_{\mathcal{J}}$ is the solution to $\min_{u\geq 0} \|A_{\bullet\mathcal{J}} u - A_{\bullet\mathcal{S}} z_{\mathcal{S}}\|^2$, and $A_{\bullet\mathcal{J}}$ has full column rank. Therefore, $x^*_{\mathcal{J}}$ is a solution to the linear complementarity problem: $0 \leq u \perp A^T_{\bullet\mathcal{J}} A_{\bullet\mathcal{J}} u - A^T_{\bullet\mathcal{J}} A_{\bullet\mathcal{S}} z_{\mathcal{S}} \geq 0$. In view of $\text{supp}(x^*) = \mathcal{J}^* \subseteq \mathcal{J}$, we deduce that $x^*_{\mathcal{J}^*} = \big(A^T_{\bullet\mathcal{J}^*} A_{\bullet\mathcal{J}^*}\big)^{-1} A^T_{\bullet\mathcal{J}^*} A_{\bullet\mathcal{S}} z_{\mathcal{S}} > 0$. Using $\mathcal{I} = \mathcal{S} \setminus \mathcal{J}^*$ and $A_{\bullet\mathcal{S}} z_{\mathcal{S}} = A_{\bullet\mathcal{J}^*} z_{\mathcal{J}^*} + A_{\bullet\mathcal{I}} z_{\mathcal{I}}$, we further have $x^*_{\mathcal{J}^*} = z_{\mathcal{J}^*} + \big(A^T_{\bullet\mathcal{J}^*} A_{\bullet\mathcal{J}^*}\big)^{-1} A^T_{\bullet\mathcal{J}^*} A_{\bullet\mathcal{I}} z_{\mathcal{I}}$. Hence,

$$A(z-x^*) = A_{\bullet\mathcal{S}}(z_{\mathcal{S}} - x^*_{\mathcal{S}}) = A_{\bullet\mathcal{J}^*}\big(z_{\mathcal{J}^*} - x^*_{\mathcal{J}^*}\big) + A_{\bullet\mathcal{I}} z_{\mathcal{I}} = \big[-A_{\bullet\mathcal{J}^*}\big(A^T_{\bullet\mathcal{J}^*} A_{\bullet\mathcal{J}^*}\big)^{-1} A^T_{\bullet\mathcal{J}^*} A_{\bullet\mathcal{I}} + A_{\bullet\mathcal{I}}\big] z_{\mathcal{I}}.$$

Direct calculations yield $A^T_{\bullet\mathcal{I}} A(z-x^*) = \big[A^T_{\bullet\mathcal{I}} A_{\bullet\mathcal{I}} - A^T_{\bullet\mathcal{I}} A_{\bullet\mathcal{J}^*}\big(A^T_{\bullet\mathcal{J}^*} A_{\bullet\mathcal{J}^*}\big)^{-1} A^T_{\bullet\mathcal{J}^*} A_{\bullet\mathcal{I}}\big] z_{\mathcal{I}} = \big(M/M_{\mathcal{J}^*\mathcal{J}^*}\big) \cdot z_{\mathcal{I}}$; the other equation also follow readily.

Since $M/M_{\mathcal{J}^*\mathcal{J}^*}$ is positive definite and $z_{\mathcal{I}} > 0$, it follows from Lemma 5.3 and the expression for $A^T_{\bullet\mathcal{I}} A(z-x^*)$ derived above that there exists an index $j \in \mathcal{I}$ such that $A^T_{\bullet j} A(z-x^*) > 0$. Hence, $\max_{j\in\mathcal{I}}[A^T_{\bullet j} A(z-x^*)]_+ > 0$. Furthermore, since $\mathcal{I} = \mathcal{S} \setminus \mathcal{J}^*$ and $\mathcal{J}^* \subseteq \mathcal{J} \subset \mathcal{S}$, we have $\mathcal{I} = (\mathcal{S} \setminus \mathcal{J}) \cup (\mathcal{J} \setminus \mathcal{J}^*)$. However, it follows from the linear complementarity condition for $x^*_{\mathcal{J}}$ that $A^T_{\bullet\mathcal{J}} A(x^* - z) = A^T_{\bullet\mathcal{J}} A_{\bullet\mathcal{J}} x^*_{\mathcal{J}} - A^T_{\bullet\mathcal{J}} A_{\bullet\mathcal{S}} z_{\mathcal{S}} \geq 0$, which implies that $A^T_{\bullet\mathcal{J}} A(z-x^*) \leq 0$ or equivalently $[A^T_{\bullet j} A(z-x^*)]_+ = 0$ for all $j \in \mathcal{J}$. Therefore, $\max_{j\in\mathcal{I}}[A^T_{\bullet j} A(z-x^*)]_+ = \max_{j\in\mathcal{S}\setminus\mathcal{J}}[A^T_{\bullet j} A(z-x^*)]_+$. ☐

Lemma 5.5. *Let* $U := \begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix} \in \mathbb{R}^{2\times 2}$ *be a positive definite matrix for real numbers* α, β *and* γ. *Define the set* $\mathcal{W} := \big\{(u_1, u_2) \in \mathbb{R}^2_{++} \,|\, \big(\alpha u_1 + \gamma u_2\big)_+ \geq \big(\gamma u_1 + \beta u_2\big)_+\big\}$. *Then* $\mathcal{W}$ *is nonempty if and only if* $\alpha > \gamma$. *Furthermore, if* $\mathcal{W}$ *is nonempty, then* $\{u_2 \,|\, (u_1, u_2) \in \mathcal{W}\} = \mathbb{R}_{++}$.

Proof. Since U is positive definite, we have $\alpha > 0$, $\beta > 0$, and $\alpha\beta > \gamma^2$. To show the "if" part, suppose $\alpha > \gamma$. Then for a fixed $u_1 > 0$, we have $\alpha u_1 > \gamma u_1$ and $\alpha u_1 > 0$. Therefore, for a sufficiently small $u_2 > 0$, it is easy to see that $(\alpha u_1 + \gamma u_2)_+ \geq (\gamma u_1 + \beta u_2)_+$. This shows that $\mathcal{W}$ is nonempty. To prove the "only if" part, suppose $\mathcal{W}$ is nonempty but $\alpha \leq \gamma$. Note that this implies that $\gamma > 0$. Since $\alpha \cdot \beta > \gamma^2$ (due to the positive definiteness of U), we have $\beta > \dfrac{\gamma}{\alpha} \cdot \gamma \geq \gamma$. Therefore, $\beta > \gamma \geq \alpha > 0$. Hence, for any $(u_1, u_2) > 0$, we have $\alpha u_1 \leq \gamma u_1$ and $\gamma u_2 < \beta u_2$ such that $0 < \alpha u_1 + \gamma u_2 < \gamma u_1 + \beta u_2$. This implies that $\mathcal{W}$ is empty, yielding a contradiction. Finally, when $\mathcal{W}$ is nonempty, we see, in view of $\alpha > \gamma$ proven above, that for any $u_2 > 0$, there exists a sufficiently large $u_1 > 0$ such that $\alpha u_1 + \gamma u_2 > 0$ and $\alpha u_1 + \gamma u_2 > \gamma u_1 + \beta u_2$. This shows that $(\alpha u_1 + \gamma u_2)_+ > (\gamma u_1 + \beta u_2)_+$. Hence, $\{u_2 \,|\, (u_1, u_2) \in \mathcal{W}\} = \mathbb{R}_{++}$. ☐

Theorem 5.4. *Given a matrix* $A \in \mathbb{R}^{m\times N}$ *with unit columns and the index set* $\mathcal{S} = \{1, 2, 3\}$, *let* $M := A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}}$. *Then every nonzero vector* $x \in \mathbb{R}^N_+$ *with supp*$(x) = \mathcal{S}$ *is recovered from* $y = Ax$ *via constrained matching pursuit if and only if each of the following conditions holds:*

(i) $A_{\bullet\mathcal{S}}$ *has full column rank;*

(ii) $\big\|(A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}} u)_+\big\|_\infty > \big\|(A^T_{\bullet\mathcal{S}^c} A_{\bullet\mathcal{S}} u)_+\big\|_\infty$ *for all* $u \in \mathbb{R}^3_{++}$*;*

(iii) For any $\mathcal{J} \in \{\{1\}, \{2\}, \{3\}\}$, $\big\|(M/M_{\mathcal{J}\mathcal{J}}\, v)_+\big\|_\infty > \big\|(A^T_{\bullet\mathcal{S}^c}[I - A^T_{\bullet\mathcal{J}} A_{\bullet\mathcal{J}}] A_{\mathcal{S}\setminus\mathcal{J}} v)_+\big\|_\infty$ *for all* $v \in \mathbb{R}^2_{++}$*;*

(iv) All the following implications hold:

$$\Big[1 - \vartheta_{12}^2 > \min(\Delta_{13}, \Delta_{23})\Big] \implies \Big[\det M > \max_{i\in\mathcal{S}^c} \big(\vartheta_{i3}(1-\vartheta_{12}^2) - \vartheta_{i1}\Delta_{13} - \vartheta_{i2}\Delta_{23}\big)_+\Big],$$

$$\Big[1 - \vartheta_{13}^2 > \min(\Delta_{12}, \Delta_{23})\Big] \implies \Big[\det M > \max_{i\in\mathcal{S}^c} \big(\vartheta_{i2}(1-\vartheta_{13}^2) - \vartheta_{i1}\Delta_{12} - \vartheta_{i3}\Delta_{23}\big)_+\Big],$$

$$\Big[1 - \vartheta_{23}^2 > \min(\Delta_{12}, \Delta_{13})\Big] \implies \Big[\det M > \max_{i\in\mathcal{S}^c} \big(\vartheta_{i1}(1-\vartheta_{23}^2) - \vartheta_{i2}\Delta_{12} - \vartheta_{i3}\Delta_{13}\big)_+\Big],$$

where $\Delta_{12} := \vartheta_{12} - \vartheta_{13}\vartheta_{23}$, $\Delta_{13} := \vartheta_{13} - \vartheta_{12}\vartheta_{23}$, *and* $\Delta_{23} := \vartheta_{23} - \vartheta_{12}\vartheta_{13}$.

Remark 5.2. We comment on the above conditions before presenting a proof:

(a) Since the matrix $M = \begin{bmatrix} 1 & \vartheta_{12} & \vartheta_{13} \\ \vartheta_{12} & 1 & \vartheta_{23} \\ \vartheta_{13} & \vartheta_{23} & 1 \end{bmatrix}$, its determinant $\det M = 1 + 2\vartheta_{12}\vartheta_{13}\vartheta_{23} - \vartheta_{12}^2 - \vartheta_{13}^2 - \vartheta_{23}^2$.

(b) If the hypothesis of an implication in condition (iv) fails, then that implication holds even when the conclusion statement is false. Hence, that implication is vacuously true and can be neglected.

(c) Since each Schur complement of $M := A_{\bullet\mathcal{S}}^T A_{\bullet\mathcal{S}}$ is positive definite, we have $(1-\vartheta_{13}^2)(1-\vartheta_{23}^2) \geq \Delta_{12}^2$, $(1-\vartheta_{12}^2)(1-\vartheta_{23}^2) \geq \Delta_{13}^2$, and $(1-\vartheta_{12}^2)(1-\vartheta_{13}^2) \geq \Delta_{23}^2$. By virtue of these inequalities, it is easy to verify that at least two hypotheses of the three implications in condition (iv) must hold.

Proof of Theorem 5.4. "If". Suppose conditions (i)-(iv) hold. Fix an arbitrary $z = (z_{\mathcal{S}}, 0) \in \mathbb{R}_+^N$ with $z_{\mathcal{S}} = (z_1, z_2, z_3) \in \mathbb{R}_{++}^3$, and let $y = Az = A_{\bullet\mathcal{S}} z_{\mathcal{S}}$. Consider the following three steps of Algorithm 1:

• Step 1: Let $x^0 = 0$. Since $y = A_{\bullet\mathcal{S}} z_{\mathcal{S}}$, it follows from condition (ii) that $\max_{i=1,2,3}(A_{\bullet i}^T A_{\bullet\mathcal{S}} z_{\mathcal{S}})_+ > \max_{j\in\mathcal{S}^c}(A_{\bullet j}^T A_{\bullet\mathcal{S}} z_{\mathcal{S}})_+$. Hence, it follows from Algorithm 1 that $j_1^* \in \mathcal{S} = \{1,2,3\}$, and the index set $\mathcal{J}_1 = \{j_1^*\}$. Further, $x^1 := \arg\min_{x\geq 0, \mathrm{supp}(x)\subseteq\mathcal{J}_1} \|y - Ax\|_2^2$ is given by $x^1 = \langle A_{\bullet j_1^*}, A_{\bullet\mathcal{S}} z_{\mathcal{S}}\rangle_+ \mathbf{e}_{j_1^*}$, where $\langle A_{\bullet j_1^*}, A_{\bullet\mathcal{S}} z_{\mathcal{S}}\rangle_+ > 0$ in view of Proposition 3.1.

• Step 2: By observing that x^1 is the optimal solution obtained from Step 1 with $\mathrm{supp}(x^1) = \mathcal{J}_1 = \{j_1^*\}$, it follows from Lemma 5.4 and $A_{\bullet\mathcal{J}_1}^T A_{\bullet\mathcal{J}_1} = 1$ that by letting the index set $\mathcal{I} := \mathcal{S} \setminus \mathcal{J}_1$,

$$\begin{aligned} \max_{i\in\mathcal{S}} \left(A_{\bullet i}^T A(z - x^1)\right)_+ &= \left\|\left(A_{\bullet\mathcal{S}}^T A(z-x^1)\right)_+\right\|_\infty = \left\|\left((M/M_{\mathcal{J}_1\mathcal{J}_1}) \cdot z_{\mathcal{I}}\right)_+\right\|_\infty, \\ \max_{j\in\mathcal{S}^c} \left(A_{\bullet j}^T A(z - x^1)\right)_+ &= \left\|\left(A_{\bullet\mathcal{S}^c}^T A(z-x^1)\right)_+\right\|_\infty = \left\|(A_{\bullet\mathcal{S}^c}^T [I - A_{\bullet\mathcal{J}_1} A_{\bullet\mathcal{J}_1}^T] A_{\mathcal{I}} \cdot z_{\mathcal{I}})_+\right\|_\infty. \end{aligned}$$

Noting that the Schur complement $M/M_{\mathcal{J}_1\mathcal{J}_1}$ is positive definite and $z_{\mathcal{I}} > 0$, we deduce via Lemma 5.3 that $\left\|\left((M/M_{\mathcal{J}_1\mathcal{J}_1}) \cdot z_{\mathcal{I}}\right)_+\right\|_\infty > 0$. By $z_{\mathcal{I}} > 0$ and condition (iii), we have $\max_{i\in\mathcal{S}} \left(A_{\bullet i}^T A(z-x^1)\right)_+ > \max_{j\in\mathcal{S}^c} \left(A_{\bullet j}^T A(z-x^1)\right)_+$. In light of Algorithm 1, we see that $j_2^* := \arg\max_{i\in\mathcal{S}} \left(A_{\bullet i}^T A(z-x^1)\right)_+$ satisfies $j_2^* \in \mathcal{I}$, and $\mathcal{J}_2 = \{j_1^*, j_2^*\} \subset \mathcal{S}$ with $j_1^* \neq j_2^*$. Moreover, let x^2 be the unique optimal solution to $\min_{w\geq 0, \mathrm{supp}(w)\subseteq\mathcal{J}_2} \|y - Aw\|^2$. Then it follows from Proposition 3.1 that $\mathrm{supp}(x^2) = \mathcal{J}_2$.

• Step 3: Let the index j_3 be such that $\{j_3\} = \mathcal{S} \setminus \mathcal{J}_2$. Note that $\{j_2^*, j_3\} = \mathcal{I}$. Hence, the Schur complement $U := M/M_{\mathcal{J}_1\mathcal{J}_1}$ is one of the following 2×2 positive definite matrices:

$$U^1 := \begin{bmatrix} 1-\vartheta_{12}^2 & \Delta_{23} \\ \Delta_{23} & 1-\vartheta_{13}^2 \end{bmatrix}, \quad U^2 := \begin{bmatrix} 1-\vartheta_{12}^2 & \Delta_{13} \\ \Delta_{13} & 1-\vartheta_{23}^2 \end{bmatrix}, \quad U^3 := \begin{bmatrix} 1-\vartheta_{13}^2 & \Delta_{12} \\ \Delta_{12} & 1-\vartheta_{23}^2 \end{bmatrix}, \tag{23}$$

where Δ_{ij}'s are defined in condition (iv), $U^1 = M/M_{11}$, $U^2 = M/M_{22}$, and $U^3 = M/M_{33}$. Hence, $U = \begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix}$ is positive definite, where $\alpha, \beta \in \{1-\vartheta_{12}^2, 1-\vartheta_{13}^2, 1-\vartheta_{23}^2\}$ with $\alpha \neq \beta$, and $\gamma \in \{\Delta_{12}, \Delta_{13}, \Delta_{23}\}$. Furthermore, either $(U_{1\bullet} z_{\mathcal{I}})_+ \geq (U_{2\bullet} z_{\mathcal{I}})_+$ or $(U_{2\bullet} z_{\mathcal{I}})_+ \geq (U_{1\bullet} z_{\mathcal{I}})_+$, where $U_{i\bullet}$ denotes the ith row of U. Since $z_{\mathcal{I}} > 0$, it follows from Lemma 5.5 that either $\alpha > \gamma$ or $\beta > \gamma$. We first consider the case where $\alpha > \gamma$. In this case, $\alpha = 1 - \vartheta_{j_1^*, j_2^*}^2$, $\beta = 1 - \vartheta_{j_1^*, j_3}^2$, and $\gamma = \Delta_{j_2^*, j_3}$. In light of the implications given by condition (iv), we have that

$$\det M > \max_{i\in\mathcal{S}^c} \left(\vartheta_{i,j_3}(1 - \vartheta_{j_1^*, j_2^*}^2) - \vartheta_{i,j_1^*}\Delta_{j_1^*, j_3} - \vartheta_{i,j_2^*}\Delta_{j_2^*, j_3}\right)_+. \tag{24}$$

Additionally, since x^2 is the unique solution to $\min_{w\geq 0, \mathrm{supp}(w)\subseteq\mathcal{J}_2} \|y - Aw\|^2$ with $\mathrm{supp}(x^2) = \mathcal{J}_2$, it follows from Lemma 5.4 that by letting $\widetilde{\mathcal{I}} := \mathcal{S} \setminus \mathcal{J}_2 = \{j_3\}$,

$$\begin{aligned} \max_{i\in\mathcal{S}} \left(A_{\bullet i}^T A(z - x^2)\right)_+ &= \left\|\left((M/M_{\mathcal{J}_2\mathcal{J}_2}) \cdot z_{\widetilde{\mathcal{I}}}\right)_+\right\|_\infty, \\ \max_{i\in\mathcal{S}^c} \left(A_{\bullet i}^T A(z - x^2)\right)_+ &= \left\|(A_{\bullet\mathcal{S}^c}^T [I - A_{\bullet\mathcal{J}_2}^T (A_{\bullet\mathcal{J}_2} A_{\bullet\mathcal{J}_2})^{-1} A_{\bullet\mathcal{J}_2}^T] A_{\widetilde{\mathcal{I}}} \cdot z_{\widetilde{\mathcal{I}}})_+\right\|_\infty. \end{aligned}$$

Note that $z_{\widetilde{\mathcal{I}}}$ and $M/M_{\mathcal{J}_2\mathcal{J}_2}$ are positive scalars. It follows from the Schur determinant formula that $M/M_{\mathcal{J}_2\mathcal{J}_2} = \det(M/M_{\mathcal{J}_2\mathcal{J}_2}) = \det M/\det(M_{\mathcal{J}_2\mathcal{J}_2})$. Thus $\max_{i\in\mathcal{S}}\big(A^T_{\bullet i}A(z-x^2)\big)_+ = \det M/\det(M_{\mathcal{J}_2\mathcal{J}_2})\cdot z_{\widetilde{\mathcal{I}}}$. Further, direct calculations show that $(A_{\bullet\mathcal{J}_2}A_{\bullet\mathcal{J}_2})^{-1}A^T_{\bullet\mathcal{J}_2}A_{\widetilde{\mathcal{I}}} = (\Delta_{j_1^*,j_3},\Delta_{j_2^*,j_3})^T/\det(M_{\mathcal{J}_2\mathcal{J}_2})$. In view of this result and $\det(M_{\mathcal{J}_2\mathcal{J}_2}) = 1-\vartheta^2_{j_1^*,j_2^*}$, we have, for each $i\in\mathcal{S}^c$,

$$\big((A^T_{\bullet i}[I - A^T_{\bullet\mathcal{J}_2}(A_{\bullet\mathcal{J}_2}A_{\bullet\mathcal{J}_2})^{-1}A^T_{\bullet\mathcal{J}_2}]A_{\widetilde{\mathcal{I}}}\cdot z_{\widetilde{\mathcal{I}}}\big)_+ = \frac{z_{\widetilde{\mathcal{I}}}}{\det(M_{\mathcal{J}_2\mathcal{J}_2})}\Big(\vartheta_{i,j_3}(1-\vartheta^2_{j_1^*,j_2^*}) - \vartheta_{i,j_1^*}\Delta_{j_1^*,j_3} - \vartheta_{i,j_2^*}\Delta_{j_2^*,j_3}\Big)_+.$$

These results and the inequality (24) imply that $\max_{i\in\mathcal{S}}\big(A^T_{\bullet i}A(z-x^2)\big)_+ > \max_{i\in\mathcal{S}^c}\big(A^T_{\bullet i}A(z-x^2)\big)_+$. The other case where $\beta>\gamma$ can be established by the similar argument. Therefore, following Algorithm 1, $j_3^* := \arg\max_{i\in\mathcal{S}}\big(A^T_{\bullet i}A(z-x^2)\big)_+$ satisfies $j_3^* = j_3$. This yields $\mathcal{J}_3 = \mathcal{S}$. Since $A_{\bullet\mathcal{S}}$ has full column rank, the exact vector recovery is achieved.

"Only if". Suppose every nonzero vector $x\in\mathbb{R}^N_+$ with $\mathrm{supp}(x)=\mathcal{S}$ is recovered from $y=Ax$ via constrained matching pursuit for a given matrix $A\in\mathbb{R}^{m\times N}$ and the index set $\mathcal{S}=\{1,2,3\}$. It follows from Lemma 5.2 that condition (i) must hold. Besides, by setting $x^0=0$, we see via Corollary 3.1 that $\max_{i\in\mathcal{S}}(A^T_{\bullet i}A_{\bullet\mathcal{S}}z_{\mathcal{S}})_+ > \max_{j\in\mathcal{S}^c}(A^T_{\bullet j}A_{\bullet\mathcal{S}}z_{\mathcal{S}})_+$ holds for all $z_{\mathcal{S}}\in\mathbb{R}^3_{++}$. This yields condition (ii).

For each $p\in\mathcal{S}$, define the set $\mathcal{W}_p := \{z_{\mathcal{S}}\in\mathbb{R}^3_{++}\,|\,(A^T_{\bullet p}A_{\bullet\mathcal{S}}z_{\mathcal{S}})_+ = \max_{i\in\mathcal{S}}(A^T_{\bullet i}A_{\bullet\mathcal{S}}z_{\mathcal{S}})_+\}$. Clearly, $\mathbb{R}^3_{++} = \mathcal{W}_1\cup\mathcal{W}_2\cup\mathcal{W}_3$. Since the matrix $M := A^T_{\bullet\mathcal{S}}A_{\bullet\mathcal{S}}$ given by (c) of Remark 5.2 is positive definite, we observe $|\vartheta_{ij}|<1$ for any $i\neq j$. Based on this observation, it is easy to show that for any given $(z_2,z_3)>0$, there exists a sufficiently large $z_1>0$ such that $(z_1,z_2,z_3)\in\mathcal{W}_1$. Hence, $\mathcal{W}_1$ is nonempty and $\{(z_2,z_3)\,|\,z_{\mathcal{S}}=(z_1,z_2,z_3)\in\mathcal{W}_1\} = \mathbb{R}^2_{++}$. By a similar argument, we deduce that $\mathcal{W}_2$ and $\mathcal{W}_3$ are nonempty and $\{(z_1,z_3)\,|\,z_{\mathcal{S}}=(z_1,z_2,z_3)\in\mathcal{W}_2\}=\mathbb{R}^2_{++}$ and $\{(z_1,z_2)\,|\,z_{\mathcal{S}}=(z_1,z_2,z_3)\in\mathcal{W}_3\}=\mathbb{R}^2_{++}$. Since $\mathbb{R}^3_{++}=\mathcal{W}_1\cup\mathcal{W}_2\cup\mathcal{W}_3$, $z_{\mathcal{S}}$ belongs to one of $\mathcal{W}_i$'s for any $z_{\mathcal{S}}\in\mathbb{R}^3_{++}$. For each $p\in\mathcal{S}$, it follows from Algorithm 1 that for any $z\in\mathcal{W}_p$, the corresponding unique $x^1 = (A^T_{\bullet p}A_{\bullet\mathcal{S}}z_{\mathcal{S}})_+\mathbf{e}_p$, where $(A^T_{\bullet p}A_{\bullet\mathcal{S}}z_{\mathcal{S}})_+>0$. Moreover, we must have $\max_{i\in\mathcal{S}}\big(A^T_{\bullet i}A(z-x^1)\big)_+ > \max_{j\in\mathcal{S}^c}\big(A^T_{\bullet j}A(z-x^1)\big)_+$. This condition, as shown at Step 2 of the "if" part, is equivalent to $\big\|(M/M_{\mathcal{J}_1\mathcal{J}_1}\,z_{\mathcal{I}})_+\big\|_\infty > \big\|(A^T_{\bullet\mathcal{S}^c}[I-A^T_{\bullet\mathcal{J}_1}A_{\bullet\mathcal{J}_1}]A_{\mathcal{I}}z_{\mathcal{I}})_+\big\|_\infty$, where $\mathcal{J}_1=\{p\}$ and $\mathcal{I}=\mathcal{S}\setminus\mathcal{J}_1$. Since $\{z_{\mathcal{I}}\,|\,z_{\mathcal{S}}\in\mathcal{W}_p\}=\mathbb{R}^2_{++}$ as shown before, we obtain condition (iii).

To establish condition (iv), we first show the following claim: if $1-\vartheta^2_{12}>\min(\Delta_{13},\Delta_{23})$ holds true, then there exists $z\in\mathbb{R}^N_+$ with $\mathrm{supp}(z)=\mathcal{S}$ such that when $y=Az$, Algorithm 1 give rises to $\mathcal{J}_2=\{1,2\}$. To prove this claim, it is noted that $1-\vartheta^2_{12}>\min(\Delta_{13},\Delta_{23})$ is equivalent to $1-\vartheta^2_{12}>\Delta_{23}$ or $1-\vartheta^2_{12}>\Delta_{13}$. For the former case, i.e., $1-\vartheta^2_{12}>\Delta_{23}$, it follows from Lemma 5.5 and $U^1=M/M_{11}$ given in (23) that there exists $v:=(v_1,v_2)^T\in\mathbb{R}^2_{++}$ such that $(U^1_{1\bullet}v)_+\geq(U^1_{2\bullet}v)_+$. Further, as shown previously, there exists a sufficiently large $v_0>0$ such that $\widetilde{z}=(\widetilde{z}_{\mathcal{S}},0)$ with $\widetilde{z}_{\mathcal{S}}:=(\widetilde{z}_1,\widetilde{z}_2,\widetilde{z}_3)=(v_0,v_1,v_2)$ satisfies $\widetilde{z}\in\mathcal{W}_1$. This implies via Lemma 5.4 and the argument for Step 1 of the "if" part that when $y=A\widetilde{z}$, Algorithm 1 give rises to $(j_1^*,j_2^*)=(1,2)$ and $\mathcal{J}_2=\{1,2\}$. The similar argument can be used to show that if $1-\vartheta^2_{12}>\Delta_{13}$ holds, then there exists $z\in\mathbb{R}^N_+$ with $\mathrm{supp}(z)=\mathcal{S}$ such that when $y=Az$, Algorithm 1 give rises to $(j_1^*,j_2^*)=(2,1)$ and $\mathcal{J}_2=\{1,2\}$. The above proof can be extended to show that if $1-\vartheta^2_{13}>\min(\Delta_{12},\Delta_{23})$ (respectively $1-\vartheta^2_{13}>\min(\Delta_{12},\Delta_{23})$) holds, then there exists $z\in\mathbb{R}^N_+$ with $\mathrm{supp}(z)=\mathcal{S}$ such that when $y=Az$, Algorithm 1 yields $\mathcal{J}_2=\{1,3\}$ (respectively $\mathcal{J}_2=\{2,3\}$).

As indicated in Remark 5.2, if the hypothesis of an implication in condition (iv) is false, then that implication holds true vacuously. Now consider an implication in condition (iv) whose hypothesis holds true. Then there exists $z\in\mathbb{R}^N_+$ with $\mathrm{supp}(z)=\mathcal{S}$ such that Algorithm 1 yields $\mathcal{J}_2:=\{j_1^*,j_2^*\}$ from $y=Az$. Hence, the corresponding x^2 obtained from $y=Az$ via Algorithm 1 satisfies $\mathrm{supp}(x^2)=\mathcal{J}_2$. Since the exact support recovery implies that $\max_{i\in\mathcal{S}}(A^T_{\bullet i}A(z-x^2))_+ > \max_{j\in\mathcal{S}^c}(A^T_{\bullet i}A(z-x^2))_+$, we deduce, in view of $\mathrm{supp}(x^2)=\mathcal{J}_2$, Lemma 5.4 and the argument for Step 3 of the "if" part, that

$$\frac{\det M}{\det(M_{\mathcal{J}_2\mathcal{J}_2})}\cdot z_{\widetilde{\mathcal{I}}} > \frac{z_{\widetilde{\mathcal{I}}}}{\det(M_{\mathcal{J}_2\mathcal{J}_2})}\Big(\vartheta_{i,j_3}(1-\vartheta^2_{j_1^*,j_2^*}) - \vartheta_{i,j_1^*}\Delta_{j_1^*,j_3} - \vartheta_{i,j_2^*}\Delta_{j_2^*,j_3}\Big)_+,$$

where $\widetilde{\mathcal{I}}=\{j_3\}=\mathcal{S}\setminus\mathcal{J}_2$, $z_{\widetilde{\mathcal{I}}}\in\mathbb{R}_{++}$, and $\det(M_{\mathcal{J}_2\mathcal{J}_2})=1-\vartheta^2_{j_1^*,j_2^*}$. This yields condition (iv). □

5.2.3 Sufficient Conditions for Exact Vector Recovery on $\mathbb{R}^N_+$ for a Fixed Support

When a given support $\mathcal{S}$ is of size greater than or equal to 4, necessary *and* sufficient conditions are difficult to obtain due to increasing complexities. Hence, we seek neat sufficient conditions in this subsection.

Theorem 5.5. *Given a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns and the index set $\mathcal{S} \subset \{1,\ldots,N\}$, let $M := A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}}$. Then every nonzero vector $z \in \mathbb{R}^N_+$ with $\mathrm{supp}(z) = \mathcal{S}$ is recovered from $y = Az$ via constrained matching pursuit if the following conditions hold:*

(i) $A_{\bullet\mathcal{S}}$ has full column rank or equivalently M is positive definite; and

(ii) For any (possibly empty) index set $\mathcal{J} \subset \mathcal{S}$,

$$\big\|(M/M_{\mathcal{J}\mathcal{J}}\, x)_+\big\|_\infty > \big\|\big(A^T_{\bullet\mathcal{S}^c}[I - A^T_{\bullet\mathcal{J}}(A^T_{\bullet\mathcal{J}}A_{\bullet\mathcal{J}})^{-1}A_{\bullet\mathcal{J}}]A_{\bullet\mathcal{S}\setminus\mathcal{J}}\, x\big)_+\big\|_\infty, \quad \forall\, x \in \mathbb{R}^{|\mathcal{S}\setminus\mathcal{J}|}_{++}, \tag{25}$$

where $M/M_{\mathcal{J}\mathcal{J}}$ is the Schur complement of $M_{\mathcal{J}\mathcal{J}}$ in M.

Proof. Due to condition (i), it suffices to show the exact support recovery of each $z \in \mathbb{R}^N_+$ with $\mathrm{supp}(z) = \mathcal{S}$ via Algorithm 1 from $y = Az$. Toward this end, we see via a similar argument for Corollary 3.1 that condition (**H**) given by (9) holds if for any $0 \neq u \in \mathbb{R}^N_+$ with $\mathrm{supp}(u) = \mathcal{S}$, any index set $\mathcal{J} \subset \mathcal{S}$, and the (unique) optimal solution $v = \arg\min_{w\geq 0, \mathrm{supp}(w)\subseteq\mathcal{J}} \|A(u-w)\|_2^2$, the following holds:

$$\max_{i\in\mathcal{S}} \big(A^T_{\bullet i}A(u-v)\big)_+ = \max_{i\in\mathcal{S}\setminus\mathcal{J}} \big(A^T_{\bullet i}A(u-v)\big)_+ > \max_{j\in\mathcal{S}^c} \big(A^T_{\bullet j}A(u-v)\big)_+,$$

where the first equation follows from Lemma 3.1. Let $\mathcal{J}^* := \mathrm{supp}(v)$. Hence, $\mathcal{J}^* \subseteq \mathcal{J} \subset \mathcal{S}$. Since v is the optimal solution to $\min_{w\geq 0, \mathrm{supp}(w)\subseteq\mathcal{J}} \|A(u-w)\|_2^2$, we deduce via Lemma 5.4 that

$$\begin{aligned}
\max_{i\in\mathcal{S}} \big(A^T_{\bullet i}A(u-v)\big)_+ &= \big\|\big(A^T_{\bullet\mathcal{S}}A(u-v)\big)_+\big\|_\infty = \big\|\big((M/M_{\mathcal{J}^*\mathcal{J}^*})\cdot u_{\mathcal{I}}\big)_+\big\|_\infty,\\
\max_{j\in\mathcal{S}^c} \big(A^T_{\bullet j}A(u-v)\big)_+ &= \big\|\big(A^T_{\bullet\mathcal{S}^c}A(u-v)\big)_+\big\|_\infty = \big\|\big(A^T_{\bullet\mathcal{S}^c}[I - A^T_{\bullet\mathcal{J}^*}(A^T_{\bullet\mathcal{J}^*}A_{\bullet\mathcal{J}^*})^{-1}A_{\bullet\mathcal{J}^*}]A_{\bullet\mathcal{I}}\cdot u_{\mathcal{I}}\big)_+\big\|_\infty,
\end{aligned}$$

where $\mathcal{I} := \mathcal{S}\setminus\mathcal{J}^*$ is nonempty. Since $u_{\mathcal{I}} > 0$, we see that $\max_{i\in\mathcal{S}} \big(A^T_{\bullet i}A(u-v)\big)_+ > \max_{j\in\mathcal{S}^c} \big(A^T_{\bullet j}A(u-v)\big)_+$ holds under condition (ii). This leads to the desired result. □

In what follows, we develop conditions to verify the inequality given in (25), which leads to a numerical scheme to check (25). Fix an index set $\mathcal{J} \subset \mathcal{S}$, and let $r := |\mathcal{S}\setminus\mathcal{J}|$. Further, let $M/M_{\mathcal{J}\mathcal{J}} = [p_1, \cdots, p_r]$, and $E := \big(A^T_{\bullet\mathcal{S}^c}[I - A^T_{\bullet\mathcal{J}}(A^T_{\bullet\mathcal{J}}A_{\bullet\mathcal{J}})^{-1}A_{\bullet\mathcal{J}}]A_{\bullet\mathcal{S}\setminus\mathcal{J}}\big)^T = [q_1, \cdots, q_{|\mathcal{S}^c|}]$, namely, $p_i \in \mathbb{R}^r$ is the ith column of $M/M_{\mathcal{J}\mathcal{J}}$ and $q_j \in \mathbb{R}^r$ is the jth column of E.

Lemma 5.6. *The inequality (25) for a fixed index set $\mathcal{J} \subset \mathcal{S}$ holds if and only if for each $q_j \in \mathbb{R}^r$, there exist $w \in \mathbb{R}^r_+$ and $0 \neq (w', \beta) \in \mathbb{R}^r_+ \times \mathbb{R}_+$ such that $[p_1 - q_j, p_2 - q_j, \cdots, p_r - q_j]w = w' + \beta\cdot q_j$.*

Proof. Since the Schur complement $M/M_{\mathcal{J}\mathcal{J}}$ is symmetric, it is easy to see that the inequality (25) fails if and only if there exists $v > 0$ such that $\max_{i=1,\ldots,r}(p_i^T v)_+ \leq (q_j^T v)_+$ for some j. In view of Lemma 5.3, we deduce that $\max_{i=1,\ldots,r}(p_i^T v)_+ > 0$ such that $q_j^T v > 0$ for this j. Hence, the inequality system $\max_{i=1,\ldots,r}(p_i^T v)_+ \leq (q_j^T v)_+, v > 0$ is equivalent to the following linear inequality system:

$$(\mathrm{I}): \quad v > 0, \quad q_j^T v > 0, \quad q_j^T v \geq p_i^T v, \quad \forall\, i = 1,\ldots,r.$$

By Motzkin's Transposition Theorem, (I) has no solution if and only if there exist $w \in \mathbb{R}^r_+$ and $0 \neq (w', \beta) \in \mathbb{R}^r_+ \times \mathbb{R}_+$ such that $[p_1 - q_j, p_2 - q_j, \cdots, p_r - q_j]w = w' + \beta\cdot q_j$, yielding the desired result. □

The condition derived in the above lemma can be effectively verified via a linear program for the given matrices $M/M_{\mathcal{J}\mathcal{J}}$ and E.

5.3 Exact Vector Recovery on $\mathbb{R}^{N_1} \times \mathbb{R}^{N_2}_+$ for a Fixed Support

In this subsection, we briefly discuss an extension of the preceding exact vector recovery results to a Cartesian product of copies of $\mathbb{R}$ and $\mathbb{R}_+$. Let $\mathcal{I}_1$ and $\mathcal{I}_+$ be two nonempty index subsets that form a disjoint union of $\{1, \ldots, N\}$. Consider the constraint set $\mathcal{P} = \mathbb{R}_{\mathcal{I}_1} \times (\mathbb{R}_+)_{\mathcal{I}_+}$. The following preliminary result can be easily extended from Corollary 3.1 and Lemma 5.2; its proof is thus omitted.

Lemma 5.7. *Let $A \in \mathbb{R}^{m\times N}$ be a matrix with unit columns, and $\mathcal{P} = \mathbb{R}_{\mathcal{I}_1} \times (\mathbb{R}_+)_{\mathcal{I}_+}$. The following hold:*

(i) Let $0 \neq z \in \Sigma_K \cap \mathcal{P}$ with $|supp(z)| = r$. Then the exact support recovery of z is achieved if and only if for any sequence $\big((x^k, j^*_k, \mathcal{J}_k)\big)_{k\in\mathbb{N}}$ generated by Algorithm 1 with $y = Az$,

$$\max\Big(\max_{j\in(supp(z)\setminus\mathcal{J}_k)\cap\mathcal{I}_1} |A^T_{\bullet j}A(z-x^k)|, \max_{j\in(supp(z)\setminus\mathcal{J}_k)\cap\mathcal{I}_+} [A^T_{\bullet j}A(z-x^k)]_+ \Big)$$
$$> \max\Big(\max_{j\in[supp(z)]^c\cap\mathcal{I}_1} |A^T_{\bullet j}A(z-x^k)|, \max_{j\in[supp(z)]^c\cap\mathcal{I}_+} [A^T_{\bullet j}A(z-x^k)]_+\Big), \quad \forall\, k = 0, 1, \ldots, r-1.$$

(ii) Let $\mathcal{S}$ be a nonempty index subset of $\{1, \ldots, N\}$. The exact vector recovery of every vector $x \in \mathcal{P}$ with $supp(x) = \mathcal{S}$ is achieved via constrained matching pursuit only if $A_{\bullet\mathcal{S}}$ has full column rank.

The next result characterizes the exact vector recovery on $\mathcal{P}$ for a given support $\mathcal{S}$ of size 2.

Theorem 5.6. *Given a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns and the index set $\mathcal{S} = \{1, 2\}$ with $1 \in \mathcal{I}_1$ and $2 \in \mathcal{I}_+$, every vector $x \in \mathcal{P} = \mathbb{R}_{\mathcal{I}_1} \times (\mathbb{R}_+)_{\mathcal{I}_+}$ with $supp(x) = \mathcal{S}$ is recovered from $y = Ax$ via constrained matching pursuit if and only if the following conditions hold:*

(i) $A_{\bullet\mathcal{S}}$ has full column rank or equivalently $|\vartheta_{12}| < 1$;

(ii) $\max\big(|z_1 + \vartheta_{12}z_2|, (\vartheta_{12}z_1 + z_2)_+\big) > \max\Big(\max_{j\in\mathcal{S}^c\cap\mathcal{I}_1} |\vartheta_{j1}z_1 + \vartheta_{j2}z_2|, \max_{j\in\mathcal{S}^c\cap\mathcal{I}_+} \big(\vartheta_{j1}z_1 + \vartheta_{j2}z_2\big)_+\Big)$, $\forall\, (z_1, z_2)^T \in \big(\mathbb{R}\setminus\{0\}\big) \times \mathbb{R}_{++}$;

(iii) $1 - \vartheta^2_{12} > \max\Big(\max_{j\in\mathcal{S}^c\cap\mathcal{I}_1} |\vartheta_{j2} - \vartheta_{12}\vartheta_{j1}|, \max_{j\in\mathcal{S}^c\cap\mathcal{I}_+}(\vartheta_{j2} - \vartheta_{12}\vartheta_{j1})_+, \max_{j\in\mathcal{S}^c} |\vartheta_{j1} - \vartheta_{12}\vartheta_{j2}|\Big)$.

Proof. "Only if". Suppose the exact vector recovery is achieved for any $x \in \mathcal{P}$ with $\text{supp}(x) = \mathcal{S}$. Condition (i) follows from statement (ii) of Lemma 5.7, and condition (ii) follows from Step 1 of Algorithm 1 and statement (i) of Lemma 5.7 with $x^0 = 0$ and $\mathcal{J}_0 = \emptyset$. To establish condition (iii), we first notice via $|\vartheta_{12}| < 1$ that for any $z \in \mathcal{P}$ with $\text{supp}(z) = \mathcal{S}$, i.e., $z_1 \neq 0$ and $z_2 > 0$, $|z_1 + \vartheta_{12}z_2| \geq (\vartheta_{12}z_1 + z_2)_+$ if and only if $|z_1| \geq z_2 > 0$, and $|z_1 + \vartheta_{12}z_2| \leq (\vartheta_{12}z_1 + z_2)_+$ if and only if $z_2 \geq |z_1| > 0$. When the former holds, i.e., $|z_1| \geq z_2 > 0$, we have $j^*_1 = 1$ and $x^1 = (z_1 + \vartheta_{12}z_2) \cdot \mathbf{e}_1$. Hence, $A^T_{\bullet j}A(z - x^1) = (\vartheta_{j2} - \vartheta_{j1}\vartheta_{12})z_2$. Using Step 2 of Algorithm 1 and statement (i) of Lemma 5.7 with $\mathcal{J}_1 = \{1\}$, it is easy to obtain $1 - \vartheta^2_{12} > \max\big(\max_{j\in\mathcal{S}^c\cap\mathcal{I}_1} |\vartheta_{j2} - \vartheta_{12}\vartheta_{j1}|, \max_{j\in\mathcal{S}^c\cap\mathcal{I}_+}(\vartheta_{j2} - \vartheta_{12}\vartheta_{j1})_+\big)$. We next consider the case where $z_2 \geq |z_1| > 0$. In this case, $j^*_1 = 2$ such that $x^1 = (\vartheta_{12}z_1 + z_2)_+ \cdot \mathbf{e}_2$, where $\vartheta_{12}z_1 + z_2 > 0$. Hence, $A^T_{\bullet j}A(z - x^1) = (\vartheta_{j1} - \vartheta_{j2}\vartheta_{12})z_1$. Applying Step 2 of Algorithm 1 and statement (i) of Lemma 5.7 with $\mathcal{J}_1 = \{2\}$, we have that

$$(1 - \vartheta^2_{12})|z_1| > \max\big(\max_{j\in\mathcal{S}^c\cap\mathcal{I}_1} |(\vartheta_{j1} - \vartheta_{12}\vartheta_{j2})z_1|, \max_{j\in\mathcal{S}^c\cap\mathcal{I}_+} [(\vartheta_{j1} - \vartheta_{12}\vartheta_{j2})z_1]_+\big).$$

It is easy to show that $(1 - \vartheta^2_{12})|z_1| > \max_{j\in\mathcal{S}^c\cap\mathcal{I}_+}[(\vartheta_{j1} - \vartheta_{12}\vartheta_{j2})z_1]_+$ for any $z_1 \neq 0$ if and only if $1 - \vartheta^2_{12} > \max_{j\in\mathcal{S}^c\cap\mathcal{I}_+} |\vartheta_{j1} - \vartheta_{12}\vartheta_{j2}|$. This yields $1 - \vartheta^2_{12} > \max_{j\in\mathcal{S}^c} |\vartheta_{j1} - \vartheta_{12}\vartheta_{j2}|$, and condition (iii).

"If". This part can be shown in a similar way by reversing the previous argument. □

Necessary and sufficient conditions for the exact vector recovery on $\mathcal{P}$ for a given support $\mathcal{S}$ of size 3 can be established via a similar argument for Theorem 5.4. Instead doing this, we provide a sufficient condition for a given support of arbitrary size. To simplify notation, we define the following function $F_{\mathcal{I},\mathcal{J}} : \mathbb{R}_{\mathcal{I}} \times (\mathbb{R}_+)_{\mathcal{J}} \to \mathbb{R}$ for given index sets $\mathcal{I}$ and $\mathcal{J}$: $F_{\mathcal{I},\mathcal{J}}(v) := \max\big(\max_{i\in\mathcal{I}} |v_i|, \max_{i\in\mathcal{J}}(v_i)_+\big)$.

Theorem 5.7. *Given a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns and the index set $\mathcal{S} \subset \{1,\ldots,N\}$, let $M := A^T_{\bullet\mathcal{S}} A_{\bullet\mathcal{S}}$, $\mathcal{S}_1 := \mathcal{S} \cap \mathcal{I}_1$, and $\mathcal{S}_+ := \mathcal{S} \cap \mathcal{I}_+$. Then every vector $z \in \mathcal{P}$ with supp$(z) = \mathcal{S}$ is recovered from $y = Az$ via constrained matching pursuit if the following conditions hold:*

(i) $A_{\bullet\mathcal{S}}$ has full column rank or equivalently M is positive definite; and

(ii) For any (possibly empty) index sets $\mathcal{L}_1 \subset \mathcal{S}_1$ and $\mathcal{L}_+ \subset \mathcal{S}_+$, letting $\widetilde{\mathcal{L}} := \mathcal{L}_1 \cup \mathcal{L}_+$,

$$
\begin{aligned}
& F_{\mathcal{S}_1\setminus\mathcal{L}_1,\, \mathcal{S}_+\setminus\mathcal{L}_+}\left(M/M_{\widetilde{\mathcal{L}}\widetilde{\mathcal{L}}} \begin{pmatrix} v_{\mathcal{S}_1\setminus\mathcal{L}_1} \\ v_{\mathcal{S}_+\setminus\mathcal{L}_+} \end{pmatrix} \right) \\
> \ & F_{\mathcal{S}^c\cap\mathcal{I}_1,\, \mathcal{S}^c\cap\mathcal{I}_+}\left(\big(A^T_{\bullet\mathcal{S}^c}[I - A^T_{\bullet\widetilde{\mathcal{L}}}(A^T_{\bullet\widetilde{\mathcal{L}}}A_{\bullet\widetilde{\mathcal{L}}})^{-1}A_{\bullet\widetilde{\mathcal{L}}}]A_{\bullet\mathcal{S}\setminus\widetilde{\mathcal{L}}} \begin{pmatrix} v_{\mathcal{S}_1\setminus\mathcal{L}_1} \\ v_{\mathcal{S}_+\setminus\mathcal{L}_+} \end{pmatrix} \right)
\end{aligned}
$$

for all $v_{\mathcal{S}_+\setminus\mathcal{L}_+} > 0$ and all $v_{\mathcal{S}_1\setminus\mathcal{L}_1}$ whose each element is nonzero.

Proof. Let $\mathcal{J} \subset \mathcal{S}$ be a nonempty index set. Since $\mathcal{P} = \mathbb{R}_{\mathcal{I}_1} \times (\mathbb{R}_+)_{\mathcal{I}_+}$ is a closed convex cone, it follows from the discussions at the end of Section 2 that the necessary and sufficient optimality condition for an optimal solution $x^* = (x^*_{\mathcal{J}}, 0)$ of the underlying minimization problem $\min_{w\in\mathcal{P},\mathrm{supp}(w)\subseteq\mathcal{J}} \|Aw - Az\|_2^2$ is given by: $\mathcal{C} \in x^*_{\mathcal{J}} \perp A^T_{\bullet\mathcal{J}}(A_{\bullet\mathcal{J}}x^*_{\mathcal{J}} - Az) \in \mathcal{C}^*$, where $z \in \mathcal{P}$ is such that supp$(z) = \mathcal{S}$, the convex cone $\mathcal{C} := \{w_{\mathcal{J}} \,|\, (w_{\mathcal{J}}, 0) \in \mathcal{P}\} = \mathbb{R}_{\mathcal{I}_1\cap\mathcal{J}} \times (\mathbb{R}_+)_{\mathcal{I}_+\cap\mathcal{J}}$ and the dual cone $\mathcal{C}^*$ is given by $\mathcal{C}^* = \{0\} \times (\mathbb{R}_+)_{\mathcal{I}_+\cap\mathcal{J}}$. Hence, we have that

$$
A^T_{\bullet\mathcal{I}_1\cap\mathcal{J}}(A_{\bullet\mathcal{J}}x^*_{\mathcal{J}} - Az) = A^T_{\bullet\mathcal{I}_1\cap\mathcal{J}}A(x^* - z) = 0,
$$

where $(\mathcal{I}_1 \cap \mathcal{J}) \subset \mathcal{S}_1$, and

$$
0 \le x^*_{\mathcal{I}_+\cap\mathcal{J}} \perp A^T_{\bullet\mathcal{I}_+\cap\mathcal{J}}(A_{\bullet\mathcal{J}}x^*_{\mathcal{J}} - Az) \ge 0,
$$

where $x^*_{\mathcal{J}} = (x^*_{\mathcal{I}_1\cap\mathcal{J}}, x^*_{\mathcal{I}_+\cap\mathcal{J}})$ with $x^*_{\mathcal{I}_+\cap\mathcal{J}} \ge 0$. Let the index set $\mathcal{L}_+ := \{i \in \mathcal{I}_+ \cap \mathcal{J} \,|\, x^*_i > 0\}$. Thus $\mathcal{L}_+ \subset \mathcal{S}_+$ and $A^T_{\bullet\mathcal{L}_+}A(x^* - z) = 0$. Set $\mathcal{L}_1 := \mathcal{I}_1 \cap \mathcal{J}$, and $\widetilde{\mathcal{L}} := \mathcal{L}_1 \cup \mathcal{L}_+$. Hence, $\mathcal{L}_1$ and $\mathcal{L}_+$ are disjoint subsets of $\mathcal{S}$ with $A^T_{\bullet\widetilde{\mathcal{L}}}A(z - x^*) = 0$. Further, $x^*_{\mathcal{S}\setminus\widetilde{\mathcal{L}}} = 0$. Hence, $A^T_{\bullet\mathcal{S}\setminus\widetilde{\mathcal{L}}}A(z - x^*) = M/M_{\widetilde{\mathcal{L}}\widetilde{\mathcal{L}}}(z - x^*)_{\mathcal{S}\setminus\widetilde{\mathcal{L}}} = M/M_{\widetilde{\mathcal{L}}\widetilde{\mathcal{L}}}\, z_{\mathcal{S}\setminus\widetilde{\mathcal{L}}}$, and $A^T_{\bullet\mathcal{S}^c}A(z - x^*) = A^T_{\bullet\mathcal{S}^c}[I - A^T_{\bullet\widetilde{\mathcal{L}}}(A^T_{\bullet\widetilde{\mathcal{L}}}A_{\bullet\widetilde{\mathcal{L}}})^{-1}A_{\bullet\widetilde{\mathcal{L}}}]A_{\bullet\mathcal{S}\setminus\widetilde{\mathcal{L}}}\, z_{\mathcal{S}\setminus\widetilde{\mathcal{L}}}$. Since $\mathcal{S}$ is a disjoint union of $\mathcal{S}_1$ and $\mathcal{S}_+$, $z_{\mathcal{S}\setminus\widetilde{\mathcal{L}}} = (z_{\mathcal{S}_1\setminus\mathcal{L}_1}, z_{\mathcal{S}_+\setminus\mathcal{L}_+})$, where $z_{\mathcal{S}_+\setminus\mathcal{L}_+} > 0$ and each element of $z_{\mathcal{S}_1\setminus\mathcal{L}_1}$ is nonzero. Further,

$$
\max\Big(\max_{j\in\mathcal{S}_1\setminus\mathcal{J}} |A^T_{\bullet j}A(z - x^*)|, \max_{j\in\mathcal{S}_+\setminus\mathcal{J}} [A^T_{\bullet j}A(z - x^*)]_+ \Big) \ = \ F_{\mathcal{S}_1\setminus\mathcal{L}_1,\, \mathcal{S}_+\setminus\mathcal{L}_+}\Big(M/M_{\widetilde{\mathcal{L}}\widetilde{\mathcal{L}}}\, z_{\mathcal{S}\setminus\widetilde{\mathcal{L}}} \Big),
$$

and

$$
\begin{aligned}
& \max\Big(\max_{j\in\mathcal{S}^c\cap\mathcal{I}_1} |A^T_{\bullet j}A(z - x^*)|, \max_{j\in\mathcal{S}^c\cap\mathcal{I}_+} [A^T_{\bullet j}A(z - x^*)]_+ \Big) \\
= \ & F_{\mathcal{S}^c\cap\mathcal{I}_1,\, \mathcal{S}^c\cap\mathcal{I}_+}\Big(\big(A^T_{\bullet\mathcal{S}^c}[I - A^T_{\bullet\widetilde{\mathcal{L}}}(A^T_{\bullet\widetilde{\mathcal{L}}}A_{\bullet\widetilde{\mathcal{L}}})^{-1}A_{\bullet\widetilde{\mathcal{L}}}]A_{\bullet\mathcal{S}\setminus\widetilde{\mathcal{L}}}\, z_{\mathcal{S}\setminus\widetilde{\mathcal{L}}} \Big).
\end{aligned}
$$

Consequently, under the condition (ii), condition (**H**) holds, leading to the exact vector recovery. □

6 Sufficient Conditions for Uniform Exact Recovery on Convex, CP Admissible Sets via Constrained Matching Pursuit

In this section, we derive sufficient conditions for uniform exact support and vector recovery via constrained matching pursuit using the restricted isometry-like and restricted orthogonality-like constants. For this purpose, we introduce the following constants.

Definition 6.1. *For a given (possible non-CP admissible) set $\mathcal{P}$, a matrix $A \in \mathbb{R}^{m\times N}$, and disjoin index sets $\mathcal{S}_1, \mathcal{S}_+, \mathcal{S}_-$ whose union is $\{1, \dots, N\}$, we say that*

(i) A real number δ is of Property RI on $\mathcal{P}$ *if $0 < \delta < 1$ and $(1-\delta)\cdot\|u-v\|_2^2 \le \|A(u-v)\|_2^2$ for all $u, v \in \Sigma_K \cap \mathcal{P}$ with supp$(v) \subset$ supp(u);*

(ii) A real number θ is of Property RO on $\mathcal{P}$ corresponding to $\mathcal{S}_1, \mathcal{S}_+, \mathcal{S}_-$ *if $\theta > 0$ and for all $u, v \in \Sigma_K \cap \mathcal{P}$ with supp$(v) \subset$ supp(u), the following holds:*

$$\max\Big(\max_{j\in[supp(u)]^c\cap\mathcal{S}_1} |\langle A(u-v), A_{\bullet j}\rangle|, \max_{j\in[supp(u)]^c\cap\mathcal{S}_+} \langle A(u-v), A_{\bullet j}\rangle_+, \\ \max_{j\in[supp(u)]^c\cap\mathcal{S}_-} \langle A(u-v), A_{\bullet j}\rangle_-\Big) \ \le\ \theta\cdot\|u-v\|_2.$$

We also denote these two constants by $\delta_{K,\mathcal{P}}$ and $\theta_{K,\mathcal{P}}$ respectively to emphasize their dependence on $\mathcal{P}$.

When $\mathcal{P} = \mathbb{R}^N$, the constant $\delta_{K,\mathcal{P}}$ resembles the restricted isometry constant, and the constant $\theta_{K,\mathcal{P}}$ is closely related to the $(K, 1)$-restricted orthogonality constant [10, Definition 6.4].

6.1 Cone Case

We first consider the case where $\mathcal{P}$ is an irreducible, closed, convex and CP admissible cone; see Definition 4.2 for the irreducibility. It follows from Proposition 4.2 that $\mathcal{P}$ is a Cartesian product of copies of $\mathbb{R}, \mathbb{R}_+$ and $\mathbb{R}_-$, i.e., $\mathcal{P} = \mathbb{R}_{\mathcal{I}_1}\times(\mathbb{R}_+)_{\mathcal{I}_+}\times(\mathbb{R}_-)_{\mathcal{I}_-}$, where $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$ form a disjoint union of $\{1, \dots, N\}$. The following theorem gives a sufficient condition for condition (**H**) on $\mathcal{P}$, and thus for the exact support recovery on $\mathcal{P}$, in terms of the constants $\theta_{K,\mathcal{P}}$ and $\theta_{K,\mathcal{P}}$ introduced in Definition 6.1.

Theorem 6.1. *Let $\mathcal{P} = \mathbb{R}_{\mathcal{I}_1}\times(\mathbb{R}_+)_{\mathcal{I}_+}\times(\mathbb{R}_-)_{\mathcal{I}_-}$, where the index sets $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$ form a disjoint union of $\{1, \dots, N\}$, and let $A \in \mathbb{R}^{m\times N}$ be a matrix with unit columns. Suppose there exist constants $\delta_{K,\mathcal{P}}$ of Property RI on $\mathcal{P}$ and $\theta_{K,\mathcal{P}}$ of Property RO on $\mathcal{P}$ corresponding to $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$ such that*

$$1 - \delta_{K,\mathcal{P}} > \sqrt{K}\cdot\theta_{K,\mathcal{P}}. \tag{26}$$

Then condition (**H**) *given by (9) holds on $\mathcal{P}$.*

Proof. Given any $0 \ne u \in \Sigma_K \cap \mathcal{P}$ and any index set $\mathcal{J} \subset \text{supp}(u)$, let v be an arbitrary optimal solution to $\min_{w\in\mathcal{P},\ \text{supp}(w)\subseteq\mathcal{J}} \|A(u-w)\|_2^2$. Hence, for each $j \in \text{supp}(u)\setminus\mathcal{J}$, either $j \in \mathcal{I}_1$, $j \in \mathcal{I}_+$ or $j \in \mathcal{I}_-$. For any $j \in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_1$, we have $\mathbb{I}_j(v) = \mathbb{R}$, where $\mathbb{I}_j(v)$ is defined in (2). For any $j \in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_+$, it follows from $j \notin \mathcal{J}$ that $v_j = 0$ and $\mathbb{I}_j(v) = \mathbb{R}_+$. Similarly, for any $j \in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_-$, we have $\mathbb{I}_j(v) = \mathbb{R}_-$. Further, in light of $\|A_{\bullet j}\|_2 = 1, \forall\, j$ and the expressions for $f_j^*(u, v)$ given below (4), we have

$$\begin{aligned} f_j^*(u,v) &= \|A(u-v)\|_2^2 - |\langle A(u-v), A_{\bullet j}\rangle|^2, && \forall\, j \in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_1, \\ f_j^*(u,v) &= \|A(u-v)\|_2^2 - [\langle A(u-v), A_{\bullet j}\rangle_+]^2, && \forall\, j \in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_+, \\ f_j^*(u,v) &= \|A(u-v)\|_2^2 - [\langle A(u-v), A_{\bullet j}\rangle_-]^2, && \forall\, j \in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_-. \end{aligned}$$

Define the following quantities:

$$\begin{aligned} \Gamma_1 &:= \max\Big(\max_{j\in[\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_1} |\langle A(u-v), A_{\bullet j}\rangle|, \max_{j\in[\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_+} \langle A(u-v), A_{\bullet j}\rangle_+, \\ &\qquad \max_{j\in[\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_-} \langle A(u-v), A_{\bullet j}\rangle_-\Big), \qquad (27) \\ \Gamma_2 &:= \max\Big(\max_{j\in[\text{supp}(u)]^c\cap\mathcal{I}_1} |\langle A(u-v), A_{\bullet j}\rangle|, \max_{j\in[\text{supp}(u)]^c\cap\mathcal{I}_+} \langle A(u-v), A_{\bullet j}\rangle_+, \\ &\qquad \max_{j\in[\text{supp}(u)]^c\cap\mathcal{I}_-} \langle A(u-v), A_{\bullet j}\rangle_-\Big). \end{aligned}$$

Note that if $\Gamma_1 > \Gamma_2$, then $\min_{j\in \text{supp}(u)\setminus\mathcal{J}} f_j^*(u,v) < \min_{j\in[\text{supp}(u)]^c} f_j^*(u,v)$ such that condition (**H**) given by (9) holds. Hence, it suffices to show that $\Gamma_1 > \Gamma_2$ as follows.

By virtue of the definition of the constant $\theta_{K,\mathcal{P}}$ corresponding to $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$, we deduce that $\Gamma_2 \le \theta_{K,\mathcal{P}} \cdot \|u-v\|_2$. Besides, in view of Proposition 4.5 and the definition of Γ_1 in (27), we have

$$\begin{aligned}
\|A(u-v)\|_2^2 &\le \sum_{j\in \text{supp}(u)\setminus\mathcal{J}} \langle A(u-v), A_{\bullet j}\rangle \cdot (u-v)_j \\
&= \sum_{j\in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_1} \langle A(u-v), A_{\bullet j}\rangle \cdot (u-v)_j + \sum_{j\in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_+} \langle A(u-v), A_{\bullet j}\rangle \cdot (u-v)_j \\
&\quad + \sum_{j\in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_-} \langle A(u-v), A_{\bullet j}\rangle \cdot (u-v)_j \\
&\le \sum_{j\in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_1} |\langle A(u-v), A_{\bullet j}\rangle| \cdot |(u-v)_j| + \sum_{j\in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_+} \langle A(u-v), A_{\bullet j}\rangle_+ \cdot (u-v)_j \\
&\quad + \sum_{j\in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_-} \langle A(u-v), A_{\bullet j}\rangle_- \cdot |(u-v)_j| \\
&\le \Gamma_1 \cdot \big\|(u-v)_{\text{supp}(u)\setminus\mathcal{J}}\big\|_1 \le \Gamma_1 \cdot \sqrt{|\text{supp}(u)\setminus\mathcal{J}|} \cdot \|u-v\|_2, \\
&\le \Gamma_1 \cdot \sqrt{K} \cdot \|u-v\|_2,
\end{aligned}$$

where the second inequality follows from the fact that $u_j > 0 = v_j$ for each $j \in [\text{supp}(u)\setminus\mathcal{J}]\cap\mathcal{I}_+$. Therefore, by the definition of the constant $\delta_{K,\mathcal{P}}$, we have

$$(1-\delta_{K,\mathcal{P}}) \cdot \|u-v\|_2^2 \le \|A(u-v)\|_2^2 \le \Gamma_1 \cdot \sqrt{K} \cdot \|u-v\|_2.$$

Since $\text{supp}(v) \subset \text{supp}(u)$, we have $\|u-v\|_2 > 0$. This further implies that $[(1-\delta_{K,\mathcal{P}})/\sqrt{K}] \cdot \|u-v\|_2 \le \Gamma_1$. Using $\Gamma_2 \le \theta_{K,\mathcal{P}} \cdot \|u-v\|_2$ and the assumption that $1-\delta_{K,\mathcal{P}} > \sqrt{K}\cdot\theta_{K,\mathcal{P}}$ given in (26), we obtain $\Gamma_1 > \Gamma_2$. As a result, condition (**H**) holds. $\square$

Since $\theta_{K,\mathcal{P}}$ and $\theta_{K,\mathcal{P}}$ may be difficult to find numerically due to the conditions such as $\text{supp}(v) \subset \text{supp}(u)$ in their definitions, it is desired that similar constants independent of the above mentioned conditions can be used. This leads to the following quantities.

Definition 6.2. *Let a matrix $A \in \mathbb{R}^{m\times N}$ with unit columns and the index sets $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$ which form a disjoint union of $\{1,\dots,N\}$ be given.*

(i) The constant $\widehat{\delta}_K \in (0,1)$ is such that $(1-\widehat{\delta}_K)\cdot\|x\|_2^2 \le \|Ax\|_2^2$ for all $x \in \Sigma_K$;

(ii) The constant $\widehat{\theta}_K > 0$ corresponding to the index set $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$ is such that for any $x \in \Sigma_K$,

$$\max\Big(\max_{j\in\mathcal{I}_1} |\langle Ax, A_{\bullet j}\rangle|,\ \max_{j\in\mathcal{I}_+} \langle Ax, A_{\bullet j}\rangle_+,\ \max_{j\in\mathcal{I}_-} \langle Ax, A_{\bullet j}\rangle_- \Big) \le \widehat{\theta}_K \cdot \|x\|_2.$$

To emphasize the dependence of the above constants on A (when $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$ are fixed), we also write them as $\widehat{\delta}_K(A)$ and $\widehat{\theta}_K(A)$, respectively.

Based on Definition 6.2, it is easy to see that $\widehat{\delta}_K$ is of Property RI and $\widehat{\theta}_K$ is of Property RO, both on $\mathcal{P}$. Hence, by Theorem 6.1, we obtain the following corollary immediately; its proof is omitted.

Corollary 6.1. *For a given matrix $A \in \mathbb{R}^{m\times N}$ with unit columns and a closed, convex, and CP admissible cone $\mathcal{P}$ defined by the index sets $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$, if there exist positive constants $\widehat{\delta}_K$ and $\widehat{\theta}_K$ given by Definition 6.2 such that $1-\widehat{\delta}_K > \sqrt{K}\cdot\widehat{\theta}_K$, then condition (**H**) given by (9) holds.*

In what follows, we discuss the constants $\widehat{\delta}_K$ and $\widehat{\theta}_K$ subject to perturbations of A.

Proposition 6.1. *Let a matrix $A^\diamond \in \mathbb{R}^{m\times N}$ be such that there exist constants $\widehat{\delta}_K(A^\diamond) \in (0,1)$ and $\widehat{\theta}_K(A^\diamond) > 0$ satisfying $1 - \widehat{\delta}_K(A^\diamond) > \sqrt{K}\cdot\widehat{\theta}_K(A^\diamond)$. Then there exists a constant $\eta > 0$ such that for any A with $\|A - A^\diamond\|_2 < \eta$, there exist constants $\widehat{\delta}_K(A) > 0$ and $\widehat{\theta}_K(A) > 0$ satisfying the conditions given by Definition 6.2 such that $1 - \widehat{\delta}_K(A) > \sqrt{K}\cdot\widehat{\theta}_K(A)$.*

Proof. For the given matrix $A^\diamond$ and the positive constants $\widehat{\delta}_K(A^\diamond)$ and $\widehat{\theta}_K(A^\diamond)$, it suffices to show that for any $\varepsilon > 0$, there exist constants $\eta' > 0$ and $\eta'' > 0$ such that (i) for each A with $\|A - A^\diamond\|_2 < \eta'$, there exists a constant $\widehat{\delta}_K(A) > 0$ satisfying condition (i) of Definition 6.2 such that $|\widehat{\delta}_K(A) - \widehat{\delta}_K(A^\diamond)| < \varepsilon$; and (ii) for each A with $\|A - A^\diamond\|_2 < \eta''$, there exists a constant $\widehat{\theta}_K(A) > 0$ satisfying condition (ii) of Definition 6.2 such that $|\widehat{\theta}_K(A) - \widehat{\theta}_K(A^\diamond)| < \varepsilon$.

To show the existence of η', we use the inequality $\big|\|Ax\|_2 - \|A^\diamond x\|_2\big| \le \|A - A^\diamond\|_2\cdot\|x\|_2$ for any A and x [22, Proposition 5.3]. Hence, for all A in the neighborhood $\mathcal{U}$ of $A^\diamond$ given by $\mathcal{U} = \{A \,|\, \|A - A^\diamond\|_2 < \alpha\}$ for some $\alpha > 0$, we have $\big|\|Ax\|_2^2 - \|A^\diamond x\|_2^2\big| = \big|\|Ax\|_2 - \|A^\diamond x\|_2\big|\cdot(\|Ax\|_2 + \|A^\diamond x\|_2) \le \|A - A^\diamond\|_2\cdot\|x\|_2\cdot(2\|A^\diamond\|_2 + \alpha)\cdot\|x\|_2 \le c'\cdot\|A - A^\diamond\|_2\cdot\|x\|_2^2$ for all x, where $c' := 2\|A^\diamond\|_2 + \alpha > 0$. Hence, $\|Ax\|_2^2 \ge \|A^\diamond x\|_2^2 - c'\cdot\|A - A^\diamond\|_2\cdot\|x\|_2^2 \ge [1 - \widehat{\delta}_K(A^\diamond) - c'\cdot\|A - A^\diamond\|_2]\cdot\|x\|_2^2$ for all x. Letting $\widehat{\delta}_K(A) := \widehat{\delta}_K(A^\diamond) + c'\cdot\|A - A^\diamond\|_2$, we can obtain a positive constant η' with $0 < \eta' < \min(\varepsilon/c', \alpha)$ such that for each A with $\|A - A^\diamond\|_2 < \eta'$, $|\widehat{\delta}_K(A) - \widehat{\delta}_K(A^\diamond)| < \varepsilon$.

To show the existence of η'', define the function h_j for a fixed index j and a matrix A:

$$h_j(A, x) := \begin{cases} |\langle Ax, A_{\bullet j}\rangle|, & \text{if } j \in \mathcal{I}_1; \\ \langle Ax, A_{\bullet j}\rangle_+, & \text{if } j \in \mathcal{I}_+; \\ \langle Ax, A_{\bullet j}\rangle_-, & \text{if } j \in \mathcal{I}_-. \end{cases}$$

Using the fact that $|x_+ - y_+| \le |x - y|$ and $|x_- - y_-| \le |x - y|$ for any $x, y \in \mathbb{R}$, we have, for each j,

$$\begin{aligned} |h_j(A,x) - h_j(A^\diamond, x)| &\le |\langle Ax, A_{\bullet j}\rangle - \langle A^\diamond x, A^\diamond_{\bullet j}\rangle| \\ &= \Big|\langle A^\diamond x, (A - A^\diamond)_{\bullet j}\rangle + \langle (A - A^\diamond)x, A^\diamond_{\bullet j}\rangle + \langle (A - A^\diamond)x, (A - A^\diamond)_{\bullet j}\rangle\Big| \\ &\le |\langle A^\diamond x, (A - A^\diamond)\mathbf{e}_j\rangle| + |\langle (A - A^\diamond)x, A^\diamond\mathbf{e}_j\rangle| + |\langle (A - A^\diamond)x, (A - A^\diamond)\mathbf{e}_j\rangle| \\ &\le \|A - A^\diamond\|_2\cdot\big[2\|A^\diamond\|_2 + \|A - A^\diamond\|_2\big]\cdot\|x\|_2, \end{aligned}$$

where the last inequality follows from Cauchy-Schwarz inequality and $\|\mathbf{e}_j\|_2 = 1$. Therefore, for all A in the neighborhood $\mathcal{U}$ of $A^\diamond$ given by $\mathcal{U} = \{A \,|\, \|A - A^\diamond\|_2 < \beta\}$ for some $\beta > 0$, we obtain the constant $c := 2\|A^\diamond\|_2 + \beta > 0$ such that for each j, $h_j(A,x) \le h_j(A^\diamond, x) + c\cdot\|A - A^\diamond\|_2\cdot\|x\|_2$. In view of

$$\max_j h_j(A,x) = \max\Big(\max_{j\in\mathcal{I}_1}|\langle Ax, A_{\bullet j}\rangle|,\ \max_{j\in\mathcal{I}_+}\langle Ax, A_{\bullet j}\rangle_+,\ \max_{j\in\mathcal{I}_-}\langle Ax, A_{\bullet j}\rangle_-\Big),$$

we further have

$$\begin{aligned} \max_j h_j(A,x) &\le \max_j h_j(A^\diamond, x) + c\cdot\|A - A^\diamond\|_2\cdot\|x\|_2 \le \widehat{\theta}_K(A^\diamond)\cdot\|x\|_2 + c\cdot\|A - A^\diamond\|_2\cdot\|x\|_2 \\ &\le \big[\widehat{\theta}_K(A^\diamond) + c\cdot\|A - A^\diamond\|_2\big]\cdot\|x\|_2. \end{aligned}$$

By letting $\widehat{\theta}_K(A) := \widehat{\theta}_K(A^\diamond) + c\cdot\|A - A^\diamond\|_2$, it is easy to obtain a positive constant η'' with $0 < \eta'' < \min(\varepsilon/c, \beta)$ such that for each A with $\|A - A^\diamond\|_2 < \eta''$, $|\widehat{\theta}_K(A) - \widehat{\theta}_K(A^\diamond)| < \varepsilon$. ∎

Remark 6.1. The above proposition shows that for fixed index sets $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$, $\mathcal{A} := \{A \in \mathbb{R}^{m\times N} \,|\, 1 - \widehat{\delta}_K(A) > \sqrt{K}\cdot\widehat{\theta}_K(A)\}$ is an open set in the matrix space $\mathbb{R}^{m\times N}$. Since the set of matrices of completely full rank, i.e., $A \in \mathbb{R}^{m\times N}$ is such that every $m\times m$ submatrix of A is invertible [22], is open and dense in the matrix space $\mathbb{R}^{m\times N}$, we conclude that for any $A \in \mathcal{A}$ and an arbitrarily small $\varepsilon > 0$, there exists a matrix $A' \in \mathcal{A}$ of complete full rank such that $\|A' - A\| < \varepsilon$. An advantage of using the matrix A' is that it leads to a unique x^k in each step (cf. Lemma 2.1) and thus gives rise to the exact vector recovery, provided that the sparsity level $K \le m$.

6.2 Non-cone Case

In this subsection, we consider the case where an irreducible, convex and CP admissible set $\mathcal{P}$ is not a cone. We exploit the positive homogeneous property of the functions used to characterize the two constants $\delta_{K,\mathcal{P}}$ and $\theta_{K,\mathcal{P}}$ in Definition 6.1 and obtain sufficient conditions for exact support recovery on $\mathcal{P}$. Toward this end, we recall that a function $f : \mathbb{R}^N \to \mathbb{R}$ is positively homogeneous of degree $p \in \mathbb{N}$ if for any $\lambda \geq 0$, $f(\lambda x) = \lambda^p \cdot f(x)$ for all $x \in \mathbb{R}^N$. We start from a technical lemma.

Lemma 6.1. *Let $\mathcal{P}$ be a convex set in $\mathbb{R}^N$ containing the zero vector, $g : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ be a positively homogeneous function of degree p, and the set $\mathcal{K} \subset \mathbb{R}^N \times \mathbb{R}^N$ be such that $(0,0) \in \mathcal{K}$ and $\mathcal{K} = \lambda\mathcal{K}$ for any $\lambda > 0$. Then $g(u,v) \leq 0$ holds for all $(u,v) \in (\mathcal{P} \times \mathcal{P}) \cap \mathcal{K}$ if and only if $g(x,y) \leq 0$ holds for all $(x,y) \in (cone(\mathcal{P}) \times cone(\mathcal{P})) \cap \mathcal{K}$.*

Proof. Since $\mathcal{P}$ is a subset of cone($\mathcal{P}$), the "if" part holds trivially. To show the "only if" part, suppose $g(u,v) \leq 0$ holds for all $(u,v) \in (\mathcal{P} \times \mathcal{P}) \cap \mathcal{K}$. Since $\mathcal{P}$ is convex, we have $\text{cone}(\mathcal{P}) = \{\lambda x \,|\, x \in \mathcal{P}, \lambda \geq 0\}$ [21, Corollary 2.6.3]. Hence, for any $(x,y) \in (\text{cone}(\mathcal{P}) \times \text{cone}(\mathcal{P})) \cap \mathcal{K}$, there exist (possibly distinct) real numbers $\alpha, \beta \in \mathbb{R}_+$ and $(u,v) \in (\mathcal{P} \times \mathcal{P}) \cap \mathcal{K}$ such that $x = \alpha u$ and $y = \beta v$. We claim that there exist a pair $(\widehat{u}, \widehat{v}) \in (\mathcal{P} \times \mathcal{P}) \cap \mathcal{K}$ and a positive constant λ such that $(x,y) = \lambda \cdot (\widehat{u}, \widehat{v})$. We show this claim for four possible cases as follows:

(a) $x = y = 0$. Then we choose $\widehat{u} = \widehat{v} = 0$ and any $\lambda > 0$, using the fact that $0 \in \mathcal{P}$ and $(0,0) \in \mathcal{K}$.

(b) $x \neq 0$ and $y = 0$. This implies that α must be positive. Since $(x,y) = (\alpha u, 0) \in \mathcal{K}$ and $\mathcal{K} = \lambda\mathcal{K}$ for any $\lambda > 0$, we have $(u,0) = (1/\alpha)(x,y) \in \mathcal{K}$. Further, since $\mathcal{P}$ contains the zero vector, we have $(u,0) \in (\mathcal{P} \times \mathcal{P}) \cap \mathcal{K}$. Therefore, by letting $(\widehat{u}, \widehat{v}) = (u,0)$ and $\lambda = \alpha > 0$, the desired result holds.

(c) $x = 0$ and $y \neq 0$. This follows readily by interchanging the roles of x and y in case (b).

(d) $x \neq 0$ and $y \neq 0$. In this case, both $\alpha > 0$ and $\beta > 0$. Without loss of generality, we assume that $\alpha \geq \beta$. Since $0 < \beta/\alpha \leq 1$ and $\mathcal{P}$ is a convex set containing the zero vector and v, we see that the vector $\widehat{v} := (\beta/\alpha)v$ belongs to $\mathcal{P}$. Hence, $y = \alpha\widehat{v}$ such that $(x,y) = \alpha(u, \widehat{v}) \in \mathcal{K}$. Letting $\widehat{u} := u$ and $\lambda = \alpha > 0$, we have $(\widehat{u}, \widehat{v}) \in (\mathcal{P} \times \mathcal{P}) \cap \mathcal{K}$ and $(x,y) = \lambda \cdot (\widehat{u}, \widehat{v})$.

In light of the above claim, we deduce that for any $(x,y) \in (\text{cone}(\mathcal{P}) \times \text{cone}(\mathcal{P})) \cap \mathcal{K}$, $g(x,y) = g(\lambda(\widehat{u}, \widehat{v})) = \lambda^p \cdot g(\widehat{u}, \widehat{v}) \leq 0$ by the positive homogeneity of $g(\cdot,\cdot)$. $\square$

Proposition 6.2. *Let $\mathcal{P}$ be a convex set in $\mathbb{R}^N$ containing the zero vector, $A \in \mathbb{R}^{m \times N}$ be a matrix, and the index sets $\mathcal{S}_1, \mathcal{S}_+, \mathcal{S}_-$ form a disjoint union of $\{1, \ldots, N\}$. Then the following hold:*

(i) A real number δ is of Property RI on $\mathcal{P}$ if and only if it is of Property RI on cone($\mathcal{P}$);

(ii) A real number θ is of Property RO on $\mathcal{P}$ corresponding to $\mathcal{S}_1, \mathcal{S}_+, \mathcal{S}_-$ if and only if it is of Property RO on cone($\mathcal{P}$) corresponding to $\mathcal{S}_1, \mathcal{S}_+, \mathcal{S}_-$.

Proof. Define the set $\mathcal{K} := \{(0,0)\} \cup \{(u,v) \in \Sigma_K \times \Sigma_K \,|\, \text{supp}(v) \subset \text{supp}(u)\} \subset \mathbb{R}^N \times \mathbb{R}^N$. It is easy to verify that $(0,0) \in \mathcal{K}$, and $\mathcal{K} = \lambda\mathcal{K}$ for any positive number λ. For any fixed real numbers $\delta \in (0,1)$ and $\theta > 0$, define the functions $g_\delta : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ and $h_\theta : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$:

$$\begin{aligned} g_\delta(u,v) &:= (1-\delta)\|u-v\|_2^2 - \|A(u-v)\|_2^2, \\ h_\theta(u,v) &:= \max\Big(\max_{j \in [\text{supp}(u)]^c \cap \mathcal{S}_1} |\langle A(u-v), A_{\bullet j}\rangle|, \max_{j \in [\text{supp}(u)]^c \cap \mathcal{S}_+} \langle A(u-v), A_{\bullet j}\rangle_+, \\ &\qquad \max_{j \in [\text{supp}(u)]^c \cap \mathcal{S}_-} \langle A(u-v), A_{\bullet j}\rangle_- \Big) - \theta\|(u-v)\|_2. \end{aligned}$$

Clearly, g_δ is positively homogeneous of degree two. Furthermore, when $\lambda = 0$, we see that for any u and v in $\mathbb{R}^N$, $h_\theta(\lambda u, \lambda v) = h_\theta(0,0) = 0 = \lambda h_\theta(u,v)$. Besides, in view of $\text{supp}(u) = \text{supp}(\lambda u)$ for any $\lambda > 0$ and any u, we also have that $h_\theta(\lambda u, \lambda v) = \lambda h_\theta(u,v)$ for any u, v and any $\lambda > 0$. Therefore, h_θ is positively homogeneous of degree one. As a result, we obtain the following equivalent implications:

$$\begin{aligned}
\Big[\delta \text{ is of Property RI on } \mathcal{P}\Big] \quad &\Longleftrightarrow \quad \Big[g_\delta(u,v) \le 0, \ \forall\, (u,v) \in (\mathcal{P} \times \mathcal{P}) \cap \mathcal{K}\Big] \\
&\Longleftrightarrow \quad \Big[g_\delta(u,v) \le 0, \ \forall\, (u,v) \in (\text{cone}(\mathcal{P}) \times \text{cone}(\mathcal{P})) \cap \mathcal{K}\Big] \\
&\Longleftrightarrow \quad \Big[\delta \text{ is of Property RI on cone}(\mathcal{P})\Big],
\end{aligned}$$

where the first and last double implications follow from the definition of δ on $\mathcal{P}$ or cone($\mathcal{P}$) given by Definition 6.1, and the second double implication follow from Lemma 6.1. Similarly, we can show that θ is of Property RO on $\mathcal{P}$ if and only if it is of Property RO on cone($\mathcal{P}$) using h_θ. $\square$

By applying the above proposition and the conic hull of a closed, convex, and CP admissible set given by Proposition 4.4, we obtain sufficient conditions for exact support recovery in the following theorem.

Theorem 6.2. *Let $A \in \mathbb{R}^{m \times N}$ be a matrix with unit columns, and $\mathcal{P}$ be an irreducible, closed, convex, and CP admissible set in $\mathbb{R}^N$ whose conic hull is given by cone($\mathcal{P}$) $= \mathbb{R}_{\mathcal{I}_1} \times (\mathbb{R}_+)_{\mathcal{I}_+} \times (\mathbb{R}_-)_{\mathcal{I}_-}$, where $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$ form a disjoint union of $\{1, \dots, N\}$. Then condition* (**H**) *holds on $\mathcal{P}$ under either one of the following conditions:*

(i) There exist constants $\delta_{K,cone(\mathcal{P})}$ of Property RI and $\theta_{K,cone(\mathcal{P})}$ of Property RO corresponding to $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$ such that $1 - \delta_{K,cone(\mathcal{P})} > \sqrt{K} \cdot \theta_{K,cone(\mathcal{P})}$;

(ii) There exist constants $\delta_{K,\mathcal{P}}$ of Property RI and $\theta_{K,\mathcal{P}}$ of Property RO corresponding to $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$ such that $1 - \delta_{K,\mathcal{P}} > \sqrt{K} \cdot \theta_{K,\mathcal{P}}$.

Proof. (i) It follows from Theorem 6.1 that if $1 - \delta_{K,\text{cone}(\mathcal{P})} > \sqrt{K} \cdot \theta_{K,\text{cone}(\mathcal{P})}$, then condition (**H**) holds on cone($\mathcal{P}$). Since $\mathcal{P}$ is a subset of cone($\mathcal{P}$), condition (**H**) also holds on $\mathcal{P}$.

(ii) Suppose $1 - \delta_{K,\mathcal{P}} > \sqrt{K} \cdot \theta_{K,\mathcal{P}}$ holds for the constants $\delta_{K,\mathcal{P}}$ of Property RI and $\theta_{K,\mathcal{P}}$ of Property RO corresponding to $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$. Since $\mathcal{P}$ is CP admissible, it contains the zero vector. Since $\mathcal{P}$ is also convex, we deduce via Proposition 6.2 that $\delta_{K,\mathcal{P}}$ is a constant of Property RI on cone($\mathcal{P}$) and $\theta_{K,\mathcal{P}}$ is a constant of Property RO on cone($\mathcal{P}$) corresponding to $\mathcal{I}_1, \mathcal{I}_+$ and $\mathcal{I}_-$. Therefore, by Theorem 6.1, condition (**H**) holds on cone($\mathcal{P}$). It thus follows from statement (i) that condition (**H**) holds on $\mathcal{P}$. $\square$

Remark 6.2. Theorem 6.2 gives a sufficient condition for (**H**) and thus exact support recovery on a closed, convex, and CP admissible set by leveraging its conic hull. Despite the simplicity of its proof, Theorem 6.2 provides a potentially effective way to establish the exact support recovery for the following reasons. It is usually difficult to find and compute the constants $\delta_{K,\mathcal{P}}$ and $\theta_{K,\mathcal{P}}$ for a general closed, convex and CP admissible set $\mathcal{P}$. On the other hand, computing the constants $\delta_{k,\text{cone}(\mathcal{P})}$ and $\theta_{k,\text{cone}(\mathcal{P})}$ is easier, due to the simple structure of cone($\mathcal{P}$) as illustrated in Propositions 4.2 and 4.4. Note that the conditions $1 - \delta_{K,\mathcal{P}} > \sqrt{K} \cdot \theta_{K,\mathcal{P}}$ and $1 - \delta_{K,\text{cone}(\mathcal{P})} > \sqrt{K} \cdot \theta_{K,\text{cone}(\mathcal{P})}$ are equivalent in view of Proposition 6.2. Hence, the latter condition in term of cone($\mathcal{P}$) does not lead to conservativeness.

Theorem 6.2 can be extended to a non-CP admissible set as long as the closure of its conic hull is CP admissible. This is shown in the following corollary.

Corollary 6.2. *Let $A \in \mathbb{R}^{m \times N}$ and $\mathcal{P}$ be a closed convex set containing the zero vector. Suppose the closure of cone(P), denoted by $\mathcal{C}$, is CP admissible. Then the following hold:*

(i) If there exist constants $\delta_{K,\mathcal{C}}$ of Property RI on $\mathcal{C}$ and $\theta_{K,\mathcal{C}}$ of Property RO on $\mathcal{C}$ such that $1 - \delta_{K,\mathcal{C}} > \sqrt{K} \cdot \theta_{K,\mathcal{C}}$, then condition (**H**) *holds on $\mathcal{P}$.*

(ii) If $K \leq m$ and the constants $\widehat{\delta}_K(A)$ and $\widehat{\theta}_K(A)$ corresponding to the index sets for $\mathcal{C}$ given in Definition 6.2 are such that $1 - \widehat{\delta}_K(A) > \sqrt{K} \cdot \widehat{\theta}_K(A)$, then there exists a matrix $A' \in \mathbb{R}^{m \times N}$ sufficiently close to A such that the exact vector recovery on $\Sigma_K \cap \mathcal{P}$ is achieved using A'.

Proof. (i) Let $\mathrm{cl}(\cdot)$ denote the closure of a set. It follows directly from the fact that $\mathcal{P} \subseteq \mathcal{C} := \mathrm{cl}(\mathrm{cone}(\mathcal{P}))$ and the similar argument for statement (i) of Theorem 6.2.

(ii) This result follows from Corollary 6.1, statement (i), Proposition 6.1, and Remark 6.1. $\square$

For illustration, consider the set $\mathcal{P} := \{x \in \mathbb{R}^N \mid \|x - \mathbf{e}_1\|_2 \leq 1\}$, which is a convex set containing the zero vector. As indicated right after the proof of Proposition 4.4, $\mathcal{P}$ is not CP admissible but the closure of its conic hull is given by $\mathcal{C} = \mathbb{R}_+ \times \mathbb{R}^{N-1}$ and is thus CP admissible. Suppose $1 - \delta_{K,\mathcal{C}} > \sqrt{K} \cdot \theta_{K,\mathcal{C}}$. Then by Corollary 6.2, condition (**H**) holds on $\mathcal{P}$. Another example is a convex set whose interior contains the zero vector. In this case, the closure of its conic hull is $\mathbb{R}^N$ for which a similar sufficient condition in terms of $\delta_{K,\mathcal{C}}$ and $\theta_{K,\mathcal{C}}$ with $\mathcal{C} = \mathbb{R}^N$ can be established.

Remark 6.3. It is interesting to ask whether the sufficient condition $1 - \delta_{K,\mathcal{P}} > \sqrt{K} \cdot \theta_{K,\mathcal{P}}$ derived in Theorem 6.2 for condition (**H**) can be improved using similar techniques for the cone case given in the proof of Theorem 6.1. In spite of many tries, our efforts show that these techniques do not yield better (i.e., less restrictive) sufficient conditions in terms of $\delta_{K,\mathcal{P}}$ and $\theta_{K,\mathcal{P}}$. Although this finding does not rule out the possibility of the existence of better sufficient conditions in terms of $\delta_{K,\mathcal{P}}$ and $\theta_{K,\mathcal{P}}$ because it only gives certain sufficient conditions, it demonstrates a potential difficulty of further improving the obtained sufficient conditions using the same line of ideas given in the proof of Theorem 6.1. It also justifies the importance of the sufficient conditions in terms of $\delta_{K,\mathrm{cone}(\mathcal{P})}$ and $\theta_{K,\mathrm{cone}(\mathcal{P})}$ in Theorem 6.2.

7 Conclusions

This paper studies the exact support and vector recovery on a constraint set via constrained matching pursuit. We show the exact recovery critically relies on a constraint set, and introduce the class of CP admissible sets. Rich properties of these sets are exploited, and various exact recovery conditions are developed for convex CP admissible cones or sets. Future research includes the exact recovery of constrained sparse vectors subject to noise and errors via constrained matching pursuit.

Acknowledgements. The authors would like to thank Dr. Joel A. Tropp for a helpful discussion on the counterexample given in Section 5.1.

References

[1] S. Bahmanai, B. Raj, and P. Boufounos. Greedy sparsity-constrained optimization. *Journal of Machine Learning Research*, Vol. 14(1), pp. 707–841, 2013.

[2] A. Beck and Y.C. Eldar. Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. *SIAM Journal on Optimization*, Vol. 23(3), pp. 1480–1509, 2013.

[3] A. Beck and N. Hallak. On the minimization over sparse symmetric sets: projections, optimality conditions, and algorithms. *Mathematics of Operations Research*, Vol. 41(1), pp. 196–223, 2015.

[4] A. Bruckstein, M. Elad, and M. Zibulevsky. On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations. *IEEE Transactions on Information Theory*, Vol. 54(11), pp. 4813–4820, 2008.

[5] T. T. Cai and L. Wang. Orthogonal matching pursuit for sparse signal recovery with noise. *IEEE Transactions on Information Theory*, Vol. 57(7), pp. 4680–4689, 2011.

[6] R.W. Cottle, J.-S. Pang, and R.E. Stone. *The Linear Complementarity Problem*, Academic Press Inc., (Cambridge 1992).

[7] W. Dai and O. Milenkovic. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Transactions on Information Theory*, Vol. 55(5), pp. 2230–2249, 2009.

[8] D. L. Donoho, Y. Tsaig, I. Drori, and J. L. Starck. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Transactions on Information Theory*, Vol. 58(2), pp. 1094–1121, 2012.

[9] *Compressed Sensing: Theory and Applications*, edited by Y. C. Eldar and G. Kutyniok. Combridge Unversity Press, 2012.

[10] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Birkhäuser, Basel, 2013.

[11] S. Foucart and D. Koslicki. Sparse recovery by means of nonnegative least squares. *IEEE Signal Processing Letters*, Vol. 21, pp. 498–502, 2014.

[12] Y. Itoh, M.F. Duarte, and M. Parente. Perfect recovery conditions for non-negative sparse modeling. *IEEE Transactions on Signal Processing*, Vol. 65, pp. 69–80, 2017.

[13] S. Kwon, J. Wang, and B. Shim. Multipath matching pursuit. *IEEE Transactions on Information Theory*, Vol. 60(5), pp. 2986–3001, 2014.

[14] F. Locatello, M. Tschannen, G. Ratsch, and M. Jaggi. Greedy algorithms for cone constrained optimization with convergence guarantees. *Advances in Neural Information Processing Systems*, pp. 773–784, 2017.

[15] S. G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, Vol. 41(12), pp. 3397–3415, 1993.

[16] Q. Mo and Y. Shen. A remark on the restricted isometry property in orthogonal matching pursuit. *IEEE Transactions on Information Theory*, Vol. 58(6), pp. 3654–3656, 2012.

[17] S. Mousavi and J. Shen. Solution uniqueness of convex piecewise affine functions based optimization with applications to constrained ℓ_1 minimization. *ESAIM: Control, Optimisation and Calculus of Variations*, in press, 2018. Also available at ArXiv preprint arXiv:1711.05882.

[18] D. Needell and J. A. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, Vol. 26, pp. 301–321, 2009.

[19] D. Needell and R. Vershynin. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. *IEEE Journal of Selected Topics in Signal Processing*, Vol. 4(2), pp. 310–316, 2010.

[20] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. *The 27th IEEE Asilomar Conference on Signals, Systems and Computers*, pp. 40–44, November, 1993.

[21] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[22] J. Shen and S. Mousavi. Least sparsity of p-norm based optimization problems with $p > 1$. *SIAM Journal on Optimization*, Vol. 28(3), pp. 2721–21751, 2018.

[23] J. Shen and T. M. Lebair. Shape restricted smoothing splines via constrained optimal control and nonsmooth Newton's methods. *Automatica*, Vol. 53, pp. 216–224, 2015.

[24] J. Shen and X. Wang. Estimation of monotone functions via P-splines: A constrained dynamical optimization approach. *SIAM Journal on Control and Optimization*, Vol. 49(2), pp. 646–671, 2011.

[25] G. Swirszcz, N. Abe and A. Lozano. Grouped orthogonal matching pursuit for variable selection and prediction. *Advances in Neural Information Processing Systems*, pp. 1150–1158, 2009.

[26] J. A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, Vol. 50(10), pp. 2231–2242, 2004.

[27] J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, Vol. 53(12), pp. 4655–4566, 2007.

[28] J. A. Tropp, A. Gilbert and M. Strauss. Algorithms for simultaneous sparse approximations. Part I: Greedy pursuit. *Signal Processing*, Vol. 86(3), pp. 572–588, 2006.

[29] J. Wang, S. Kwon, and B. Shim. Generalized orthogonal matching pursuit. *IEEE Transactions on Signal Processing*, Vol. 60(12), pp. 6202–6216, 2012.

[30] M. Wang, W. Xu, and A. Tang. A unique "nonnegative" solution to an underdetermined system: From vectors to matrices. *IEEE Transactions on Signal Processing*, Vol. 59(3), pp. 1007–1016, 2011.

[31] J. Wen, Z. Zhou, J. Wang, X. Tang, and Q. Mo. A sharp condition for exact support recovery of sparse signals with orthogonal matching pursuit. *IEEE International Symposium on Information Theory*, pp. 2364–2368, July, 2016.

[32] F. Xu, Z. Lu, and Z. Xu. An efficient optimization approach for cardinality-constrained index tracking problems. *Optimization Methods and Software*, Vol. 31(2), pp. 258–271, 2016.

[33] T. Zhang. On the consistency of feature selection using greedy least squares regression. *Journal of Machine Learning Research*, Vol. 10, pp. 555–568, 2009.

[34] T. Zhang. Sparse recovery with orthogonal matching pursuit under RIP. *IEEE Transactions on Information Theory*, Vol. 57(9), pp. 6215–6221, 2011.