Research Article

Chunyan Liu, Feng Zhang*, Wei Qiu, Chuan Li and Zhenbei Leng

A perturbation analysis based on group sparse representation with orthogonal matching pursuit

https://doi.org/10.1515/jiip-2019-0043
Received June 24, 2019; revised August 14, 2020; accepted September 25, 2020

Abstract: In this paper, by exploiting orthogonal projection matrix and block Schur complement, we extend the study to a complete perturbation model. Based on the block-restricted isometry property (BRIP), we establish some sufficient conditions for recovering the support of the block $K$-sparse signals via block orthogonal matching pursuit (BOMP) algorithm. Under some constraints on the minimum magnitude of the nonzero elements of the block $K$-sparse signals, we prove that the support of the block $K$-sparse signals can be exactly recovered by the BOMP algorithm in the case of $\ell_2$ and $\ell_2/\ell_\infty$ bounded total noise if $A$ satisfies the BRIP of order $K + 1$ with

$$\delta_{K+1} < \frac{1}{\sqrt{K + 1}(1 + e^{(K+1)})^2} + \frac{1}{(1 + e^{(K+1)})^2} - 1.$$ 

In addition, we also show that this is a sharp condition for exactly recovering any block $K$-sparse signal with the BOMP algorithm. Moreover, we also give the reconstruction upper bound of the error between the recovered block-sparse signal and the original block-sparse signal. In the noiseless and perturbed case, we also prove that the BOMP algorithm can exactly recover the block $K$-sparse signal under some constraints on the block $K$-sparse signal and

$$\delta_{K+1} < \frac{2 + \sqrt{2}}{2(1 + e^{(K+1)})^2} - 1.$$ 

Finally, we compare the actual performance of perturbed OMP and perturbed BOMP algorithm in the numerical study. We also present some numerical experiments to verify the main theorem by using the completely perturbed BOMP algorithm.

Keywords: Compressed sensing, block sparsity, perturbation of measurement matrix, block-restricted isometry property, block orthogonal matching pursuit, block Schur complement

MSC 2010: 49N45, 47A55, 94A12

1 Introduction

Compressive Sensing (CS), pioneered by Donoho and Candès, Romberg and Tao [6, 23], is an efficient data acquisition paradigm which has captured lots of attention in different fields, including machine learning, sig-
nal processing, pattern recognition, information theory, image processing, engineering and communication [2, 5, 13, 18, 21].

In essence, a crucial concern in CS is how to recover a sparse or compressible sparse signal from an under-determined system of linear equations. Generally, we consider the model \( y = Ax \), where \( y \in \mathbb{C}^n \) (\( \mathbb{C} \) denotes the complex field) are the observed measurements, \( A \in \mathbb{C}^{n \times m} \) \((n \ll m)\) is a given measurement matrix and \( x \in \mathbb{C}^m \) is an unknown signal that needs to be recovered. In practice, we usually run into the case that the observation vector \( y \) is disturbed by \( v \), and consider the CS model

\[
\hat{y} = Ax + v,
\]

where \( v \in \mathbb{C}^n \) is the additive noise independent of \( x \). However, the measurement matrix is affected by various external environment and internal interference in the compression sampling so that the actual measurement matrix is perturbed more or less, and hence the measurement results have a large deviation. In practical application, the measurement matrix error is often unavoidable, such as the error in analog-to-digital signal conversion and the lattice error in signal sampling. The perturbed case can be realized in source separation [4], remote sensing [12], radar [14] and others. That is to say, not only the observation vector \( y \) is perturbed by \( v \), but also the measurement matrix \( A \) is perturbed by \( E \) in practical applications \((E \in \mathbb{C}^{n \times m} \) is a perturbation matrix\). In such a scenario, there will be an extra multiplicative noise \( Ex \) in (1.1) when the perturbed sensing matrix is \( \hat{A} = A + E \). In 2010, Herman and Strohmer [15] proposed the following completely perturbed \( \ell_1 \) minimization model to recover stably the signal \( x \) from the measurements \( \hat{y} \):

\[
\min \| x \|_1 \quad \text{such that} \quad \| \hat{A}x - \hat{y} \|_2 \leq \epsilon'_{A,K,Y},
\]

where \( \epsilon'_{A,K,Y} \) is the total noise. They showed that the signal \( x \) could be reliably recovered by solving (1.2) under the condition \( \delta_{2K} \leq \sqrt{2} / (1 + \epsilon_{A}^2) - 1 \). In [16], Ince and Nacaroglu established a sufficient condition to guarantee stable recovery of the signal by solving a completely perturbed \( \ell_p \) minimization problem.

However, nonzero elements of some real-world signals may appear in a few fixed blocks [20], which is different from the structure of a conventional sparse signal. The block-sparse signals appear in various applications, such as DNA microarrays [24], equalization of sparse communication channels [26], multiple measurement signals [7] and clustering of data in multiple subspaces [11]. From the mathematical viewpoint, a block-structured signal \( x \) over the block index set \( J = \{d_1, \ldots, d_N\} \) can be modeled as follows:

\[
x = [x_{1} \cdots x_{d_1} x_{d_1+1} \cdots x_{d_1+d_2} \cdots x_{m-d_2+1} \cdots x_{m}]^T,
\]

where \( x[i] \) denotes the \( i \)th block of \( x \), \( d_i \) is the block size for the \( i \)th block. If \( x \) has at most \( K \) nonzero blocks, i.e., \( \| x \|_{2,0} = \sum_{i=1}^{N} I(\| x[i] \|_2) \leq K \), we call such \( x \) block \( K \)-sparse signal. Next, the mixed \( \ell_2/\ell_1 \) norm is defined as \( \| x \|_{2,1} = \sum_{i=1}^{N} \| x[i] \|_2 \); the mixed \( \ell_2/\ell_{\infty} \) norm is \( \| x \|_{2,\infty} = \max(\| x[1] \|_2, \ldots, \| x[N] \|_2) \), and the \( \| x \|_{2,3} \) is replaced by \( \| x \|_2 \), i.e., \( \| x \|_{2,2} = \| x \|_2 \). Especially, the block-sparse signal is converted into a conventional sparse signal when the block size \( d_1 = \cdots = d_N = 1 \). In contrast to the block-sparse signal, the conventional sparse signal will be called sparse signal in this paper.

Block orthogonal matching pursuit (BOMP) is a widely used greedy algorithm which can quickly get a local optimal solution instead of seeking the global optimal solution. The basic idea of the BOMP algorithm is to select the column with the strongest linear correlation between the measurement matrix and the current residual, add all columns of the group to which the column belongs to the label set at a time, and finally determine the support of the block-sparse signal by iteration [8]. Compared with the standard OMP algorithm [27], Eldar and Kuppinger [8] have shown that the BOMP algorithm can reconstruct block-sparse signals better.

For any set \( \Lambda \subset \{1, 2, \ldots, N\} \), let \( \hat{A}[\Lambda] \) denote the submatrix of \( \hat{A} \) that contains only the blocks indexed by \( \Lambda \), and let \( x[\Lambda] \) be the subvector of \( x \) that contains only the blocks indexed by \( \Lambda \). For instance, if \( \Lambda = \{2, 3, 5\} \), we have \( \hat{A}[\Lambda] = [\hat{A}[2], \hat{A}[3], \hat{A}[5]] \) and \( x[\Lambda] = (x[2], x[3], x[5])^T \). Then the BOMP algorithm for \( \hat{A} \) is formally described in Algorithm 1 [25].

In the following, the block-restricted isometry property (BRIP) of sensing matrix \( A \) has been proposed to analyze the recovery performance of algorithms for block-sparse signals [9], that is, a matrix \( A \) satisfies the
exists a matrix perturbation. In this paper, with the BRIP condition, we study exact support recovery of block-sparse signals.

fails to recover recovered by using the BOMP algorithm in Zhou and Wang [31] established a sufficient condition to guarantee exact or stable recovery of $\supp(x) = \{i \mid \|x[i]\|_2 \neq 0\} = \Lambda_k$.

Whenever this causes no confusion, we simply denote $\Lambda_k$ for the completely perturbed model. As there are several types of noise, we only consider bounded noise and $\ell_2/\ell_{\infty}$ bounded noise for the block-sparse compressed sensing of the completely perturbed model, respectively, i.e.,

$$
\|A\hat{x} - \hat{y}\|_2 \leq c'_{A,Y},
$$

$$
\|A^* (A\hat{x} - \hat{y})\|_{2,\infty} \leq c''_{A,Y},
$$

where $c''_{A,Y} = \|A\|_2 (1 + c_A)c'_{A,Y}$ (the proof is relegated to Section 2), and the matrix $A$ can be formed as follows:

$$
A = [\hat{A}_1 \cdots \hat{A}_{d_1} \hat{A}_{d_1+1} \cdots \hat{A}_{d_1+d_2} \cdots \hat{A}_{m-d_k+1} \cdots \hat{A}_m].
$$

To investigate the theoretical analysis of the block-sparse compressed sensing with the BOMP algorithm for the completely perturbed model in (1.3) and (1.4), the restricted isometry property (RIP) for $A$ (see [15]) is extended to the BRIP for $A$ in Section 2 (see [29]).

In recent years, many researchers have been trying to find out the RIP-based condition for guaranteeing the exact or stable recovery of $K$-sparse signals by using the OMP algorithm. For the unperturbed case, if $\delta_{K+1} < 1/\sqrt{K+1}$, the $K$-sparse signal $x$ can be reliably recovered by using the OMP algorithm to solve (1.1) in $K$ iterations; see, e.g., [28]. Mo [22] has proved that $\delta_{K+1} < 1/\sqrt{K+1}$ is a sharp condition for exact or stable recovery of $K$-sparse signal via the OMP algorithm. Especially, Wen and Zhu [32] have shown that $\delta_{K+1} \geq 1/\sqrt{K+1}$ may fail to recover a $K$-sparse signal $x$ with the OMP algorithm in $K$ iterations. Wen, Zhou and Wang [31] established a sufficient condition to guarantee exact or stable recovery of $\supp(x)$ with the OMP algorithm in $K$ iterations. In particular, Liu and Fang have improved the sufficient condition in [19]. In addition, under some constraints on the minimum magnitude of the nonzero block elements of block $K$-sparse signal $x$ (i.e., $\min_{i \in \supp(x)} \|x[i]\|_2$) and assuming $A$ satisfies the BRIP of order $K+1$ with $\delta_{K+1} < 1/\sqrt{K+1}$, Wen, Zhou and Liu [30] have proved that the block $K$-sparse signal can be exactly or stably recovered by using the BOMP algorithm in $K$ iterations. If $K \geq 1$ and $1/\sqrt{K+1} \leq t < 1$, [30] shows that there exists a matrix $A$ satisfying the BRIP with $\delta_{K+1} = t$ and a block $K$-sparse signal $x$ such that the BOMP may fail to recover $x$ in $K$ iterations.

From the applied viewpoint, it is important to recover the block-sparse signal in the setting of complete perturbation. In this paper, with the BRIP condition, we study exact support recovery of block-sparse signals.
for the block-sparse compressed sensing of the completely perturbed model with the BOMP algorithm in complex settings. The main results of this paper are summarized as follows. Firstly, if $A$ satisfies the BRIP of order $K+1$ with

$$\delta_{K+1} < \frac{1}{\sqrt{K+1} \| \varepsilon_{A}^{(K+1)} \|^2} + \frac{1}{(1 + \varepsilon_{A}^{(K+1)})^2} - 1 \tag{1.5}$$

and $\min_{i \in \supp \{ x \}} \| x[i] \|_2$ exceeds a certain lower bound in the block-sparse compressed sensing of the completely perturbed model, we show that the support of the block $K$-sparse signal can be exactly or stably recovered with a certain stopping criterion under either the $\ell_2$ or $\ell_2/\ell_\infty$ bounded total noise. In addition, we provide the reconstruction upper bound of the error between the recovered block-sparse signal and the original block-sparse signal. Secondly, we show that (1.5) is a sharp condition for exactly recovering any block $K$-sparse signal with the BOMP algorithm. Thirdly, when the measurement matrix $A$ is perturbed by $E$ in the noiseless case, we also prove that the BOMP algorithm can exactly recover the block $K$-sparse signal under some constraints on the block $K$-sparse signal and

$$\delta_{K+1} < \frac{2 + \sqrt{2}}{2(1 + \varepsilon_{A}^{(K+1)})^2} - 1.$$ 

Finally, we conduct some numerical experiments to verify Theorem 3.5 by using the block-sparse algorithm with complete perturbation.

The rest of the paper is organized as follows. In Section 2, we give some notations and lemmas. Section 3 gives our main theoretical results. Section 4 provides a series of numerical experiments to support the theoretical results. Finally, we summarize this paper in Section 5.

## 2 Notations and preliminaries

In this section, we first introduce some symbols which will be used in the rest of this paper.

**Notation.** Let $\Omega = \{1, 2, \ldots, k\}$ with $k \in \mathbb{N}^+$ and $\Omega_0 = \emptyset$. For two sets $\Lambda$ and $\Gamma$, let $\Lambda \setminus \Gamma = \{ i \mid i \in \Lambda, i \notin \Gamma \}$. Let $\Lambda = \{ \lambda_1, \lambda_2, \ldots, \lambda_{|\Lambda|} \}$. Let $A[\Lambda]$ be the submatrix of the block matrix $A$ to the index set $\Lambda$. Let $x[\Lambda]$ be the subvector of the block vector $x$ to the index set $\Lambda$. Let $A^T[\Lambda]$ be the transpose of $A[\Lambda]$, and let $A^*[\Lambda]$ denote the conjugate transposition of $A[\Lambda]$, where $|\Lambda|$ is the cardinality of $\Lambda$. Let $H$ denote a subspace of $\mathbb{C}^n$. Let $P[H]$ represent an orthogonal projection block matrix onto $H$. Let $P^*[H] = I_n - P[H]$ be an orthogonal projection block matrix onto the orthogonal complement of $H$, where $I_n$ denotes the identity block matrix. Let $H_A = \text{span}[A[\lambda_1], A[\lambda_2], \ldots, A[|\Lambda|]]$ represent the column space of $A[\Lambda]$, $H_\emptyset = 0$, and let

$$P[H_A] = A[\Lambda][A^*[\Lambda]|A[\Lambda])^{-1} A^*[\Lambda].$$

Let the scalar $\langle x, y \rangle = y^*x$ denote the Euclidean inner product of the complex block vectors $x$ and $y$. For a block matrix $B \in \mathbb{C}^{n \times n}$,

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

where the block $B_{11}$ is a $k$-order invertible block matrix with $0 < k < n$. From [33], letting $C_k(B)$ denote the Schur complement of $B_{11}$ in $B$, we have $C_k(B) = B_{22} - B_{21} B_{11}^{-1} B_{12}$ and $C_0(B) = B$.

Similar to [15], we define some notations which will be used in the rest of this paper. We first quantify the perturbations $E$ and $v$ with the relative bounds

$$\frac{\|E\|_2}{\|A\|_2} \leq c_A, \quad \frac{\|E\|_2^{(K)}}{\|A\|_2^{(K)}} \leq c_A^{(K)}, \quad \frac{\|v\|_2}{\|y\|_2} \leq c_y, \tag{2.1}$$

where $\| \cdot \|_2$ represents the spectral norm and $\| \cdot \|_2^{(K)}$ stands for the largest spectral norm taken over all $K$-block submatrices. We also define the constants

$$y_K = \frac{\|x_K\|_2}{\|x\|_2}, \quad s_K = \frac{\|x_K\|_2,1}{\sqrt{K} \|x\|_2}, \quad a_A = \frac{\|A\|_2}{\sqrt{1 - \delta_K}}, \quad s_A^{(K)} = \frac{\sqrt{1 + \delta_K}}{\sqrt{1 - \delta_K}}.$$

The $x_K$ is the best $K$-term approximation of $x$, and it contains the $K$-largest blocks of $x$ in index set $J$. 
Before presenting our main lemmas, we first give the BRIP [29] of matrix $\tilde{A}$.

**Definition 2.1 (BRIP for $\tilde{A}$).** For the BRIC $\delta_K$ ($K = 1, 2, \ldots$) for $A$ and $e^{(K)}_A$ associated with matrix $E$, fix the constant $\tilde{\delta}_{K,\text{max}} = (1 + \tilde{\delta}_K)(1 + e^{(K)}_A)^2 - 1$. Then the matrix $\tilde{A} = A + E$ satisfies the BRIP of order $K$ if there exists a smallest nonnegative number $\delta_K \leq \tilde{\delta}_{K,\text{max}}$ such that $(1 - \tilde{\delta}_K\|x\|_2^2) \leq \|\tilde{A}x\|_2^2 \leq (1 + \tilde{\delta}_K\|x\|_2^2)$ for all $x \in \mathbb{R}^N$ that are block $K$-sparse over the block index set $\mathcal{I}$. Similar to [15], we have the following inequality: $1 - (1 - \tilde{\delta}_K)(1 - e^{(K)}_A)^2 \leq \delta_K \leq \tilde{\delta}_{K,\text{max}} = (1 + \tilde{\delta}_K)(1 + e^{(K)}_A)^2 - 1$.

Next, we will give some evident properties of the block-sensing matrix $\tilde{A}$, orthogonal projection matrix and the block Schur complement of matrices.

**Lemma 2.2.** For any complex vectors $\alpha, \beta \in \mathbb{C}^n$, it holds that $\langle \tilde{P}[H]\alpha, \tilde{P}[H]\beta \rangle = \langle P[H]\alpha, P[H]\beta \rangle$ and $\|P[H]\alpha\|_2 \leq \|\alpha\|_2$.

**Lemma 2.3.** If the matrix $\tilde{A}$ satisfies the BRIP of orders $K_1$ and $K_2$ with $K_1 < K_2$, we have $\tilde{\delta}_{K_1} \leq \tilde{\delta}_{K_2}$.

**Lemma 2.4.** Let the matrix $\tilde{A}$ satisfy the BRIP of order $K + 1$, and let $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_{|\Lambda|}\} \subseteq \Omega_N$ with $|\Lambda| \leq K + 1$. Then it holds that
\[
(1 - \tilde{\delta}_{K+1})\|\xi\|_2^2 \leq \|\tilde{A}[\Lambda]\tilde{A}[\Lambda]\|_2 \leq (1 + \tilde{\delta}_{K+1})\|\xi\|_2^2,
\]
where $\xi = (\xi[1], \xi[2], \ldots, \xi[|\Lambda|])^T$.

**Lemma 2.5 ([33]).** For any block matrices $B, C$, if $B \succeq C > 0$, then $C_k(B) \geq C_k(C)$.

**Lemma 2.6.** For any block vector $x \in \mathbb{C}^n$, $\|x\|_{2,1,2} \leq \sqrt{m}\|x\|_2$, $\|x\|_2 \leq \sqrt{m}\|x\|_{2,\infty}$ and $\|x\|_{2,\infty} \leq \|x\|_2 \leq \|x\|_{2,1}$.

**Lemma 2.7.** In (1.1), let the matrix $\tilde{A} = A + E$ be given. If the $\epsilon_2$ bounded noise for the block-sparse compressed sensing of the completely perturbed model is such that $\|\tilde{A}x - \hat{y}\|_2 \leq \epsilon_2^{A,K,y}$, we have $\|\tilde{A}^*\tilde{A}x - \hat{y}\|_2 \leq \epsilon_2^{A,K,y}$, where the total noise parameter is
\[
\epsilon_2^{A,K,y} = \frac{\epsilon_A^{K}(1 + \epsilon_A^{(K)})}{1 - \epsilon_A^{(K)}(r_K + s_K)} + 2
\]
where $\epsilon_A^{K} = \frac{\epsilon_A^2}{1 + \epsilon_A^2r_K}$.

**Proof.** Using Lemma 2.6, we have
\[
\|\tilde{A}^*\tilde{A}x - \hat{y}\|_2 \leq \|\tilde{A}^*\tilde{A}x - \hat{y}\|_2 \leq \|\tilde{A}^*\|_2\|\tilde{A}x - \hat{y}\|_2.
\]
Applying $\|\tilde{A}^*\|_2 = \|A + E\|_2$ and $\|\tilde{A}x - \hat{y}\|_2 \leq \epsilon_2^{A,K,y}$, we can easily get
\[
\|\tilde{A}^*\tilde{A}x - \hat{y}\|_2 \leq \|A + E\|_2\epsilon_2^{A,K,y} = \|A\|_2 + \|E\|_2 \leq (\|A\|_2 + \|E\|_2)\epsilon_2^{A,K,y},
\]
Using the inequality $\|E\|_2 \leq \|A\|_2\epsilon_A$, this leads to
\[
\|\tilde{A}^*\tilde{A}x - \hat{y}\|_2 \leq (\|A\|_2 + \|A\|_2\epsilon_A)\epsilon_2^{A,K,y}.
\]
So the lemma is proved.

Next, we will give some important lemmas, and the proof will be provided in Appendix A. All the following lemmas are based on the assumption that $\tilde{A}$ satisfies the BRIP of order $K + 1$.

**Lemma 2.8.** In (1.1), let
\[
P[\tilde{H}_\Omega]\hat{y} = \sum_{i=1}^N \tilde{A}[i]\hat{x}[i] = \tilde{A}[\Omega]\hat{x} \quad \text{with} \quad \hat{x} = \hat{x}[1], \hat{x}[2], \ldots, \hat{x}[s]^T.
\]
Suppose $\tilde{A}x = \tilde{A}[\Omega]\lambda\|x[1]\|_2 \|x[2]\|_2 \cdots \|x[s]\|_2 \neq 0$. For any $s < j \leq N$ and
\[
\Lambda_k = \begin{cases} \emptyset, & k = 0, \\ \{\lambda_1, \lambda_2, \ldots, \lambda_k\}, & 1 \leq k < s, \end{cases}
\]
Lemma 2.9. In\( \|\mathbf{y}\|_2 = 1\), if \(\hat{\mathbf{A}}^{\dagger}(\hat{A}\mathbf{x} - \hat{y})\|_{2,co} \leq \epsilon_{A,K,y}''\), then we have
\[
\|(\hat{P}^{\dagger}[\mathbf{\hat{H}}_{\Omega}](\hat{A}\mathbf{x} - \hat{y}))^{\dagger}\hat{A}[j]\|_2 \leq \|(\hat{A}\mathbf{x} - \hat{y})^{\dagger}(\hat{P}^{\dagger}[\mathbf{\hat{H}}_{\Omega}])\hat{A}[j]\|_2 \leq \left(1 + \frac{\sqrt{sK\delta_{K+1}}}{1 - \delta_{K+1}}\right)\epsilon_{A,K,y}''
\]
for \(s \leq K\) and \(1 \leq j \leq N\).  

Lemma 2.10. In (1.1), let \(P[\mathbf{\hat{H}}_{\Omega}]\hat{y} = \sum_{i=1}^{s} \hat{A}[i]\hat{x}[i] = \hat{A}[\Omega_{s}\hat{x}]\), and suppose that \(\hat{A}\mathbf{x} = \hat{A}[\Omega_{s}]\hat{x}[\Omega_{s}]\) with \(s \leq K\). If \(\|\hat{A}\mathbf{x} - \hat{y}\|_2 \leq \epsilon_{A,K,y}''\), then we have
\[
\min_{1 \leq i \leq s}\|\hat{x}[i]\|_2 \geq \frac{\min_{1 \leq i \leq s}\|\mathbf{x}[i]\|_2}{\sqrt{1 - \delta_{K+1}}} - \frac{\epsilon_{A,K,y}''}{\sqrt{1 - \delta_{K+1}}}
\]
For \(\|\hat{A}^{\dagger}(\hat{A}\mathbf{x} - \hat{y})\|_{2,co} \leq \epsilon_{A,K,y}''\), we also have
\[
\min_{1 \leq i \leq s}\|\hat{x}[i]\|_2 \geq \frac{\sqrt{K\epsilon_{A,K,y}''}}{1 - \delta_{K+1}}.
\]

Lemma 2.11. In (1.1), suppose that
\(\hat{A}\mathbf{x} = \hat{A}[\Omega_{s}]\hat{x}[\Omega_{s}]\) with \(s \leq K\) and \(\|\mathbf{x}[1]\|_2,\|\mathbf{x}[2]\|_2, \ldots, \|\mathbf{x}[s]\|_2 \neq 0\).

If \(\|\hat{A}\mathbf{x} - \hat{y}\|_2 \leq \epsilon_{A,K,y}''\), then we have
\[
\|\hat{P}^{\dagger}[\mathbf{\hat{H}}_{\Omega}]\hat{y}\|_2 \geq \|\hat{P}^{\dagger}[\mathbf{\hat{H}}_{\Omega}]\hat{y}\|_2 \geq \cdots \geq \|\hat{P}^{\dagger}[\mathbf{\hat{H}}_{\Omega}]\hat{y}\|_2 \geq \sqrt{1 - \delta_{K+1}}\min_{1 \leq i \leq s}\|\mathbf{x}[i]\|_2 - \epsilon_{A,K,y}'
\]
For \(\|\hat{A}^{\dagger}(\hat{A}\mathbf{x} - \hat{y})\|_{2,co} \leq \epsilon_{A,K,y}''\), we also have
\[
\|\hat{A}^{\dagger}\hat{P}^{\dagger}[\mathbf{\hat{H}}_{\Omega}]\hat{y}\|_{2,co} \leq \left(1 + \frac{\sqrt{K\delta_{K+1}}}{1 - \delta_{K+1}}\right)\epsilon_{A,K,y}'',
\]
\[
\|\hat{A}^{\dagger}\hat{P}^{\dagger}[\mathbf{\hat{H}}_{\Omega}]\hat{y}\|_{2,co} \geq (1 - \delta_{K+1})\min_{1 \leq i \leq s}\|\mathbf{x}[i]\|_2 - \left(1 + \frac{\sqrt{K(1-\delta_{K+1})}}{1 - \delta_{K+1}}\right)\epsilon_{A,K,y}''
\]
for all \(k < s\)

3 Main theoretical results

In this section, we give the following main results which will be proved in Appendix B. The following results are based on the assumption that \(\text{supp}(\mathbf{x}) = \{i \mid \|\mathbf{x}[i]\|_2 \neq 0\} = \Omega_{s} = \{1, 2, \ldots, s\}\) with \(s \leq K\).

Theorem 3.1. For given relative perturbations \(\epsilon_{A}, \epsilon_{A}^{(K)}, \epsilon_{A}^{(K+1)}\) and \(\epsilon_{y}\) in (2.1), under \(\|\hat{A}\mathbf{x} - \hat{y}\|_2 \leq \epsilon_{A,K,y}''\), assume the BRIC for matrix \(A\) satisfies
\[
\delta_{K+1} < \frac{1}{\sqrt{K + 1}(1 + \epsilon_{A}^{(K+1)})^2} + \frac{1}{(1 + \epsilon_{A}^{(K+1)})^2} - 1.
\]
Then the BOMP algorithm with the stopping criterion $\|r^k\|_2 \leq e'_{A,K,Y}$ exactly recovers $\text{supp}(x)$ in $K$ iterations provided that

$$\min_{i \in \text{supp}(x)} \|x[i]\|_2 > \frac{e'_{A,K,Y}}{\sqrt{1 - ((1 + \delta_{K+1})(1 + e_A^{(K+1)})^2 - 1) + \frac{\sqrt{1 + \delta_{K+1}}(1 + e_A^{(K+1)})^2 e'_{A,K,Y}}{1 - [(1 + \delta_{K+1})(1 + e_A^{(K+1)})^2 - 1] \sqrt{K + 1}}}}.$$ (3.2)

**Remark 3.2.** Theorem 3.1 provides a sufficient condition (3.2) for recovering the support of the block $K$-sparse signals via the BOMP algorithm in the setting of complete perturbation. It has been proven in [30] that if there exists a matrix $A$ satisfying the BRIP with $\frac{1}{\sqrt{K+1}} < \delta_{K+1} < 1$, the BOMP algorithm may fail to recover block $K$-sparse signal $x$ in $K$ iterations. Similar to the proof in [30], we can show that the BOMP algorithm may fail to recover $x$ in $K$ iterations when there exists a matrix $A$ satisfying the BRIP with $\frac{1}{\sqrt{K+1}} < \delta_{K+1} < 1$.

According to the above inequality, we have $\frac{1}{\sqrt{K+1}} \leq \delta_{K+1} < (1 + \delta_{K+1})^2 - 1$. Namely, if there exists a matrix $A$ which satisfies the BRIP with

$$\frac{1}{\sqrt{K+1}} \leq \delta_{K+1} < 1,$$

the BOMP algorithm may fail to recover a block $K$-sparse signal $x$ in $K$ iterations. The results show that (3.1) is a sharp sufficient condition for exact recovery of the block $K$-sparse signal within completely perturbed model with the BOMP algorithm in $K$ iterations. Condition (3.1) of Theorem 3.1 generalizes the condition of [30] to the block-sparse compressed sensing of the completely perturbed model.

In addition, as a special case of Theorem 3.1, $E = 0$, which means that there is no perturbation in the measurement matrix $A$; then $\delta_{K+1} = \delta_{K+1}$. Obviously, when $d_i = 1$ ($i = 1, 2, \ldots, N$), $E = 0$, and we obtain a sufficient condition for recovering $K$-sparse signals by the OMP algorithm in $K$ iterations [19]. Theorem 1 in [19] is included in our theorem.

**Theorem 3.3.** For given relative perturbations $e_A, e_A^{(K)}$ and $e_Y$ in (2.1), under $\|\hat{A}^* (\hat{A}x - \hat{y})\|_{2,\infty} \leq e'_{A,K,Y}$, assume the BRIC for matrix $A$ satisfies

$$\delta_{K+1} < \frac{1}{\sqrt{K + 1}(1 + e_A^{(K+1)})^2} + \frac{1}{(1 + e_A^{(K+1)})^2} - 1.$$ 

Then the BOMP algorithm with the stopping criterion

$$\|\hat{A}^* r^k\|_{2,\infty} \leq \left(1 + \sqrt{K} \delta_{K+1} \right) e''_{A,K,Y}$$

exactly recovers $\text{supp}(x)$ in $K$ iterations provided that

$$\min_{i \in \text{supp}(x)} \|x[i]\|_2 > \frac{\sqrt{K} e_{A,K,Y}}{1 - ((1 + \delta_{K+1})(1 + e_A^{(K+1)})^2 - 1) + \frac{\sqrt{K}(1 + \delta_{K+1})(1 + e_A^{(K+1)} - 1)}{1 - [(1 + \delta_{K+1})(1 + e_A^{(K+1)})^2 - 1] \sqrt{K + 1}}} + \frac{1}{1 - [(1 + \delta_{K+1})(1 + e_A^{(K+1)})^2 - 1] \sqrt{K + 1}}.$$ (3.3)

**Remark 3.4.** From Theorem 3.3, we can see that (3.3) is a sufficient condition for the recovery of the block $K$-sparse signals with the BOMP algorithm. Obviously, when $d_i = 1$ ($i = 1, 2, \ldots, N$), $E = 0$, we obtain a sufficient condition for recovering $K$-sparse signals by the OMP algorithm in $K$ iterations [19]. Theorem 2 in [19] is included in our theorem.

In the following, we will provide the approximation precision between the output signal and the original signal, which is characterized by a total noise and BRIC when $\text{supp}(x) = \{i \mid \|x[i]\|_2 \neq 0\}$ can be exactly and stably recovered in the presence of total noise.

**Theorem 3.5.** In (1.1), let $|\text{supp}(x)| = s \leq K$. If the matrix $A$ satisfies the BRIP of order $K$ and the BOMP algorithm can recover exactly $\text{supp}(x)$ in $s$ iterations, i.e., $A_s = \text{supp}(x)$, then, for $\|\hat{A}x - \hat{y}\|_2 \leq e'_{A,K,Y}$,

$$\|x^* - x[A_s]\|_2 \leq \frac{e_{A,K,Y}}{\sqrt{2 - (1 + \delta_K)(1 + e_A^2)}}.$$ (3.4)
For \( \|\hat{A}^* (\hat{A}x - \hat{y})\|_2 \leq \epsilon_{g,k,y}^{(K)} \),
\[
\|x^* - x[\Lambda]\|_2 \leq \frac{\sqrt{5}}{2 - (1 + \delta_{K+1})(1 + \epsilon_{A}^{(K+1)})^2} \epsilon_{g,k,y}^{(K)}.
\] (3.5)

**Remark 3.6.** Under the \( \epsilon_2 \) and \( \epsilon_3/\epsilon_{g,y} \) bounded total noise, (3.4) and (3.5) in Theorem 3.5 offer upper bound estimations, which explicitly characterize the relationship between the output block-sparse signal and the original block-sparse signal. The results show that the approximation precision is controlled by the total noise and BRIC. Particularly, they reveal that the support of the block-sparse signal can be exactly recovered by the block-sparse compressed sensing of the completely perturbed model with the BOMP algorithm without the noise.

To simplify the notation, for any \( x \in \mathbb{C}^m \), we define
\[
g(x) = \max \left\{ \frac{\|x[\Lambda]\|_2}{\|x[\Lambda]\|_2} \mid \Lambda \subset \Omega_N, \Lambda \neq \emptyset \right\}.
\]

Using Lemma 2.6, if \( x \) is a block \( K \)-sparse signal, we have \( 1 \leq g(x) \leq \sqrt{K} \). In the noiseless case, when the measurement matrix \( A \) is perturbed by \( E \), we prove that the BOMP algorithm can also exactly recover the block \( K \)-sparse signal under some constraints on the block signal and
\[
0 < \delta_{K+1} < \frac{2 + \sqrt{2}}{2(1 + \epsilon_{A}^{(K+1)})^2} - 1.
\]

**Theorem 3.7.** In (1.1), in the noiseless case and when the measurement matrix \( A \) is perturbed by \( E \), if the matrix \( A \) satisfies the BRIP of order \( K + 1 \) with
\[
\delta_{K+1} < \frac{2 + \sqrt{2}}{2(1 + \epsilon_{A}^{(K+1)})^2} - 1
\]
and
\[
g(x) < \frac{\sqrt{2(1 + \delta_{K+1})(1 + \epsilon_{A}^{(K+1)})^2 - (1 + \delta_{K+1})^2(1 + \epsilon_{A}^{(K+1)})^4}}{(1 + \delta_{K+1})(1 + \epsilon_{A}^{(K+1)})^2 - 1},
\] (3.6)
then the BOMP algorithm can exactly recover the block \( K \)-sparse signal in \( K \) iterations.

**Remark 3.8.** Since \( g(x) \geq 1 \) for any block \( K \)-sparse signal \( x \). To ensure that the condition (3.6) is holds, it is necessary to ensure
\[
\frac{\sqrt{2(1 + \delta_{K+1})(1 + \epsilon_{A}^{(K+1)})^2 - (1 + \delta_{K+1})^2(1 + \epsilon_{A}^{(K+1)})^4}}{(1 + \delta_{K+1})(1 + \epsilon_{A}^{(K+1)})^2 - 1} > 1
\]
According to the above inequality, we have
\[
(1 + \delta_{K+1})(1 + \epsilon_{A}^{(K+1)})^2 - 1 < \frac{\sqrt{2}}{2}, \quad \text{i.e.,} \quad \delta_{K+1} < \frac{2 + \sqrt{2}}{2(1 + \epsilon_{A}^{(K+1)})^2} - 1.
\]

## 4 Numerical experiments

In this section, we conduct some numerical experiments to verify our theoretical results by using the block-sparse algorithm with complete perturbation. We also investigate the performance of perturbed OMP and perturbed BOMP algorithm in processing block-sparse signals [3, 10]. Without loss of generality, we consider the block-sparse signal \( x \) with even block size, i.e., \( d_i = d \ (i = 1, \ldots, m) \), and let the signal length be \( N = 256 \). In each experiment, we randomly generate a block-sparse signal \( x \), where the elements of each \( x \) are drawn from the standard Gaussian distribution, and randomly generate a \( 128 \times 256 \) measurement matrix \( A \) from Gaussian distribution [1, 17]. We generate the measurements \( y = Ax + v \) by using \( x \) and \( A \), where \( v \) is the
Gaussian noise. In addition, $E$ is a random Gaussian matrix, where $\|E\|_2 = \epsilon_A \|A\|_2$ (the value of $\epsilon_A$ is not fixed). The perturbed matrix is $\hat{A} = A + E$. In each experiment, we select the average value of 200 independent trails as the final statistical results.

As the first experiment, we compare the performance of perturbed OMP and perturbed BOMP algorithm in processing block-sparse signals. Some reconstruction errors $\|x^* - x\|_2$ are listed in Table 1. Without loss of generality, we choose signals whose block size $d$ ranges from 2 to 32 with the block sparsity $K = 32, 16, 8, 4, 2$ as the test signals and use the reconstruction error to measure the algorithm performance for perturbation level $\epsilon_A = 0, 0.05, 0.1, 0.2$. It is easy to see that the reconstruction error decreases with the decrease of the perturbation level $\epsilon_A$. Table 1 shows that the perturbed BOMP algorithm is more effective than the perturbed OMP algorithm in reconstructing block-sparse signals. In other words, signal structure is very important in signal reconstruction. Specifically, when the perturbation level is $\epsilon_A = 0$ and block sparsity is $K = 2$, the reconstruction performance of the perturbed BOMP algorithm is better than other cases.

In this part, we conduct some numerical experiments to verify our theoretical results by using the completely perturbed BOMP algorithm for Theorem 3.5. With

$$\delta_K = 0.1 \left( \frac{1}{\sqrt{K(1 + \epsilon_A^K)^2}} + \frac{1}{(1 + \epsilon_A^K)^2} - 1 \right),$$

we produce the signals with 64 blocks uniformly at random (i.e., $d = 4$) and choose 10 as the best block number for the sparse approximation. Table 2 and Table 3 show the theoretical error bound $\|x - x^*\|_2$ and $\|x - x^*\|_{2,\infty}$ for various number of samples, where the perturbation level is $\epsilon_A = 0, 0.05, 0.1, 0.15$. The above tables show that the smaller the perturbation and the more samples, the better the performance of the perturbed BOMP algorithm. As the results show, the reconstruction error bound is lower than the theoretical error bound when the number of samples is $M = 112, 128, 144$. However, the reconstruction performance of the perturbed BOMP algorithm is poor when the number of samples is $M = 96$.

<table>
<thead>
<tr>
<th>$\epsilon_A$</th>
<th>$d = 2$</th>
<th>$d = 4$</th>
<th>$d = 8$</th>
<th>$d = 16$</th>
<th>$d = 32$</th>
</tr>
</thead>
<tbody>
<tr>
<td>OMP</td>
<td>0</td>
<td>4.69</td>
<td>4.72</td>
<td>4.53</td>
<td>4.69</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>4.71</td>
<td>5.01</td>
<td>4.82</td>
<td>4.71</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>5.32</td>
<td>5.38</td>
<td>5.39</td>
<td>5.54</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>6.38</td>
<td>6.53</td>
<td>6.23</td>
<td>6.50</td>
</tr>
<tr>
<td>Block-OMP</td>
<td>0</td>
<td>2.18</td>
<td>0.89</td>
<td>0.59</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>2.20</td>
<td>0.91</td>
<td>0.76</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>2.94</td>
<td>1.41</td>
<td>1.09</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>4.51</td>
<td>2.47</td>
<td>1.70</td>
<td>1.68</td>
</tr>
</tbody>
</table>

Table 1: Restorative comparison of reconstruction errors between perturbed OMP and BOMP algorithm.

<table>
<thead>
<tr>
<th>$\epsilon_A$</th>
<th>$M = 96$</th>
<th>$M = 112$</th>
<th>$M = 128$</th>
<th>$M = 144$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$|x - x^*|_2$</td>
<td>0</td>
<td>0.0961</td>
<td>0.4425</td>
<td>0.7674</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.0659</td>
<td>0.4021</td>
<td>0.7725</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.0722</td>
<td>0.4047</td>
<td>0.7941</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>0.0775</td>
<td>0.4095</td>
<td>0.8062</td>
</tr>
</tbody>
</table>

Table 2: Theoretical verification of $\|x - x^*\|_2$ for various number of samples.
\begin{table}[h!]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
$\epsilon_A$ & $M = 96$ & $M = 112$ & $M = 128$ & $M = 144$
\hline
$\|x - x^*\|_{2,\infty}$ & 0.0961 & 0.4425 & 0.7674 & 1.1725 \\
0.05 & 0.0659 & 0.4021 & 0.7725 & 1.1320 \\
0.1 & 0.0722 & 0.4047 & 0.7941 & 1.1739 \\
0.15 & 0.0775 & 0.4095 & 0.8062 & 1.1926 \\
\hline
Theoretical threshold & 0.0516 & 0.4470 & 0.9076 & 1.4541 \\
0.05 & 0.0556 & 0.4569 & 0.9124 & 1.4640 \\
0.1 & 0.0595 & 0.4550 & 0.9040 & 1.4684 \\
0.15 & 0.0631 & 0.4553 & 0.9137 & 1.4728 \\
\hline
\end{tabular}
\caption{Theoretical verification of $\|x - x^*\|_{2,\infty}$ for various number of samples.}
\end{table}

In Figure 1, we compare the theoretical error bound and the reconstruction error of Theorem 3.5 for various block-sparsity levels with perturbation level $\epsilon_A$ for various block-sparsity levels. (a) $\|x - x^*\|_2$ versus block-sparsity level for the block size $d = 4$. (b) The theoretical error bound of (3.4) for $\delta_K = 0.1$. (c) $\|x - x^*\|_{2,\infty}$ versus block-sparsity level for the block size $d = 4$. (d) The theoretical error bound of (3.5) for $\delta_K = 0.1$.

Figure 1: The figure plots theoretical error bound, $\|x - x^*\|_2$ and $\|x - x^*\|_{2,\infty}$ versus the perturbation level $\epsilon_A$ for various block-sparsity levels. (a) $\|x - x^*\|_2$ versus block-sparsity level for the block size $d = 4$. (b) The theoretical error bound of (3.4) for $\delta_K = 0.1$. (c) $\|x - x^*\|_{2,\infty}$ versus block-sparsity level for the block size $d = 4$. (d) The theoretical error bound of (3.5) for $\delta_K = 0.1$. 

In Figure 1, we compare the theoretical error bound and the reconstruction error of Theorem 3.5 for various block-sparsity levels with perturbation level $\epsilon_A = 0, 0.05, 0.1, 0.15$. Figure 1 (a) and (c) show that decreasing $\epsilon_A$ improves the recovery performance of the perturbed BOMP algorithm for a smaller block-sparsity level. Figure 1 (a) and (b) indicate that $\|x - x^*\|_2$ is lower than the theoretical error bound. Similarly, Figure 1 (c) and (d) present the relationship between $\|x - x^*\|_{2,\infty}$ and the theoretical error bound in different block-sparsity levels and $\epsilon_A = 0, 0.05, 0.1, 0.15$, respectively.
5 Conclusion

In this paper, we discuss the sufficient conditions for recovering the support of a block \( K \)-sparse signal under the BRIP using the BOMP algorithm in both the \( \ell_2 \) and \( \ell_2/\ell_{\infty} \) bounded total noise. Under some constraints on the minimum magnitude of the nonzero elements of the block \( K \)-sparse signals, we prove that the support of the block \( K \)-sparse signals can be exactly recovered by the BOMP algorithm in both the \( \ell_2 \) and \( \ell_2/\ell_{\infty} \) bounded total noise if \( A \) satisfies the BRIP of order \( K + 1 \) with

\[
\delta_{K+1} < \frac{1}{\sqrt{|K+1|} (1 + e_k^{(K+1)^2})} + \frac{1}{(1 + e_k^{(K+1)^2})^2} - 1.
\]

In addition, we also give the reconstruction upper bound of the error between the recovered block-sparse signal and the original block-sparse signal. In the noiseless case, when the measurement matrix \( A \) is perturbed by \( E \), we also prove that the BOMP algorithm can exactly recover the block \( K \)-sparse signal under some constraints on the block \( K \)-sparse signal and

\[
\delta_{K+1} < \frac{2 + \sqrt{2}}{2(1 + e_k^{(K+1)^2})^2} - 1.
\]

Finally, a series of numerical experiments are carried out to verify our theoretical results. The results obtained are helpful in further discussing and studying the perturbed BOMP algorithm.

A Appendix

Proof of Lemma 2.8. Let \( \Lambda_k = \Omega_k \). Then we have \( \hat{H}_{\Omega_k} = \hat{H}_{\Omega_k} \). If \( \langle P^k[\hat{H}_{\Omega_k}]y, \hat{A}[j] \rangle = e^{i\theta} \| \hat{A}[j]^* P^k[\hat{H}_{\Omega_k}]y \|_2 \), then we have \( S_{\Omega_k} = \| \hat{A}[j]^* P^k[\hat{H}_{\Omega_k}]y \|_2 \), where \( \hat{A}[j] = e^{i\theta} \hat{A}[j] \). For any \( J_a = \{j_1, j_2, \ldots, j_a\} \subset \Omega_s \), let

\[
\hat{A}_{J_a} = [\hat{A}[j_1], \hat{A}[j_2], \ldots, \hat{A}[j_a]].
\]

Let \( \hat{A}[J_a, j] = \hat{A}[J_a \cup \{j\}]U \) with \( U = \text{diag}(I_{j_1}, I_{j_2}, \ldots, I_{j_a}, e^{i\theta} I_{j}) \). Thus, we have

\[
\hat{A}^*[J_a, j] \hat{A}[J_a, j] = U^* \hat{A}^*[J_a \cup \{j\}] \hat{A}[J_a \cup \{j\}] U.
\]

It is clear that \( \| J_a \union \{j\} \|_\infty \leq K + 1 \) and \( U^* U = I_{m+1} \). Then, by Lemma 2.4, we obtain

\[
(1 - \delta_{K+1}) I_{[a + 1]} \leq \hat{A}^*[J_a, j] \hat{A}[J_a, j] \leq (1 + \delta_{K+1}) I_{[a + 1]}.
\]

(A.1)

From Lemma 2.2, we can easily get

\[
P^k[\hat{H}_{\Omega_k}]P^k[\hat{H}_{\Omega_k}] = P^k[\hat{H}_{\Omega_k}]P^k[\hat{H}_{\Omega_k}],
\]

\[
P^k[\hat{H}_{\Omega_k}] \hat{A}[i] = 0, \quad \text{with } 1 \leq i \leq k,
\]

and

\[
P^k[\hat{H}_{\Omega_k}] \tilde{y} = P^k[\hat{H}_{\Omega_k}] [\hat{A} \tilde{x} + (\tilde{y} - \hat{A} \tilde{x})] = P^k[\hat{H}_{\Omega_k}] [\hat{A} \Omega_s \tilde{x} \Omega_s + (\tilde{y} - \hat{A} \tilde{x})] = P^k[\hat{H}_{\Omega_k}] [\tilde{y} - \hat{A} \tilde{x}].
\]

For any \( t > 0 \), we have

\[
\left\| \left( t + \frac{1}{\| \hat{y}[\Omega_s \setminus \Omega_k] \|_1} \right) P^k[\hat{H}_{\Omega_k}] \hat{y} - P^k[\hat{H}_{\Omega_k}] \hat{A}[j] \right\|_2^2 - \left( t - \frac{1}{\| \hat{y}[\Omega_s \setminus \Omega_k] \|_1} \right) P^k[\hat{H}_{\Omega_k}] \hat{y} + P^k[\hat{H}_{\Omega_k}] \hat{A}[j] \right\|_2^2
\]

\[
= \frac{4t}{\| \hat{y}[\Omega_s \setminus \Omega_k] \|_1} \| P^k[\hat{H}_{\Omega_k}] \hat{y} \|_2^2 - 4t \text{ Re} [\hat{y}[j], P^k[\hat{H}_{\Omega_k}] \hat{y}] + \frac{4t}{\| \hat{y}[\Omega_s \setminus \Omega_k] \|_1} \langle P^k[\hat{H}_{\Omega_k}] \hat{y}, P^k[\hat{H}_{\Omega_k}] \hat{y} \rangle - 4t S_{\Omega_k}
\]
by (A.2), (A.3), (A.4), this leads to

\[
\frac{4t}{2} \left\langle P^+[\hat{H}_{\Omega_1}]\hat{y}, P^+[\hat{H}_{\Omega_1}](P[\hat{H}_{\Omega_1}]\hat{y}) \right\rangle \\
+ \frac{4t}{2} \left\langle P^+[\hat{H}_{\Omega_1}]\hat{y}, P^+[\hat{H}_{\Omega_1}] \left( \sum_{i=1}^{s} \hat{x}[i]\tilde{A}[i] \right) \right\rangle \\
+ \frac{4t}{2} \left\langle P^+[\hat{H}_{\Omega_1}]\hat{y}, P^+[\hat{H}_{\Omega_1}](\hat{y} - \hat{A}\hat{x}) \right\rangle - 4tS_{\Omega_1} \\
= \frac{4t}{2} \left\langle P^+[\hat{H}_{\Omega_1}]\hat{y}, \sum_{i=1}^{s} \hat{x}[i]\tilde{A}[i] \right\rangle \\
+ \frac{4t}{2} \left\langle \hat{y}, P^+[\hat{H}_{\Omega_1}] \left( \sum_{i=1}^{s} \hat{x}[i]\tilde{A}[i] \right) \right\rangle - 4tS_{\Omega_1} \\
= \frac{4t}{2} \left\langle P^+[\hat{H}_{\Omega_1}]\hat{y}, \sum_{i=1}^{s} \hat{x}[i]\tilde{A}[i] \right\rangle + \frac{4t}{2} \left\langle \hat{y}, P^+[\hat{H}_{\Omega_1}]\hat{y} - \hat{A}\hat{x} \right\rangle - 4tS_{\Omega_1}. \quad (A.2)
\]

Since

\[
\left\| \left( \sum_{i=1}^{s} \hat{x}[i]\tilde{A}[i] \right) \right\|_2 \leq \left\| \hat{x}[\Omega_1] \right\|_2 S_{\Omega_1} \quad (A.3)
\]

and

\[
\left\langle \hat{y}, P^+[\hat{H}_{\Omega_1}]\hat{y} - \hat{A}\hat{x} \right\rangle = \left\langle P^+[\hat{H}_{\Omega_1}]\hat{y}, P^+[\hat{H}_{\Omega_1}]\hat{y} - \hat{A}\hat{x} \right\rangle = \left\| P^+[\hat{H}_{\Omega_1}]\hat{y} - \hat{A}\hat{x} \right\|_2^2, \quad (A.4)
\]

by (A.2), (A.3), (A.4), this leads to

\[
\begin{align*}
\left\| \left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) P^+[\hat{H}_{\Omega_1}]\hat{y} - P^+[\hat{H}_{\Omega_1}]\hat{y} \right\|_2^2 \\
- \left\| \left( t - \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) P^+[\hat{H}_{\Omega_1}]\hat{y} + P^+[\hat{H}_{\Omega_1}]\hat{y} \right\|_2^2 \\
&\leq 4t(S_{\Omega_1} - S_{\Omega_1}) + \frac{4t}{\left\| \hat{x}[\Omega_1] \right\|_1} \left\| P^+[\hat{H}_{\Omega_1}]\hat{y} - \hat{A}\hat{x} \right\|_2^2. \quad (A.5)
\end{align*}
\]

Let
\[
\xi = \left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) \hat{x}[k + 1], \ldots, \left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) \hat{x}[s], -1^T, \\
\eta = \left( t - \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) \hat{x}[k + 1], \ldots, \left( t - \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) \hat{x}[s], 1^T.
\]

According to the above equality, we have

\[
\begin{align*}
\left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) P^+[\hat{H}_{\Omega_1}]\hat{y} - P^+[\hat{H}_{\Omega_1}]\hat{y} \\
= \left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) P^+[\hat{H}_{\Omega_1}] \left( P[\hat{H}_{\Omega_1}]\hat{y} - P^+[\hat{H}_{\Omega_1}]\hat{A}[j] \right) + \left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) P^+[\hat{H}_{\Omega_1}] \left( P^+[\hat{H}_{\Omega_1}]\hat{y} \right) \\
= P^+[\hat{H}_{\Omega_1}] \left( \sum_{i=1}^{s} \left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) \hat{x}[i]\tilde{A}[i] - \hat{A}[j] \right) + \left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) P^+[\hat{H}_{\Omega_1}] \left( \hat{y} - \hat{A}\hat{x} \right) \\
= P^+[\hat{H}_{\Omega_1}] \hat{A}[\Omega_1 \setminus \Omega_k, j] \xi + \left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) P^+[\hat{H}_{\Omega_1}] \left( \hat{y} - \hat{A}\hat{x} \right).
\end{align*}
\]

It is easy to see that

\[
\begin{align*}
\left\langle P^+[\hat{H}_{\Omega_1}]\hat{A}[\Omega_1 \setminus \Omega_k, j] \xi, \left( t + \frac{1}{\left\| \hat{x}[\Omega_1] \right\|_1} \right) P^+[\hat{H}_{\Omega_1}] \left( \hat{y} - \hat{A}\hat{x} \right) \right\rangle \\
= \left( t + \frac{1}{\left\| \hat{x}[\Omega_1 \setminus \Omega_k] \right\|_1} \right) \left\langle P^+[\hat{H}_{\Omega_1}]\hat{A}[\Omega_1 \setminus \Omega_k, j], \left( \hat{y} - \hat{A}\hat{x} \right) \right\rangle \\
= -\left( t + \frac{1}{\left\| \hat{x}[\Omega_1 \setminus \Omega_k] \right\|_1} \right) \left\langle P^+[\hat{H}_{\Omega_1}]\hat{A}[j], \left( \hat{y} - \hat{A}\hat{x} \right) \right\rangle.
\end{align*}
\]
From the above equality, we thus have

\[
\left\| \left( t + \frac{1}{\| \hat{x} \|_2} \right) P^+ [\hat{H}_{\Omega_1}] \hat{y} - P^+ [\hat{H}_{\Omega_1}] \hat{A} [j] \right\|_2^2 + \left( 1 + \frac{1}{\| \hat{x} \|_2} \right) \left\| P^+ [\hat{H}_{\Omega_1}] (\hat{y} - \hat{A} \hat{x}) \right\|_2^2
\]

\[
= \left\| P^+ [\hat{H}_{\Omega_1}] \hat{A} [\Omega_\setminus \Omega_k, j] \right\|_2^2 + \left( 1 + \frac{1}{\| \hat{x} \|_2} \right) \left\| P^+ [\hat{H}_{\Omega_1}] (\hat{y} - \hat{A} \hat{x}) \right\|_2^2
\]

\[
- 2 \left( t + \frac{1}{\| \hat{x} \|_2} \right) \text{Re} \left\{ \hat{A} [j], P^+ [\hat{H}_{\Omega_1}] (\hat{y} - \hat{A} \hat{x}) \right\}
\]

\[
= \xi^* \hat{A}^* [\Omega_\setminus \Omega_k, j] P^+ [\hat{H}_{\Omega_1}] \hat{A} [\Omega_\setminus \Omega_k, j] \xi + \left( 1 + \frac{1}{\| \hat{x} \|_2} \right) \left\| P^+ [\hat{H}_{\Omega_1}] (\hat{y} - \hat{A} \hat{x}) \right\|_2^2
\]

\[
- 2 \left( t + \frac{1}{\| \hat{x} \|_2} \right) \text{Re} \left\{ \hat{A} [j], P^+ [\hat{H}_{\Omega_1}] (\hat{y} - \hat{A} \hat{x}) \right\}.
\]

(A.6)

Similar to the proof of (A.6), we can easily get

\[
\left\| \left( t - \frac{1}{\| \hat{x} \|_2} \right) P^+ [\hat{H}_{\Omega_1}] \hat{y} + P^+ [\hat{H}_{\Omega_1}] \hat{A} [j] \right\|_2^2
\]

\[
= \eta^* \hat{A}^* [\Omega_\setminus \Omega_k, j] P^+ [\hat{H}_{\Omega_1}] \hat{A} [\Omega_\setminus \Omega_k, j] \eta + \left( t - \frac{1}{\| \hat{x} \|_2} \right) \left\| P^+ [\hat{H}_{\Omega_1}] (\hat{y} - \hat{A} \hat{x}) \right\|_2^2
\]

\[
+ 2 \left( t - \frac{1}{\| \hat{x} \|_2} \right) \text{Re} \left\{ \hat{A} [j], P^+ [\hat{H}_{\Omega_1}] (\hat{y} - \hat{A} \hat{x}) \right\}.
\]

(A.7)

By using (A.6) and (A.7), this leads to

\[
\left\| \left( t + \frac{1}{\| \hat{x} \|_2} \right) P^+ [\hat{H}_{\Omega_1}] \hat{y} - P^+ [\hat{H}_{\Omega_1}] \hat{A} [j] \right\|_2^2 - \left\| \left( t - \frac{1}{\| \hat{x} \|_2} \right) P^+ [\hat{H}_{\Omega_1}] \hat{y} + P^+ [\hat{H}_{\Omega_1}] \hat{A} [j] \right\|_2^2
\]

\[
= \xi^* \hat{A}^* [\Omega_\setminus \Omega_k, j] P^+ [\hat{H}_{\Omega_1}] \hat{A} [\Omega_\setminus \Omega_k, j] \xi - \eta^* \hat{A}^* [\Omega_\setminus \Omega_k, j] P^+ [\hat{H}_{\Omega_1}] \hat{A} [\Omega_\setminus \Omega_k, j] \eta
\]

\[
+ 4t \left\| P^+ [\hat{H}_{\Omega_1}] (\hat{y} - \hat{A} \hat{x}) \right\|_2^2 - 4t \text{Re} \left\{ \hat{A} [j], P^+ [\hat{H}_{\Omega_1}] (\hat{y} - \hat{A} \hat{x}) \right\}.
\]

(A.8)

Matrix \( \tilde{A} [\Omega_\setminus j] = [\hat{A} [1], \hat{A} [2], \ldots, \hat{A} [s], \hat{A} [j]] \) is divided into matrix \( \tilde{A} [\Omega_\setminus j] = [\hat{A} [\Omega_\setminus k], \hat{A} [\Omega_\setminus \Omega_k, j]] \). According to some simple calculations, we have

\[
\tilde{A}^* [\Omega_\setminus j] \tilde{A} [\Omega_\setminus j] = \begin{bmatrix}
\hat{A}^* [\Omega_\setminus k] \hat{A} [\Omega_\setminus k] & \hat{A}^* [\Omega_\setminus k] \hat{A} [\Omega_\setminus \Omega_k, j] \\
\hat{A}^* [\Omega_\setminus \Omega_k, j] \hat{A} [\Omega_\setminus k] & \hat{A}^* [\Omega_\setminus \Omega_k, j] \hat{A} [\Omega_\setminus \Omega_k, j]
\end{bmatrix}.
\]

Then the Schur complement of \( \tilde{A}^* [\Omega_\setminus j] \tilde{A} [\Omega_\setminus j] \) in \( \hat{A}^* [\Omega_\setminus k] \hat{A} [\Omega_\setminus k] \) is defined as

\[
C_k (\tilde{A}^* [\Omega_\setminus j] \tilde{A} [\Omega_\setminus j]) = \hat{A}^* [\Omega_\setminus k] \hat{A} [\Omega_\setminus k] - \hat{A}^* [\Omega_\setminus \Omega_k, j] \hat{A} [\Omega_\setminus k] \left( \hat{A}^* [\Omega_\setminus k] \hat{A} [\Omega_\setminus k] \right)^{-1} \hat{A}^* [\Omega_\setminus k] \hat{A} [\Omega_\setminus k, j]
\]

\[
= \hat{A}^* [\Omega_\setminus k] \hat{A} [\Omega_\setminus k, j] P^+ [\hat{H}_{\Omega_1}] \hat{A} [\Omega_\setminus k, j].
\]

(A.9)

From (2.5) and (A.1), this leads to

\[
(1 - \hat{\delta}_{k+1}) I_{s-k-1} \leq C_k (\tilde{A}^* [\Omega_\setminus j] \tilde{A} [\Omega_\setminus j]) \leq (1 + \hat{\delta}_{k+1}) I_{s-k+1}
\]

(A.10)

By combining (2.4), (A.9) and (A.10), we obtain

\[
\xi^* \hat{A}^* [\Omega_\setminus \Omega_k, j] P^+ [\hat{H}_{\Omega_1}] \hat{A} [\Omega_\setminus \Omega_k, j] \xi
\]

\[
= \xi^* \hat{A}^* [\Omega_\setminus \Omega_k, j] \left(I_n - \hat{A} [\Omega_\setminus k] \left( \hat{A}^* [\Omega_\setminus k] \hat{A} [\Omega_\setminus k] \right)^{-1} \hat{A}^* [\Omega_\setminus k] \right) \hat{A} [\Omega_\setminus \Omega_k, j] \xi
\]

\[
= \xi^* \xi C_k (\tilde{A}^* [\Omega_\setminus j] \tilde{A} [\Omega_\setminus j]) \xi
\]

\[
\geq (1 - \hat{\delta}_{k+1}) \xi^* \xi = (1 - \hat{\delta}_{k+1}) \left( t + \frac{1}{\| \hat{x} \|_2} \right)^2 \| \hat{x} [\Omega_\setminus k] \|_2^2 + 1.
\]
Similarly, we obtain
\[
\eta^* \tilde{A}^* [\Omega_\delta \setminus \Omega_k, j] P^\dagger [\tilde{H}_\Omega, j] \tilde{A} [\Omega_\delta \setminus \Omega_k, j] \eta = \eta^* \mathcal{C}_\delta (\tilde{A}^* [\Omega_\delta, j] \tilde{A} [\Omega_\delta, j]) \eta \\
\leq (1 + \tilde{\delta}_{K+1}) \eta^* \eta \\
= (1 + \tilde{\delta}_{K+1}) \left( t - \frac{1}{\| \hat{x} [\Omega_\delta \setminus \Omega_k] \|_2^2 + 1} \right) \| \hat{x} [\Omega_\delta \setminus \Omega_k] \|_2^2 + 1 \right). \tag{A.12}
\]

By using (A.5), (A.8), (A.11) and (A.12), we can easily get
\[
S_{0\Omega_k} - S_{j\Omega_k} \geq \| \hat{x} [\Omega_\delta \setminus \Omega_k] \|_2^2 \left( \frac{\tilde{\delta}_{K+1}}{2} \left( t - \frac{1}{\| \hat{x} [\Omega_\delta \setminus \Omega_k] \|_2^2 + 1} \right) \right) \\
- \text{Re} (\hat{A} [j], P^\dagger [\tilde{H}_\Omega, j] (\hat{A} x - \hat{y})), \tag{A.13}
\]

According to the arithmetic-geometric inequality and (A.13), we have
\[
S_{0\Omega_k} - S_{j\Omega_k} \geq \max_{t \geq 0} \left\{ \| \hat{x} [\Omega_\delta \setminus \Omega_k] \|_2^2 \left( \frac{\tilde{\delta}_{K+1}}{2} \left( t - \frac{1}{\| \hat{x} [\Omega_\delta \setminus \Omega_k] \|_2^2 + 1} \right) \right) \\
- \text{Re} (\hat{A} [j], P^\dagger [\tilde{H}_\Omega, j] (\hat{A} x - \hat{y})), \right\}
\]
\[
= \| \hat{x} [\Omega_\delta \setminus \Omega_k] \|_2^2 \left( \frac{\tilde{\delta}_{K+1}}{2} \right) \| \hat{x} [\Omega_\delta \setminus \Omega_k] \|_2^2 + 1 \right) \right) \\
- \text{Re} (\hat{A} [j], P^\dagger [\tilde{H}_\Omega, j] (\hat{A} x - \hat{y})).
\]

So the lemma is proved.

**Proof of Lemma 2.9.** When \( j \leq s \), we have
\[
\| (P^\dagger [\tilde{H}_\Omega, j] (\hat{A} x - \hat{y}))^* \hat{A} [j] \|_2 = \| (\hat{A} x - \hat{y})^* (P^\dagger [\tilde{H}_\Omega, j] \hat{A} [j]) \|_2 = 0,
\]
so the lemma holds evidently.

When \( j > s \), let \( \mathcal{P}[\tilde{H}_\Omega, j] = \sum_{i=1}^{s} \hat{A} [i] z [i] = \hat{A} [\Omega_\delta] Z \) with \( Z = (z[1], z[2], \ldots, z[s])^T \). From Lemma 2.4, this leads to
\[
\| \mathcal{P}[\tilde{H}_\Omega, j] \hat{A} [j] \|_2^2 = \| \hat{A} [\Omega_\delta] Z \|_2^2 = \| Z^* \hat{A}^* [\Omega_\delta] \hat{A} [\Omega_\delta] Z \|_2 \geq (1 - \tilde{\delta}_{K+1}) \| Z \|_2^2. \tag{A.14}
\]

For any \( t \geq 0 \), we then have
\[
\| \mathcal{P}[\tilde{H}_\Omega, j] \hat{A} [j] \|_2^2 = \langle \mathcal{P}[\tilde{H}_\Omega, j] \hat{A} [j], \mathcal{P}[\tilde{H}_\Omega, j] \hat{A} [j] \rangle = \text{Re} \left( \mathcal{P}[\tilde{H}_\Omega, j] \hat{A} [j], \sum_{i=1}^{s} \hat{A} [i] z [i] \right) \\
= \text{Re} \left( \hat{A} [j], \mathcal{P}[\tilde{H}_\Omega, j] \left( \sum_{i=1}^{s} \hat{A} [i] z [i] \right) \right) = \text{Re} \left( \hat{A} [j], \sum_{i=1}^{s} \hat{A} [i] z [i] \right) \\
= \frac{1}{4t} \left( \left\| \sum_{i=1}^{s} \hat{A} [i] z [i] + t \hat{A} [j] \right\|_2^2 - \left\| \sum_{i=1}^{s} \hat{A} [i] z [i] - t \hat{A} [j] \right\|_2^2 \right) \\
\leq \frac{1}{4t} \left( (1 + \tilde{\delta}_{K+1}) \| Z \|_2^2 + t \right) - (1 - \tilde{\delta}_{K+1}) \| Z \|_2^2 + t^2) \\
= \frac{\tilde{\delta}_{K+1}}{2t} \left( \| Z \|_2^2 + t^2 \right) - \frac{\tilde{\delta}_{K+1}}{2} \left( \frac{1}{t} \| Z \|_2^2 + t \right). \tag{A.15}
\]

By using the arithmetic-geometric inequality, (A.14) and (A.15), then
\[
(1 - \tilde{\delta}_{K+1}) \| Z \|_2^2 \leq \min_{t \geq 0} \left\{ \frac{\tilde{\delta}_{K+1}}{2} \left( \frac{1}{t} \| Z \|_2^2 + t \right) \right\} = \hat{\delta}_{K+1} \| Z \|_2.
\]
This implies
\[
\| Z \|_2 \leq \frac{\hat{\delta}_{K+1}}{1 - \tilde{\delta}_{K+1}}.
\]
From Lemma 2.6 and the inequality
\[ \| (\hat{A} - \hat{y})^* \hat{A} \|_2 = \| (\hat{A})^* (\hat{A} - \hat{y}) \|_2 \leq \| \hat{A}^*(\hat{A} - \hat{y}) \|_{2,\infty} = \max_{1 \leq i \leq N} \| (\hat{A})^* (\hat{A} - \hat{y}) \|_2 \leq \epsilon_{A,K,y}'', \]
we have
\[
\| (\hat{A} - \hat{y})^* (P^1 [\hat{H}_\Omega] \hat{A}) \|_2 = \| (\hat{A} - \hat{y})^* (\hat{A} - P^1 [\hat{H}_\Omega] \hat{A}) \|_2 \\
\leq \| (\hat{A} - \hat{y})^* \hat{A} \|_2 + \| (\hat{A} - \hat{y})^* P^1 [\hat{H}_\Omega] \hat{A} \|_2 \\
\leq \epsilon_{A,K,y}'' + \| (\hat{A} - \hat{y})^* \sum_{i=1}^s \hat{A}(i) z[i] \|_2 \\
\leq \epsilon_{A,K,y}'' + \| z \|_1 \epsilon_{A,K,y}'' \leq \epsilon_{A,K,y}'' + \sqrt{s} \| z \|_2 \epsilon_{A,K,y}'' \\
\leq \left( 1 + \frac{\sqrt{s} \delta_{K+1}}{1 - \delta_{K+1}} \right) \epsilon_{A,K,y}''.
\]
So the lemma is proved. \(\square\)

**Proof of Lemma 2.10.** Using the equality
\[ P^1 [\hat{H}_\Omega] \hat{y} = \sum_{i=1}^s \hat{A}(i) \hat{x}[i] = \hat{A}[\Omega_3] \hat{x}, \]
we have
\[ P^1 [\hat{H}_\Omega] (\hat{A} - \hat{y}) = \hat{A}[\Omega_3] (\hat{x} - \hat{\hat{x}}). \] (A.16)
Let \( V = \hat{x} - \hat{\hat{x}}, \) i.e., \( V[i] = \hat{x}[i] - \hat{\hat{x}}[i] \) for any \( 1 \leq i \leq s. \) When \( V = 0, (2.4), (2.5) \) hold evidently.

When \( V \neq 0, \) for \( \| \hat{A} - \hat{y} \|_2 \leq \epsilon_{A,K,y}', \) by Lemma 2.2, Lemma 2.4 and (A.16), we obtain
\[ (\epsilon_{A,K,y}')^2 \geq \| P^1 [\hat{H}_\Omega] (\hat{A} - \hat{y}) \|_2^2 = \| \hat{A}[\Omega_3] V \|_2^2 = V^* \hat{A}^* [\Omega_3] \hat{A}[\Omega_3] V \geq (1 - \delta_{K+1}) \| V \|_2^2. \] (A.17)
From Lemma 2.6 and (A.17), this leads to
\[ \| V \|_{2,\infty} \leq \| V \|_2 \leq \frac{\epsilon_{A,K,y}'}{\sqrt{1 - \delta_{K+1}}}. \] (A.18)
Thus, it follows from (A.18) that
\[ \min_{1 \leq i \leq s} \| \hat{x}[i] \|_2 = \min_{1 \leq i \leq s} \| \hat{x}[i] + (\hat{x}[i] - \hat{\hat{x}}[i]) \|_2 \geq \min_{1 \leq i \leq s} \| \hat{x}[i] \|_2 - \max_{1 \leq i \leq s} \| \hat{x}[i] - \hat{\hat{x}}[i] \|_2 = \min_{1 \leq i \leq s} \| \hat{x}[i] \|_2 - \| V \|_{2,\infty} \]
\[ \geq \min_{1 \leq i \leq s} \| \hat{x}[i] \|_2 - \frac{\epsilon_{A,K,y}'}{\sqrt{1 - \delta_{K+1}}}. \]
So (2.4) holds.

For \( \| \hat{A}^*(\hat{A} - \hat{y}) \|_{2,\infty} \leq \epsilon_{A,K,y}'', \) we have
\[
\| P^1 [\hat{H}_\Omega] (\hat{A} - \hat{y}) \|_2^2 = \| (P^1 [\hat{H}_\Omega] (\hat{A} - \hat{y})) \|_2^2 \\
= \left\| \left< P^1 [\hat{H}_\Omega] (\hat{A} - \hat{y}), \sum_{i=1}^s \hat{A}(i) (\hat{x}[i] - \hat{\hat{x}}[i]) \right> \right\|_2 \\
= \left\| \left< \hat{A}^*(\hat{A} - \hat{y}), P^1 [\hat{H}_\Omega] \left( \sum_{i=1}^s \hat{A}(i) (\hat{x}[i] - \hat{\hat{x}}[i]) \right) \right> \right\|_2 \\
\leq \| (\hat{A}[\Omega_3])^* (\hat{A} - \hat{y}) \|_{2,\infty} \| V \|_2 \leq \| V \|_{2,1} \leq \| V \|_{2,1} \epsilon_{A,K,y}''.
\] (A.19)
From Lemma 2.6, (A.17) and (A.19), we can easily get
\[ \| V \|_{2,\infty} \leq \frac{\epsilon_{A,K,y}'}{\sqrt{1 - \delta_{K+1}}} \| V \|_2 \leq \frac{\sqrt{s} \epsilon_{A,K,y}'}{\sqrt{1 - \delta_{K+1}}} \leq \frac{\sqrt{s} \epsilon_{A,K,y}'}{\sqrt{1 - \delta_{K+1}}}. \] (A.20)
Then it follows from (A.20) that
\[
\min_{1 \leq i \leq s} ||\hat{x}[i]||_2 = \min_{1 \leq i \leq s} ||x[i] + (\hat{x}[i] - x[i])||_2 \geq \min_{1 \leq i \leq s} ||x[i]||_2 - \max_{1 \leq i \leq s} ||\hat{x}[i] - x[i]||_2 = \min_{1 \leq i \leq s} ||x[i]||_2 - ||V||_{2,\infty}
\]
\[
\geq \min_{1 \leq i \leq s} ||x[i]||_2 - \frac{\sqrt{\mathcal{K}e_{A,K,y}}}{1 - \delta_{K+1}}.
\]
Therefore, (2.5) holds. So the lemma is proved. \(\square\)

**Proof of Lemma 2.11.** For $\|\hat{A}x - \hat{y}\|_2 \leq c'_{A,K,y}$ and $\hat{A}x = \hat{A}[\Omega_s]x[\Omega_s]$, we have
\[
\|P^+[\hat{H}_{\Omega_s}](\hat{A}x - \hat{y})\|_2 \leq \|\hat{A}x - \hat{y}\|_2 \leq c'_{A,K,y}
\]
with $0 \leq j \leq N$.

Since $\hat{H}_{\Omega_0} \subset \hat{H}_{\Omega_1} \subset \cdots \subset \hat{H}_{\Omega_{s-1}}$, we can easily get $\hat{H}_{\Omega_0} \supset \hat{H}_{\Omega_1} \supset \cdots \supset \hat{H}_{\Omega_{s-1}}$. Then we have
\[
\|P^+[\hat{H}_{\Omega_s}]\hat{y}\|_2 \geq \|P^+[\hat{H}_{\Omega_{s-1}}]\hat{y}\|_2 \geq \cdots \geq \|P^+[\hat{H}_{\Omega_1}]\hat{y}\|_2.
\]
And also noting that
\[
\|P^+[\hat{H}_{\Omega_{s-1}}]\hat{y}\|_2 = \|P^+[\hat{H}_{\Omega_{s-1}}](\hat{A}[\Omega_s]x[\Omega_s] - \hat{A}x + \hat{y})\|_2
\]
\[
\geq \|P^+[\hat{H}_{\Omega_{s-1}}]A[\Omega_s]x[\Omega_s]\|_2 - \|P^+[\hat{H}_{\Omega_{s-1}}](\hat{A}x - \hat{y})\|_2
\]
\[
\geq \|P^+[\hat{H}_{\Omega_{s-1}}]A[\Omega_s]x[\Omega_s]\|_2 - c'_{A,K,y},
\]
we can easily get
\[
P^+[\hat{H}_{\Omega_{s-1}}][\hat{A}[\Omega_s]x[\Omega_s] = P^+[\hat{H}_{\Omega_{s-1}}]A[\Omega_s \setminus \Omega_{s-1}]x[\Omega_s \setminus \Omega_{s-1}] = x[s]P^+[\hat{H}_{\Omega_{s-1}}]\hat{A}[s]
\]
and
\[
\hat{A}^*[\Omega_s][\hat{A}[\Omega_s]] = \hat{A}^*[s]\hat{A}[s] - \hat{A}^*[s]\hat{A}[\Omega_{s-1}][\hat{A}^*[\Omega_{s-1}][\hat{A}[\Omega_{s-1}]]^{-1}\hat{A}^*[\Omega_{s-1}][\hat{A}[s].
\]
From Lemma 2.4, Lemma 2.5, (A.22) and (A.23), we have
\[
\|P^+[\hat{H}_{\Omega_{s-1}}]A[\Omega_s]x[\Omega_s]\|_2^2 = \|x[s]\|^2 \hat{A}^*[\Omega_s][\hat{A}[\Omega_s]] \hat{A}[s] = \|x[s]\|^2 \|C_{s-1}(\hat{A}^*[\Omega_s][\hat{A}[\Omega_s])\|^2 
\]
\[
\geq (1 - \delta_{K+1}) \min_{1 \leq i \leq s} ||x[i]||_2^2.
\]
By combining (A.21) and (A.24), therefore (2.6) holds.

For $\|\hat{A}^*[\hat{A}x - \hat{y}]\|_{2,\infty} \leq c''_{A,K,y}$ and $\hat{A}x = \hat{A}[\Omega_s]x[\Omega_s]$, then it follows from Lemma 2.9 that
\[
\|\hat{A}^*P^+[\hat{H}_{\Omega_s}]\|_{2,\infty} = \|\hat{A}^*P^+[\hat{H}_{\Omega_s}][\hat{A}[\Omega_s]x[\Omega_s] - \hat{A}x + \hat{y}]\|_{2,\infty}
\]
\[
\leq \|\hat{A}^*P^+[\hat{H}_{\Omega_s}]\|_{2,\infty} + \|\hat{A}^*P^+[\hat{H}_{\Omega_s}][\hat{A}x - \hat{y}]\|_{2,\infty}
\]
\[
= \|\hat{A}^*P^+[\hat{H}_{\Omega_s}][\hat{A}x - \hat{y}]\|_{2,\infty} = \max_{1 \leq k \leq N} \|P^+[\hat{H}_{\Omega_s}][\hat{A}x - \hat{y}]\|_2
\]
\[
\leq \left(1 + \frac{\sqrt{\mathcal{K}e_{A,K,y}}}{1 - \delta_{K+1}}\right)c''_{A,K,y},
\]
so (2.7) holds.

Note that, for $k < s$, we have
\[
\|\hat{A}^*P^+[\hat{H}_{\Omega_s}][\hat{A}[\Omega_s]x[\Omega_s] - \hat{A}x + \hat{y}]\|_{2,\infty}
\]
\[
\geq \|\hat{A}^*P^+[\hat{H}_{\Omega_s}][\hat{A}[\Omega_s]x[\Omega_s] - \hat{A}x + \hat{y}]\|_{2,\infty}
\]
\[
\geq \|\hat{A}^*P^+[\hat{H}_{\Omega_s}][\hat{A}[\Omega_s]x[\Omega_s]] - \|\hat{A}^*P^+[\hat{H}_{\Omega_s}][\hat{A}x - \hat{y}]\|_{2,\infty}.
\]
Using Lemma 2.4 and Lemma 2.5, we have
\[
\|P^+[\hat{H}_{\Omega_s}]A[\Omega_s]x[\Omega_s]\|_2^2
\]
\[
= \|P^+[\hat{H}_{\Omega_s}]A[\Omega_s \setminus \Omega_k]x[\Omega_s \setminus \Omega_k]\|_2^2
\]
\[
= \hat{x}^*[\Omega_s \setminus \Omega_k][\hat{A}^*[\Omega_s \setminus \Omega_k]P^+[\hat{H}_{\Omega_s}]A[\Omega_s \setminus \Omega_k]x[\Omega_s \setminus \Omega_k]
\]
\[
= \hat{x}^*[\Omega_s \setminus \Omega_k][\hat{A}^*[\Omega_s \setminus \Omega_k](I_n - \hat{A}[\Omega_k][\hat{A}^*[\Omega_k]\hat{A}[\Omega_k])^{-1}\hat{A}^*[\Omega_k])\hat{A}[\Omega_s \setminus \Omega_k]x[\Omega_s \setminus \Omega_k]
\]
\[
= \hat{x}^*[\Omega_s \setminus \Omega_k]\Omega_k\hat{x}^*[\Omega_s \setminus \Omega_k])x[\Omega_s \setminus \Omega_k]
\]
\[
\geq (1 - \delta_{K+1}) \min_{1 \leq i \leq s} ||x[i]||_2^2.
\]
On the other hand, we have
\[
\|P^*[\bar{H}_{\Omega s}]\tilde{A}x\|_2^2 = \|P^*[\bar{H}_{\Omega s}]\tilde{A}x\|_2^2 = \left\langle P^*[\bar{H}_{\Omega s}]\tilde{A}x, P^*[\bar{H}_{\Omega s}]\sum_{i=1}^{s} \tilde{A}[i]x[i] \right\rangle_2
\]
\[
= \left\langle P^*[\bar{H}_{\Omega s}]\tilde{A}x, \sum_{i=1}^{s} \tilde{A}[i]x[i] \right\rangle_2
\]
\[
\leq \|x(\Omega_s \setminus \Omega_k)\|_{2, \infty} \max_{1 \leq i \leq s} \|\tilde{A}[i]^* (P^*[\bar{H}_{\Omega s}]\tilde{A}x)\|_2
\]
\[
= \|x(\Omega_s \setminus \Omega_k)\|_{2, \infty} \|\tilde{A}^* (P^*[\bar{H}_{\Omega s}]\tilde{A}x)\|_{2, \infty}.
\] (A.27)

From Lemma 2.6, (A.26) and (A.27), we get
\[
\|\tilde{A}^* P^*[\bar{H}_{\Omega s}]\tilde{A}x - \tilde{y}\|_{2, \infty} \geq (1 - \delta_{k+1}) \frac{\|x(\Omega_s \setminus \Omega_k)\|_2}{\|x(\Omega_s \setminus \Omega_k)\|_{2, \infty}} \geq \frac{1 - \delta_{k+1}}{\sqrt{s - k}} \min_{1 \leq i \leq s} \|x[i]\|_2 \geq (1 - \delta_{k+1}) \min_{1 \leq i \leq s} \|x[i]\|_2.
\] (A.28)

In the following, by Lemma 2.9 and a simple calculation, we have
\[
\|\tilde{A}^* P^*[\bar{H}_{\Omega s}]\tilde{A}x - \tilde{y}\|_{2, \infty}
\]
\[
= \max_{1 \leq i \leq N} \|\tilde{A}[i]^* (P^*[\bar{H}_{\Omega s}]\tilde{A}x - \tilde{y})\|_2 = \max_{1 \leq i \leq N} \|\tilde{A}x - \tilde{y}\|^* (P^*[\bar{H}_{\Omega s}]\tilde{A}[i])\|_2
\]
\[
\leq \left( 1 + \frac{\sqrt{s - 1}}{1 - \delta_{k+1}} \right) c_{A,K,Y} \leq \left( 1 + \frac{\sqrt{s - 1}}{1 - \delta_{k+1}} \right) c_{A,K,Y} \leq \left( 1 + \frac{\sqrt{s - 1}}{1 - \delta_{k+1}} \right) c_{A,K,Y}.
\] (A.29)

That is, the above (A.25), (A.28) and (A.29) show that (2.8) holds. So the lemma holds.

\[\Box\]

B Appendix

Proof of Theorem 3.1. The proof can be divided into two parts in the first step of Theorem 3.1. We first prove that the BOMP algorithm selects correct indexes in all iterations. Next, we prove that BOMP performs exactly s iterations. Induction is the first step in proof. Suppose that the BOMP algorithm selects correct indexes in the first k iterations, i.e., \(A_k \subset \Omega_s\) with \(k < s\). The assumption that we have \(A_0 = \emptyset \subset \Omega_s\) when \(k = 0\) holds evidently. Next, we need to show that the BOMP algorithm selects a correct index in the \(k + 1\)th iteration, i.e., \(A_{k+1} \subset \Omega_s\).

Since \(r^k = P^*[\bar{H}_{\Omega s}]\tilde{y}\), by Lemma 2.8, we have \(S_{0 \backslash A_k} - S_{A_k} > 0\) for any \(k = 0, 1, \ldots, s - 1\) with \(s < j \leq N\). Let \(P^*[\bar{H}_{\Omega s}]\tilde{y} = \sum_{i=1}^{s} \tilde{A}[i]\tilde{x}[i] = \tilde{A}[\Omega_s]x\) and \(\tilde{A}x = \sum_{i=1}^{s} \tilde{A}[i]x[i]\). Thus, from (2.4) and (3.2), we have
\[
\min_{1 \leq i \leq s} \|\tilde{x}[i]\|_2 = \min_{1 \leq i \leq s} \|x[i]\|_2 = \frac{c_{A,K,Y}}{\sqrt{1 - \delta_{k+1}}} > \frac{c_{A,K,Y}}{\sqrt{1 - ((1 + \delta_{k+1})^2 - \delta_{k+1})}}
\]
\[
= \frac{\sqrt{(1 + \delta_{k+1})^2 - \delta_{k+1}^2} \frac{c_{A,K,Y}}{\sqrt{1 - \delta_{k+1}}} - \frac{c_{A,K,Y}}{\sqrt{1 - \delta_{k+1}}}}{1 - [(1 + \delta_{k+1})^2 - \delta_{k+1}^2] \sqrt{K + 1}} > 0.
\]

Thus, it suffices to show that \(\|x[2]\|_2 \cdots \|x[s]\|_2 \neq 0\). According to the definition of BRIP, it is easy to check that
\[
\|\tilde{A}[j]\|_2 = \|\tilde{A}[j]\|_2 \leq \sqrt{1 + \delta_{k+1}}.
\]
Thus, we obtain

\[
\|(P^i[\hat{H}_{\Omega_s}](\hat{A}x - \hat{y}))^* \hat{A} [j]\|_2 \leq \|(P^i[\hat{H}_{\Omega_s}](\hat{A}x - \hat{y}))\|_2 \leq \sqrt{1 + \delta_{K+1}^{1+e'_{A,K,y}}}.
\]

(B.1)

By (2.2) and (B.1), we have

\[
S_{0A_k} - S_{jA_k} \geq \frac{\|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2}{\|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2} - \delta_{K+1} \frac{\|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2}{\|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2} \left(1 + \frac{\|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2}{\|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2} - \sqrt{1 + \delta_{K+1}^{1+e'_{A,K,y}}}\right).
\]

(B.2)

Let \(f(x) = x - \delta_{K+1} \sqrt{1 + x^2}\) with \(x \in [0, \infty)\). By simple calculations, we have \(f'(x) = 1 - \delta_{K+1} \frac{2 x}{\sqrt{1 + x^2}} > 0\), where \(x \in [0, \infty)\) and \(0 < \delta_{K+1} \leq 1\), i.e., the function \(f(x) = x - \delta_{K+1} \sqrt{1 + x^2}\) is monotonously increasing on the interval \([0, \infty)\). It is easy to verify that

\[
\|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2 \geq \frac{1}{\sqrt{S - k}} \text{ and } \|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2 \geq \frac{\sqrt{S - k}}{\min_{1 \leq i \leq S} \|\hat{x}[i]\|_2}.
\]

by Lemma 2.6. Using (B.2), we obtain

\[
S_{0A_k} - S_{jA_k} \geq \|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2 \left(1 - \delta_{K+1} \frac{\sqrt{S - k + 1}}{\sqrt{S - k}} - \sqrt{1 + \delta_{K+1}^{1+e'_{A,K,y}}}/\min_{1 \leq i \leq S} \|\hat{x}[i]\|_2\right).
\]

We use (2.4), (3.2) and (B.2) to get

\[
S_{0A_k} - S_{jA_k} \geq \|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2 \left(1 - \delta_{K+1} \frac{\sqrt{S - k + 1}}{\sqrt{S - k}} + \|\hat{x}[\Omega_s \setminus \Lambda_k]\|_2 \left(1 + \delta_{K+1}^{1+e'_{A,K,y}}\right)\right).
\]

Then the BOMP algorithm selects a correct index in the \(k+1\)-th iteration. Next, we show that the BOMP algorithm with the stopping criterion \(\|r^k\|_2 \leq e'_{A,K,y}\) performs exactly \(s\) iterations, which is equivalent to show that \(\|r^k\|_2 > e'_{A,K,y}\) for \(k = 0, 1, 2, \ldots, s-1\) and \(\|r^s\|_2 \leq e'_{A,K,y}\). Since the BOMP algorithm selects a correct index in each iteration under (3.2), we assume that \(\Lambda_k = \Omega_s\) for \(0 \leq k \leq s\). It follows from (2.6) and (3.2) that

\[
\|r^0\|_2 \geq \|r^1\|_2 \geq \cdots \geq \|r^{s-1}\|_2 \geq \sqrt{1 - \delta_{K+1}^{1+e'_{A,K,y}}} \min_{1 \leq i \leq S} \|\hat{x}[i]\|_2 - e'_{A,K,y}\]

\[
\geq \sqrt{1 - [(1 + \delta_{K+1}^{e'_{A,K,y}})(1 + e_{A,K,y}^{(K+1)})^2 - 1]} \min_{1 \leq i \leq S} \|\hat{x}[i]\|_2 - e'_{A,K,y}\]

\[
\geq \sqrt{1 - [(1 + \delta_{K+1}^{e'_{A,K,y}})(1 + e_{A,K,y}^{(K+1)})^2 - 1]} \sqrt{\delta_{K+1}^{1+e_{A,K,y}}}/\sqrt{1 + K + 1}\]

\[
\geq \sqrt{1 - [(1 + \delta_{K+1}^{e'_{A,K,y}})(1 + e_{A,K,y}^{(K+1)})^2 - 1]} e'_{A,K,y}/\sqrt{1 + K + 1}\]

\[
\geq \sqrt{1 - [(1 + \delta_{K+1}^{e'_{A,K,y}})(1 + e_{A,K,y}^{(K+1)})^2 - 1]} e'_{A,K,y}/\sqrt{1 + K + 1}\]

i.e., the BOMP algorithm does not terminate before performing the \(s\)-th iteration.

By Lemma 2.2, we have

\[
\|r^n\|_2 = \|P^i[\hat{H}_{\Omega_s}](\hat{A}x + \hat{y})\|_2 \leq \|P^i[\hat{H}_{\Omega_s}](\hat{A}x + \hat{y})\|_2 \leq \|P^i[\hat{H}_{\Omega_s}](\hat{A}x + \hat{y})\|_2 \leq \|\hat{A}x - \hat{y}\|_2 \leq e'_{A,K,y}.
\]

Thus, the BOMP algorithm terminates after performing the \(s\)-th iteration, i.e., the BOMP algorithm performs \(s\) iterations. So Theorem 3.1 is proved.
Proof of Theorem 3.3. For any $\hat{A}[j]$ with $s < j \leq N$ and any $k \in \mathbb{N}$ with $0 \leq k < s$, we first prove $S_{0,A_k} - S_{|A_k|} > 0$ for $A_k \subset \Omega_s$. It follows from (2.5) and (3.3) that

$$
\min_{1 \leq i \leq s} \|x[i]\|_2 \geq \min_{1 \leq i \leq s} \|x[i]\|_2 \geq \frac{\sqrt{\mathcal{R}_{e^r_{A,K},y}}}{1 - \hat{\delta}_{K+1}} > \frac{\sqrt{\mathcal{R}_{e^r_{A,K},y}}}{1 - ((1 + \hat{\delta}_{K+1})(1 + e_{A}^{(K+1)})^2 - 1)}
+ \frac{1 + \sqrt{\mathcal{R}(1 + \delta_{K+1})(1 + e_{A}^{(K+1)})^2 - 1}}{1 - ((1 + \delta_{K+1})(1 + e_{A}^{(K+1)})^2 - 1)} e_{A,K,y}'
- \frac{\sqrt{\mathcal{R}_{e^r_{A,K},y}}}{1 - \hat{\delta}_{K+1}} \geq \frac{1 + \sqrt{\mathcal{R}(1 + \delta_{K+1})(1 + e_{A}^{(K+1)})^2 - 1}}{1 - ((1 + \delta_{K+1})(1 + e_{A}^{(K+1)})^2 - 1)} e_{A,K,y}' > 0.
$$

Thus, it suffices to show that $\|x[2]\|_2 \cdots \|x[s]\|_2 \neq 0$. Then, by (2.3), we obtain

$$
\|\left(P^t[H_{\Omega,s}]\hat{A}x - \hat{y}\right)^t \hat{A}[j]\|_2 = \|\left(P^t[H_{\Omega,s}]\hat{A}[j]\right)\|_2 
\leq \left(1 + \frac{\sqrt{s} \hat{\delta}_{K+1}}{1 - \hat{\delta}_{K+1}}\right) e_{A,K,y}' \leq \left(1 + \frac{\sqrt{\mathcal{R} \hat{\delta}_{K+1} \delta_{K+1}}}{1 - \hat{\delta}_{K+1}}\right) e_{A,K,y}'.
$$

By (2.2) and (B.1), we have

$$
S_{0,A_k} - S_{|A_k|} \geq \frac{\|x[\Omega \setminus A_k]\|_2^2}{\|x[\Omega \setminus A_k]\|_2,2} - \hat{\delta}_{K+1} \|x[\Omega \setminus A_k]\|_2,2 \frac{\|\hat{y}[\Omega \setminus A_k]\|_2,2}{\|\hat{y}[\Omega \setminus A_k]\|_2,2} - \left(1 + \frac{\sqrt{\mathcal{R} \hat{\delta}_{K+1}}} {1 - \hat{\delta}_{K+1}}\right) e_{A,K,y}'
= \|x[\Omega \setminus A_k]\|_2 \left(\frac{\|\hat{y}[\Omega \setminus A_k]\|_2,2}{\|x[\Omega \setminus A_k]\|_2,2} - \frac{1}{\sqrt{s - k}} - \frac{1}{\sqrt{s - k}} \min_{1 \leq l \leq s} \|x[l]\|_2 \right)
\geq \|x[\Omega \setminus A_k]\|_2 \left(\frac{1}{\sqrt{s - k}} - \hat{\delta}_{K+1} \frac{\sqrt{s - k} - 1}{\sqrt{s - k}} - \frac{1}{\sqrt{s - k}} \min_{1 \leq l \leq s} \|x[l]\|_2 \right).
$$

By combining (2.5), (3.3) and (B.3), we have

$$
S_{0,A_k} - S_{|A_k|} \geq \|x[\Omega \setminus A_k]\|_2 \left(1 - \hat{\delta}_{K+1} \frac{\sqrt{s - k} + 1}{\sqrt{s - k}} - \frac{1}{\min_{1 \leq l \leq s} \|x[l]\|_2} \right)
\geq \|x[\Omega \setminus A_k]\|_2 \left(1 + \frac{\sqrt{\mathcal{R} \hat{\delta}_{K+1} \delta_{K+1}}}{1 - \hat{\delta}_{K+1}}\right) e_{A,K,y}'
\geq \|x[\Omega \setminus A_k]\|_2 \left(1 + \frac{\sqrt{\mathcal{R} \hat{\delta}_{K+1} \delta_{K+1}}}{1 - \hat{\delta}_{K+1}}\right) e_{A,K,y}'.
$$

Thus, the BOMP algorithm selects a correct index at each iteration. Next, we show that the BOMP algorithm with the stopping criterion

$$
\|\hat{A}^* r^k\|_{2,\infty} \leq \left(1 + \frac{\sqrt{\mathcal{R} \hat{\delta}_{K+1}}}{1 - \hat{\delta}_{K+1}}\right) e_{A,K,y}'
$$

performs exactly $s$ iterations, which is equivalent to show that

$$
\|\hat{A}^* r^k\|_{2,\infty} > \left(1 + \frac{\sqrt{\mathcal{R} \hat{\delta}_{K+1}}}{1 - \hat{\delta}_{K+1}}\right) e_{A,K,y}' \quad \text{for} \quad k = 0, 1, 2, \ldots, s - 1,
$$

$$
\|\hat{A}^* r^s\|_{2,\infty} \leq \left(1 + \frac{\sqrt{\mathcal{R} \hat{\delta}_{K+1}}}{1 - \hat{\delta}_{K+1}}\right) e_{A,K,y}'.
$$

Similarly, we assume that $A_k = \Omega_k$ for $0 \leq k \leq s$. It follows from (2.8) and (3.3) that, for any $0 \leq k < s$,

$$
\|\hat{A}^* r^k\|_{2,\infty} = \|\hat{A}^* P^t[H_{\Omega,s}]\|_{2,\infty} \geq [1 - ((1 + \delta_{K+1})(1 + e_{A}^{(K+1)})^2 - 1)] \min_{1 \leq l \leq s} \|x[l]\|_2 - \left(1 + \frac{\sqrt{\mathcal{R} - 1 \hat{\delta}_{K+1}}}{1 - \hat{\delta}_{K+1}}\right) e_{A,K,y}'.
$$
from Lemma 2.4, (B.4) and the definition of BRIP, for

When

It is easy to check that

\[ K < \frac{1}{\sqrt{K+1}(1+\epsilon'_A (K+1)^2 - 1)} + \frac{1}{(1+\epsilon'_A (K+1)^2 - 1)} - 1, \]

if \( K = 1 \), it is obvious that

\[ \sqrt{K - 1} + \frac{(\sqrt{K+1} - \sqrt{K-1} - 1)((1+\delta_{K+1})(1+\epsilon'_A (K+1)^2 - 1))}{1 - [(1+\delta_{K+1})(1+\epsilon'_A (K+1)^2 - 1)]} \geq 0, \]

for \( K > 1 \). It is easy to check that

\[ \sqrt{K+1} - \sqrt{K-1} - 1 < 0, \]

\[ \frac{(1+\delta_{K+1})(1+\epsilon'_A (K+1)^2 - 1)}{1 - [(1+\delta_{K+1})(1+\epsilon'_A (K+1)^2 - 1)]} \leq \frac{1}{(\sqrt{K+1} - 1)}, \]

Thus,

\[ \sqrt{K-1} + \frac{(\sqrt{K+1} - \sqrt{K-1} - 1)((1+\delta_{K+1})(1+\epsilon'_A (K+1)^2 - 1))}{1 - [(1+\delta_{K+1})(1+\epsilon'_A (K+1)^2 - 1)]} \geq \sqrt{K} - \frac{\sqrt{K-1}}{\sqrt{K+1} - 1} > 0. \]

So the BOMP algorithm does not terminate before performing the \( s \)-th iteration.

In the following, by (2.7), we can easily get

\[ \|\hat{A}^* \tau\|_{2,\infty} = \|\hat{A}^* P^j [\hat{H}_{\Omega_s}] \hat{y}\|_{2,\infty} \leq \left(1 + \sqrt{K} \delta_{K+1} \right) e''_{\hat{A},k,y} \leq \left(1 + \sqrt{K} \delta_{K+1} \right) e''_{\hat{A},k,y}. \]

Therefore, the BOMP algorithm terminates after performing the \( s \)-th iteration, i.e., the BOMP algorithm performs \( s \) iterations. So Theorem 3.3 is proved.

**Proof of Theorem 3.5.** Since

\[ \hat{A}(\Omega_s) x^* = P[H_{\Omega_s}] \hat{y} = P[H_{\Omega_s}] (\hat{A} \hat{x} - (\hat{A} \hat{x} - \hat{y})) = P[H_{\Omega_s}] (\hat{A}(\Omega_s) x[\Omega_s] - (\hat{A} \hat{x} - \hat{y})), \]

from Lemma 2.4, (B.4) and the definition of BRIP, for \( \|\hat{A} \hat{x} - \hat{y}\|_2 \leq e'_{\hat{A},k,y} \), we have

\[ \|x^* - x[\Lambda_s]\|_2 \leq \frac{\|\hat{A}(\Omega_s)(x^* - x[\Lambda_s])\|_2}{\sqrt{1 - \delta_{\hat{K}}}} = \frac{\|P[H_{\Omega_s}] (\hat{A} \hat{x} - \hat{y})\|_2}{\sqrt{1 - \delta_{\hat{K}}}}. \]
It follows from (B.5) and (B.6) that
\[ \frac{\|Ax - \hat{y}\|_2}{\sqrt{1 - \delta_k}} \leq \frac{c'_{A,K,y}}{\sqrt{1 - \delta_k}} \leq \frac{c'_{A,K,y}}{\sqrt{2 - (1 + \delta_k)(1 + \epsilon_K^{(K)})^2}}, \]
and for \( \|A^*(Ax - \hat{y})\|_2 \leq c''_{A,K,y} \), we also have
\[
(1 - \delta_k)\|x^* - x[L_\delta]\|_2 \leq \|A[\Omega_\delta](x^* - x[L_\delta])\|_2^2 = \|P[\hat{A}_\Omega](Ax - \hat{y})\|_2^2 \\
= \|A[\Omega_\delta](A^*[\Omega_\delta]\hat{A}[\Omega_\delta])^{-1}A^*[\Omega_\delta](Ax - \hat{y})\|_2^2 \\
= (Ax - \hat{y})^*\hat{A}[\Omega_\delta](A^*[\Omega_\delta]\hat{A}[\Omega_\delta])^{-1}A^*[\Omega_\delta](Ax - \hat{y}). \quad (B.5)
\]
Using \( (1 - \delta_k)I_s \leq A^*[\Omega_\delta]\hat{A}[\Omega_\delta] \leq (1 + \delta_k)I_s \), we have
\[
\frac{1}{1 + \delta_k}I_s \leq (A^*[\Omega_\delta]\hat{A}[\Omega_\delta])^{-1} \leq \frac{1}{1 - \delta_k}I_s. \quad (B.6)
\]
It follows from (B.5) and (B.6) that
\[
(1 - \delta_k)\|x^* - x[L_\delta]\|_2^2 \leq \frac{1}{1 - \delta_k}\|A^*[\Omega_\delta](Ax - \hat{y})\|_2^2.
\]
According to the above inequality, we have
\[
\|x^* - x[L_\delta]\|_2 \leq \frac{1}{1 - \delta_k}\|A^*[\Omega_\delta](Ax - \hat{y})\|_2 \leq \frac{\sqrt{S}}{1 - \delta_k}\|A^*[\Omega_\delta](Ax - \hat{y})\|_2 \leq \frac{\sqrt{S}}{1 - \delta_k}\|A^*[\Omega_\delta](Ax - \hat{y})\|_{2,\infty} \leq \frac{\sqrt{S}}{1 - \delta_k}\|A^*[\Omega_\delta](Ax - \hat{y})\|_2 \leq \frac{\sqrt{S}}{1 - \delta_k}c''_{A,K,y}.
\]
So Theorem 3.5 is proved.

**Proof of Theorem 3.7.** We first suppose that the BOMP algorithm selects correct indexes in the first \( k \) iterations, i.e., \( L_k \subset \Omega_s \) with \( k < s \). Since \( L_0 = \emptyset \subset \Omega_s \), the assumption holds obviously when \( k = 0 \). Next, we prove \( S_{L_k} - S_{L_0} > 0 \) for all \( s \leq j \leq N \). When \( \|Ax - \hat{y}\|_2 = 0 \), we have \( \hat{x} = x[L_\delta] \). It follows from (2.2) that
\[
S_{L_k} - S_{L_0} = \frac{\|x[L_k\backslash L_0]\|_2^2}{\|x[L_0]\|_2} - \delta_{K+1} - \|x[L_k\backslash L_0]\|_2^2 \|x[L_0]\|_2^2 (\|x[L_0\backslash L_k]\|_2^2 + \|x[L_0\backslash L_k]\|_2^2)^2 \\
= \|x[L_k\backslash L_0]\|_2^2 \|x[L_0\backslash L_k]\|_2^2 + \|x[L_0\backslash L_k]\|_2^2 \left( \frac{1}{\sqrt{1 + \frac{\|x[L_0\backslash L_k]\|_2^2}{\|x[L_0]\|_2^2}}^2} - \frac{\delta_{K+1}}{2} \right).
\]
According to the definition of \( g(x) \), we have \( g(x) \geq \frac{\|x[L_0\backslash L_k]\|_2^2}{\|x[L_0]\|_2^2} \). It follows from (3.6) that
\[
S_{L_k} - S_{L_0} \geq \|x[L_k\backslash L_0]\|_2^2 \|x[L_0\backslash L_k]\|_2^2 \left( \frac{1}{\sqrt{1 + g^2(x)}} - \frac{\delta_{K+1}}{2} \right) \geq (1 + \delta_{K+1})(1 + \epsilon_K^{(K+1)})^2 \left( \frac{1}{\sqrt{1 + g^2(x)}} - 1 \right) > 0.
\]
Then the BOMP algorithm can exactly recover the block \( K \)-sparse signal in \( K \) iterations, so Theorem 3.7 is proved.

**Funding:** This research was supported by Natural Science Foundation of China (No. 61673015), the Science and Technology Research Program of Chongqing Municipal Education Commission (No. KJQN201802002, No. KJQN201802001).

**References**


