Journal of Fourier Analysis and Applications (2023) 29:36
https://doi.org/10.1007/s00041-023-10022-6

Affine Phase Retrieval for Sparse Signals via ℓ_1 Minimization

Meng Huang[1] · **Shixiang Sun**[2,3] · **Zhiqiang Xu**[2,3]

Received: 21 September 2021 / Revised: 27 February 2023 / Accepted: 30 May 2023 /
Published online: 12 June 2023

Abstract

Affine phase retrieval is the problem of recovering signals from the magnitude-only measurements with a priori information. In this paper, we use the ℓ_1 minimization to exploit the sparsity of signals for affine phase retrieval, showing that $O(k \log(en/k))$ Gaussian random measurements are sufficient to recover all k-sparse signals by solving a natural ℓ_1 minimization program, where n is the dimension of signals. For the case where measurements are corrupted by noises, the reconstruction error bounds are given for both real-valued and complex-valued signals. Our results demonstrate that the natural ℓ_1 minimization program for affine phase retrieval is stable.

Keywords Phase retrieval · Sparse signals · ℓ_1 minimization · Compressed sensing

Mathematics Subject Classification 94A12 · 60B20

Communicated by Jaming Philippe.

✉ Meng Huang
menghuang@buaa.edu.cn

Shixiang Sun
sunshixiang@lsec.cc.ac.cn

Zhiqiang Xu
xuzq@lsec.cc.ac.cn

1 School of Mathematical Sciences, Beihang University, Beijing 100191, China

2 LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

3 School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

1 Introduction

1.1 Problem Setup

Affine phase retrieval for sparse signals aims to recover a k-sparse signal $\boldsymbol{x}_0 \in \mathbb{F}^n$, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, from the observed data

$$y_j = |\langle \boldsymbol{a}_j, \boldsymbol{x}_0 \rangle + b_j| + w_j, \quad j = 1, \ldots, m,$$

where $\boldsymbol{a}_j \in \mathbb{F}^n,\ j = 1, \ldots, m$ are given measurement vectors, $\boldsymbol{b} := (b_1, \ldots, b_m)^{\mathrm{T}} \in \mathbb{F}^m$ is the given bias vector, and $\boldsymbol{w} := (w_1, \ldots, w_m)^{\mathrm{T}} \in \mathbb{R}^m$ is the noise vector. The affine phase retrieval arises in several practical applications, such as holography [2, 20, 26, 27] and Fourier phase retrieval [3–5, 23], where some side information of signals is a priori known before capturing the magnitude-only measurements.

The aim of this paper is to study the following program to recover $\boldsymbol{x}_0$ from $\boldsymbol{y} := (y_1, \ldots, y_m)^{\mathrm{T}} \in \mathbb{R}^m$:

$$\min_{\boldsymbol{x} \in \mathbb{C}^n} \|\boldsymbol{x}\|_1 \quad \text{s.t. } \||\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}| - \boldsymbol{y}\|_2 \leq \epsilon, \tag{1}$$

where $\boldsymbol{A} := [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m]^* \in \mathbb{F}^{m \times n}$.

Particularly, we focus on the following questions:

Question 1: Assume that $\boldsymbol{a}_j,\ j = 1, \ldots, m$, are Gaussian random measurements with $m = O(k \log(en/k))$. In the absence of noise, i.e., $\boldsymbol{w} = 0$, $\epsilon = 0$, is the solution to (1) $\boldsymbol{x}_0$?

Question 2: In the noisy scenario, is the program (1) stable under small perturbation?

For the case where $\boldsymbol{x}_0 \in \mathbb{C}^n$ is non-sparse, it was shown that $m \geq 4n - 1$ generic measurements are sufficient to guarantee the uniqueness of solutions in [19], and several efficient algorithms with linear convergence rate was proposed to recover the non-sparse signals $\boldsymbol{x}_0$ from $\boldsymbol{y}$ under $m = O(n \log n)$ Gaussian random measurements in [25]. However, for the case where $\boldsymbol{x}_0$ is sparse, to the best of our knowledges, there is no result about it.

1.2 Related Works

1.2.1 Phase Retrieval

The noisy phase retrieval is the problem of recovering a signal $\boldsymbol{x}_0 \in \mathbb{F}^n$, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ from the magnitude-only measurements

$$y_j' = \left|\langle \boldsymbol{a}_j, \boldsymbol{x}_0 \rangle\right| + w_j, \quad j = 1, \ldots, m,$$

where $\boldsymbol{a}_j \in \mathbb{F}^n$ are given measurement vectors and $w_j \in \mathbb{R}$ are noises. It arises naturally in many areas such as X-ray crystallography [21, 22, 28], coherent diffractive imaging [30], and optics [14, 15, 32]. In these settings, optical detectors record only the intensity of a light wave while losing the phase information. Note that $\left|\langle \boldsymbol{a}_j, \boldsymbol{x}_0\rangle\right|^2 = \left|\langle \boldsymbol{a}_j, e^{i\theta}\boldsymbol{x}_0\rangle\right|^2$ for any $\theta \in \mathbb{R}$. Therefore the recovery of $\boldsymbol{x}_0$ for the classical phase retrieval is up to a global phase. In the absence of noise, it has been proved that $m \geq 2n - 1$ generic measurements suffice to guarantee the uniqueness of solutions for the real case [1], and $m \geq 4n - 4$ for the complex case [6, 13, 38], respectively. Moreover, several efficient algorithms have been proposed to reconstruct $\boldsymbol{x}_0$ from $\boldsymbol{y}' := [y_1', \ldots, y_m']^{\mathrm{T}}$, such as alternating minimization [29], truncated amplitude flow [37], smoothed amplitude flow [7], trust-region [33], and the Wirtinger flow (WF) variants [9, 10, 41].

1.2.2 Sparse Phase Retrieval

For several applications, the underlying signal is naturally sparse or admits a sparse representation after some linear transformation. This leads to the sparse phase retrieval:

$$\min_{\boldsymbol{x}\in\mathbb{F}^n} \quad \|\boldsymbol{x}\|_0 \quad \text{s.t. } \||\boldsymbol{A}\boldsymbol{x}| - \boldsymbol{y}'\|_2 \leq \epsilon, \tag{2}$$

where $\boldsymbol{A} := [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m]^*$. In the absence of noise, it has been established that $m = 2k$ generic measurements are necessary and sufficient for uniquely recovering of all k-sparse signals in the real case, and $m \geq 4k - 2$ are sufficient in the complex case [39]. In the noisy scenario, $O(k \log(en/k))$ measurements suffice for stable sparse phase retrieval [12]. Due to the hardness of ℓ_0-norm in (2), a computationally tractable approach to recover $\boldsymbol{x}_0$ is by solving the following ℓ_1 minimization:

$$\min_{\boldsymbol{x}\in\mathbb{F}^n} \quad \|\boldsymbol{x}\|_1 \quad \text{s.t. } \||\boldsymbol{A}\boldsymbol{x}| - \boldsymbol{y}'\|_2 \leq \epsilon. \tag{3}$$

For the real case, based on the strong restricted isometry property (SRIP) established by Voroninski and Xu [34], the authors in [18] proved that, if $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \sim 1/\sqrt{m} \cdot \mathcal{N}(0, I_n)$ are i.i.d. Gaussian random vectors with $m \geq O(k \log(en/k))$, then the solution $\widehat{\boldsymbol{x}} \in \mathbb{R}^n$ to (3) satisfies

$$\min\{\|\widehat{\boldsymbol{x}} - \boldsymbol{x}_0\|, \|\widehat{\boldsymbol{x}} + \boldsymbol{x}_0\|\} \lesssim \epsilon + \frac{\sigma_k(\boldsymbol{x}_0)_1}{\sqrt{k}},$$

where $\sigma_k(\boldsymbol{x}_0)_1 := \min_{|\mathrm{supp}(\boldsymbol{x})|\leq k} \|\boldsymbol{x} - \boldsymbol{x}_0\|_1$. Lately, this result was extended to the complex case by employing the "phaselift" technique in [36]. Specifically, the authors in [36] showed that, for any k-sparse signal $\boldsymbol{x}_0 \in \mathbb{C}^n$, the solution $\widehat{\boldsymbol{x}} \in \mathbb{C}^n$ to the program

$$\operatorname*{argmin}_{\boldsymbol{x}\in\mathbb{C}^n} \quad \|\boldsymbol{x}\|_1 \quad \text{s.t. } \|\mathcal{A}(\boldsymbol{x}) - \mathcal{A}(\boldsymbol{x}_0)\|_2 \leq \epsilon$$

satisfies

$$\min_{\theta\in[0,2\pi)} \|\widehat{\boldsymbol{x}} - e^{i\theta}\boldsymbol{x}_0\|_2 \lesssim \frac{\epsilon}{\sqrt{m}\|\boldsymbol{x}_0\|_2},$$

provided $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \sim \mathcal{N}(0, I_n)$ are i.i.d. complex Gaussian random vectors and $m \geq O(k \log(en/k))$. Here, $\mathcal{A}(\boldsymbol{x}) := (\left|\boldsymbol{a}_1^* \boldsymbol{x}\right|^2, \ldots, \left|\boldsymbol{a}_m^* \boldsymbol{x}\right|^2)$.

1.2.3 Affine Phase Retrieval

The affine phase retrieval aims to recover a signal $\boldsymbol{x}_0 \in \mathbb{F}^n$, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, from the measurements

$$y_j = \left|\langle \boldsymbol{a}_j, \boldsymbol{x}_0\rangle + b_j\right|, \quad j = 1, \ldots, m,$$

where $\boldsymbol{a}_j \in \mathbb{F}^n$, $j = 1, \ldots, m$ are measurement vectors, $\boldsymbol{b} := (b_1, \ldots, b_m)^{\mathrm{T}} \in \mathbb{F}^m$ is the bias vector. The problem can be regarded as the classic phase retrieval with a priori information, and is raised in many areas, such as holographic phase retrieval [16, 17, 27] and Fourier phase retrieval [3–5, 23]. In such scenarios, one needs to employ some additional information about the desired signals to ensure the uniqueness of solutions. Specifically, in holographic optics, a reference signal $\boldsymbol{r} \in \mathbb{C}^k$, whose structure is a priori known, is included in the diffraction patterns alongside the signal of interest $\boldsymbol{x}_0 \in \mathbb{C}^n$ [2, 20, 26]. Set $\boldsymbol{x}_0' = (\boldsymbol{x}_0^{\mathrm{T}}, \boldsymbol{r}^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{C}^{n+k}$. Then the magnitude-only measurements we obtain that

$$y_j = |\langle \boldsymbol{a}_j', \boldsymbol{x}_0'\rangle| = |\langle \boldsymbol{a}_j, \boldsymbol{x}_0\rangle + \langle \boldsymbol{a}_j'', \boldsymbol{r}\rangle| = |\langle \boldsymbol{a}_j, \boldsymbol{x}_0\rangle + b_j|, \quad j = 1, \ldots, m,$$

where $\boldsymbol{a}_j' = (\boldsymbol{a}_j^{\mathrm{T}}, \boldsymbol{a}_j''^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{C}^{n+k}$ are given measurement vectors and $b_j = \langle \boldsymbol{a}_j'', \boldsymbol{r}\rangle \in \mathbb{C}$ are known. Therefore, the holographic phase retrieval can be viewed as the affine phase retrieval.

Another application of affine phase retrieval arises in Fourier phase retrieval problem. For one-dimensional Fourier phase retrieval problem, it usually does not possess the uniqueness of solutions [35]. Actually, for a given signal with dimension n, beside the trivial ambiguities caused by shift, conjugate reflection and rotation, there still could be 2^{n-2} nontrivial solutions. To enforce the uniqueness of solutions, one approach is to use additionally known values of some entries [4], which can be recast as affine phase retrieval. More related works on the uniqueness of solutions for Fourier phase retrieval can be seen in [11, 31].

1.3 Our Contributions

In this paper, we focus on the recovery of sparse signals from the magnitude of affine measurements. Specifically, we aim to recover a k-sparse signal $\boldsymbol{x}_0 \in \mathbb{F}^n$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$) from the data

$$\boldsymbol{y} = |\boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{b}| + \boldsymbol{w},$$

where $\boldsymbol{A} := [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m]^* \in \mathbb{F}^{m \times n}$ is the measurement matrix, $\boldsymbol{b} \in \mathbb{F}^m$ is the bias vector, and $\boldsymbol{w} \in \mathbb{R}^m$ is the noise vector. Our aim is to present the performance of the following ℓ_1 minimization program:

$$\underset{\boldsymbol{x}\in\mathbb{F}^n}{\operatorname{argmin}}\|\boldsymbol{x}\|_1 \quad \text{s.t. } \||\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}|-\boldsymbol{y}\|_2 \le \epsilon. \tag{4}$$

We say a triple $(\boldsymbol{A}, \boldsymbol{b}, \Delta)$ is *instance optimal of order* k_0 if it holds

$$\|\Delta(|\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}|)-\boldsymbol{x}\|_p \le C\cdot\sigma_{k_0}(\boldsymbol{x})_q \tag{5}$$

for all $\boldsymbol{x} \in \mathbb{F}^n$. Here, $\Delta : \mathbb{R}^m \to \mathbb{F}^n$ is a decoder for reconstructing $\boldsymbol{x}$, $\sigma_k(\boldsymbol{x})_q := \min_{|\operatorname{supp}(z)|\le k}\|z-\boldsymbol{x}\|_q$ and $C := C_{k_0,p,q}$ is a constant depending on k_0, p and q.

Theorem 1 *Assume that there exists a matrix $\boldsymbol{A} \in \mathbb{F}^{m\times n}$, a vector $\boldsymbol{b} \in \mathbb{F}^m$, a decoder $\Delta : \mathbb{F}^m \to \mathbb{F}^n$ and positive integers $k_0,\ p,\ q$ such that (5) holds for all $\boldsymbol{x} \in \mathbb{F}^n$. Then $\boldsymbol{b} \notin \{\boldsymbol{A}z : z \in \mathbb{F}^n\}$.*

Proof We assume that $\boldsymbol{b} = \boldsymbol{A}z_0$ where $z_0 \in \mathbb{F}^n$. We next show that there exits $\boldsymbol{x} \in \mathbb{F}^n$ such that (5) does not hold. For the aim of contradiction, we assume that (5) holds. Since $\sigma_{k_0}(-\boldsymbol{x})_q = \sigma_{k_0}(\boldsymbol{x})_q$, we have

$$\|\Delta(|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}|)+\boldsymbol{x}\|_p = \|\Delta(|\boldsymbol{A}(-\boldsymbol{x})+\boldsymbol{b}|)-(-\boldsymbol{x})\|_p \le C\sigma_{k_0}(\boldsymbol{x})_q. \tag{6}$$

Assume that $\boldsymbol{x}_0 \in \mathbb{F}^n$ is k_0-sparse, i.e. $\sigma_{k_0}(\boldsymbol{x}_0)_q = 0$. According to (5) and (6), we obtain that

$$\Delta(|\boldsymbol{A}\boldsymbol{x}_0+\boldsymbol{b}|) = \boldsymbol{x}_0, \quad \Delta(|\boldsymbol{A}\boldsymbol{x}_0-\boldsymbol{b}|) = -\boldsymbol{x}_0. \tag{7}$$

Taking $\boldsymbol{x} = r\boldsymbol{x}_0 + 2z_0$ in (6), we have

$$\|\Delta(|\boldsymbol{A}(r\boldsymbol{x}_0+2z_0)-\boldsymbol{b}|)+r\boldsymbol{x}_0+2z_0\|_p \le C\sigma_{k_0}(r\boldsymbol{x}_0+2z_0)_q \le C\sigma_{k_0}(2z_0)_q, \tag{8}$$

where $r > 0$. Observe that

$$\Delta(|\boldsymbol{A}(r\boldsymbol{x}_0+2z_0)-\boldsymbol{b}|) = \Delta(|\boldsymbol{A}(r\boldsymbol{x}_0)+\boldsymbol{b}|) = r\boldsymbol{x}_0. \tag{9}$$

Here, we use $\boldsymbol{x}_0$ is k_0-sparse. Substituting (8) into (9), we obtain that

$$\|2r\boldsymbol{x}_0+2z_0\|_p \le C\sigma_{k_0}(2z_0)_q \tag{10}$$

holds for any $r > 0$. Note $\lim_{r\to\infty}\|2r\boldsymbol{x}_0+2z_0\|_p = \infty$. Hence, (10) does not hold provided r is large enough. A contradiction! □

For the case where $m \le n$ and $\boldsymbol{A}$ is full rank, we have $\boldsymbol{b} \in \{\boldsymbol{A}z : z \in \mathbb{F}^n\}$. According to Theorem 1, we know that it is impossible to build the instance-optimality result under this setting. This is quite different from the earlier results on standard phase retrieval [18], where the instance-optimality is

$$\min_{|c|=1}\|\Delta(|\boldsymbol{A}\boldsymbol{x}|)-c\boldsymbol{x}\|_p \le C\cdot\sigma_{k_0}(\boldsymbol{x})_q, \text{ for all } \boldsymbol{x} \in \mathbb{F}^n. \tag{11}$$

The instance-optimality result for the standard phase retrieval, as expressed in equation (11), is established in [18].

1.3.1 Real Case

Our first result gives an upper bound for the reconstruct error of (4) in the real case, under the assumption of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in \mathbb{R}^n$ being real Gaussian random vectors and $m \geq O(k \log(en/k))$. It means the ℓ_1-minimization program is stable under small perturbation, even for the approximately k-sparse signals. To begin with, we need the following definition of strong RIP condition, which was introduced by Voroninski and Xu [34].

Definition 1 (Strong RIP in [34]) The matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ satisfies the Strong Restricted Isometry Property (SRIP) of order k and constants $\theta_l, \theta_u > 0$ if the following inequality

$$\theta_l \|\boldsymbol{x}\|^2 \leq \min_{I\subset[m],|I|\geq m/2} \|\boldsymbol{A}_I \boldsymbol{x}\|^2 \leq \max_{I\subset[m],|I|\geq m/2} \|\boldsymbol{A}_I \boldsymbol{x}\|^2 \leq \theta_u \|\boldsymbol{x}\|^2$$

holds for all k-sparse signals $\boldsymbol{x} \in \mathbb{R}^n$. Here, $\boldsymbol{A}_I$ denotes the sub-matrix of $\boldsymbol{A}$ whose rows with indices in I are kept, $[m] := \{1, \ldots, m\}$ and $|I|$ denotes the cardinality of I.

The following result indicates that the matrix $\left[\boldsymbol{A}\ \boldsymbol{b}\right] \in \mathbb{R}^{m\times(n+1)}$ satisfies strong RIP condition with high probability under some mild conditions on $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$.

Theorem 2 *Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ be a Gaussian random matrix with entries $a_{k,j} \sim \mathcal{N}(0, 1/m)$. Suppose that the vector $\boldsymbol{b} \in \mathbb{R}^m$ satisfies $\alpha \leq \|\boldsymbol{b}_I\|_2 \leq \beta$ for all $I \subseteq [m]$ with $|I| \geq m/2$, where $\alpha \leq \beta$ are two positive constants. Set $\boldsymbol{A}' := \left[\boldsymbol{A}\ \boldsymbol{b}\right] \in \mathbb{R}^{m\times(n+1)}$. If $m \geq Ct(k+1)\log(en/k)$ with $t(k+1) \leq n$ and $1 < t \in \mathbb{Z}$, then there exist constants θ_l', θ_u', independent with t, such that the matrix $\boldsymbol{A}'$ satisfies the strong RIP of order $tk+1$ and constants θ_l', θ_u' with probability at least $1 - 4\exp(-c'm)$. Here, $C, c' > 0$ are constants depending only on α and β.*

The following theorem shows that if we add some restrictions on the signal $\boldsymbol{x}$, then the instance-optimality result can be established.

Theorem 3 *Assume that $\boldsymbol{A}' := \left[\boldsymbol{A}\ \boldsymbol{b}\right] \in \mathbb{R}^{m\times(n+1)}$ satisfies the strong RIP of order $(a+1)(k+1)$ with constants $\theta_u \geq \theta_l > 0$. If $a > \theta_u/\theta_l$, then the following holds: for any vector $\boldsymbol{x}_0 \in \mathbb{R}^n$, the solution $\widehat{\boldsymbol{x}}$ to (4) with $\boldsymbol{y} = |\boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{b}| + \boldsymbol{w}$ and $\|\boldsymbol{w}\|_2 \leq \epsilon$ obeys*

$$\|\widehat{\boldsymbol{x}} - \boldsymbol{x}_0\|_2 \leq K_1\epsilon + K_2 \frac{\sigma_k(\boldsymbol{x}_0)_1}{\sqrt{a(k+1)}},$$

provided $K_1\epsilon + K_2 \frac{\sigma_k(\boldsymbol{x}_0)_1}{\sqrt{a(k+1)}} < 2$. Here,

$$K_1 := \frac{2\left(1 + 1/\sqrt{a}\right)}{\sqrt{\theta_l} - \sqrt{\theta_u}/\sqrt{a}} > 0, \quad K_2 := \sqrt{\theta_u} K_1 + 2.$$

From Theorem 2, we know that if $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is a Gaussian random matrix with entries $a_{k,j} \sim \mathcal{N}(0, 1/m)$ and the sampling complexity $m \ge C(a+1)(k+2)\log(en/k)$, then with high probability the matrix $\boldsymbol{A}' := \begin{bmatrix}\boldsymbol{A} \ \boldsymbol{b}\end{bmatrix}$ satisfies strong RIP condition of order $(a+1)(k+1)$ with constants $\theta_l,\ \theta_u > 0$ under some mild conditions on $\boldsymbol{b}$. Here, the constants $\theta_l,\ \theta_u$ are independent with a. Therefore, taking the constant $a > \theta_u/\theta_l$, the conclusion of Theorem 3 holds with high probability.

In the absence of noise, i.e., $\boldsymbol{w} = 0,\ \epsilon = 0$, Theorem 3 shows that if $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \sim 1/\sqrt{m} \cdot \mathcal{N}(0, I_n)$ are real Gaussian random vectors and $m \ge O(k\log(en/k))$, then all the k-sparse signals $\boldsymbol{x}_0 \in \mathbb{R}^n$ could be reconstructed exactly by solving the program (4) under some mild conditions on $\boldsymbol{b}$. We state it as the following corollary:

Corollary 1 *Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ be a Gaussian random matrix with entries $a_{jk} \sim \mathcal{N}(0, 1/m)$, and $\boldsymbol{b} \in \mathbb{R}^m$ be a vector satisfying $\alpha \le \|\boldsymbol{b}_I\|_2 \le \beta$ for all $I \subseteq [m]$ with $|I| \ge m/2$, where $\alpha \le \beta$ are two positive universal constants. If $m \ge Ck\log(en/k)$, then with probability at least $1-4\exp(-cm)$ it holds: for any k-sparse signal $\boldsymbol{x}_0 \in \mathbb{R}^n$, the ℓ_1 minimization*

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\operatorname{argmin}} \|\boldsymbol{x}\|_1 \quad \text{s.t.} \quad |\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}| = \boldsymbol{y}$$

with $\boldsymbol{y} = |\boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{b}|$ has a unique solution $\boldsymbol{x}_0$. Here $C,\ c > 0$ are constants depending only on α and β.

1.3.2 Complex Case

We next turn to consider the estimation performance of (4) for the complex-valued signals. Let $\mathbb{H}^{n\times n}$ be the set of Hermitian matrix in $\mathbb{C}^{n\times n}$ and $\|\boldsymbol{H}\|_{0,2}$ denotes the number of non-zero rows in $\boldsymbol{H}$. Given $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in \mathbb{C}^n$ and $b_1, \ldots, b_m \in \mathbb{C}$, we define a linear map $\mathcal{A}' : \boldsymbol{H}' \in \mathbb{H}^{(n+1)\times(n+1)} \to \mathbb{R}^m$ as follows:

$$\mathcal{A}'(\boldsymbol{H}') = (\boldsymbol{a}_1'^{*}\boldsymbol{H}'\boldsymbol{a}_1', \ldots, \boldsymbol{a}_m'^{*}\boldsymbol{H}'\boldsymbol{a}_m'), \tag{12}$$

where $\boldsymbol{a}_j' := \begin{pmatrix}\boldsymbol{a}_j \\ b_j\end{pmatrix} \in \mathbb{C}^{n+1}$.

Definition 2 We say the linear map $\mathcal{A}'$ defined in (12) satisfies the restricted isometry property of order (r, k) with constants $c,\ C > 0$ if the following holds

$$c\|\boldsymbol{H}'\|_F \le \frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}')\|_1 \le C\|\boldsymbol{H}'\|_F \tag{13}$$

for all $\boldsymbol{H}' := \begin{bmatrix}\boldsymbol{H} & \boldsymbol{h} \\ \boldsymbol{h}^* & 0\end{bmatrix} \in \mathbb{H}^{(n+1)\times(n+1)}$ with $\mathrm{rank}(\boldsymbol{H}) \le r$, $\|\boldsymbol{H}\|_{0,2} \le k$ and $\|\boldsymbol{h}\|_0 \le k$.

The following theorem shows that the linear map $\mathcal{A}'$ satisfies the restricted isometry property over low-rank and sparse matrices, provided $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in \mathbb{C}^n$ are i.i.d. complex Gaussian random vectors and $\boldsymbol{b} := (b_1, \ldots, b_m)^{\mathrm{T}} \in \mathbb{C}^m$ satisfies some mild conditions.

Theorem 4 *Suppose* $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \sim 1/\sqrt{2}\cdot\mathcal{N}(0, I_n)+i/\sqrt{2}\cdot\mathcal{N}(0, I_n)$ *are i.i.d. complex Gaussian random vectors and* $\boldsymbol{b} \in \mathbb{C}^m$ *is a independent sub-gaussian random vector (it also may be deterministic) with sub-gaussian norm* $\|\boldsymbol{b}\|_{\psi_2} \le C$ *and* $\mathbb{E}\|\boldsymbol{b}\|_1 \ge c_1 m$, $\mathbb{E}\|\boldsymbol{b}\|_2 \le c_2\sqrt{m}$, *where* $C > 0$, $c_2 \ge c_1 > 0$ *are universal constants. If* $m \ge C'k\log(en/k)$, *then with probability at least* $1 - 5\exp(-c'm)$, *the linear map* $\mathcal{A}'$ *defined in (12) obeys*

$$\frac{\theta^-}{12}\|\boldsymbol{H}'\|_F \le \frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}')\|_1 \le 3\theta^+\|\boldsymbol{H}'\|_F$$

for all $\boldsymbol{H}' := \begin{bmatrix} \boldsymbol{H} & \boldsymbol{h} \\ \boldsymbol{h}^* & 0 \end{bmatrix} \in \mathbb{H}^{(n+1)\times(n+1)}$ *with* $\operatorname{rank}(\boldsymbol{H}) \le 2$, $\|\boldsymbol{H}\|_{0,2} \le k$ *and* $\|\boldsymbol{h}\|_0 \le k$. *Here,* $\theta^- := \min(1, c_1/\sqrt{2})$, $\theta^+ := \max(\sqrt{6}, c_2)$, *and* $C', c' > 0$ *are constants depending only on* c_1, c_2.

With abuse of notation, we denote $\mathcal{A}'(\boldsymbol{x}') := \mathcal{A}'(\boldsymbol{x}'\boldsymbol{x}'^*)$ for any vector $\boldsymbol{x}' \in \mathbb{C}^{n+1}$. Then we have

Theorem 5 *Assume that the linear map* $\mathcal{A}'(\cdot)$ *satisfies the RIP condition (13) of order* $(2, 2ak)$ *with constants* $c, C > 0$. *For any* k*-sparse signal* $\boldsymbol{x}_0 \in \mathbb{C}^n$, *if*

$$c - C\left(\frac{4}{\sqrt{a}} + \frac{1}{a}\right) > 0,$$

then the solution $\widehat{\boldsymbol{x}} \in \mathbb{C}^n$ *to*

$$\operatorname*{argmin}_{\boldsymbol{x}\in\mathbb{C}^n} \|\boldsymbol{x}\|_1 \quad \text{s.t.} \quad \|\mathcal{A}'(\boldsymbol{x}') - \tilde{\boldsymbol{y}}\| \le \epsilon \quad \textit{and} \quad \boldsymbol{x}' = (\boldsymbol{x}^{\mathrm{T}}, 1)^{\mathrm{T}}$$

with $\tilde{\boldsymbol{y}} = \mathcal{A}'(\boldsymbol{x}_0') + \boldsymbol{w}$, $\|\boldsymbol{w}\| \le \epsilon$ *and* $\boldsymbol{x}_0' = (\boldsymbol{x}_0^{\mathrm{T}}, 1)^{\mathrm{T}}$ *obeys*

$$\min_{\theta\in\mathbb{R}}\left(\|\widehat{\boldsymbol{x}} - e^{i\theta}\boldsymbol{x}_0\|_2 + \left|1 - e^{i\theta}\right|\right) \le \frac{C_0\epsilon}{(\|\boldsymbol{x}_0\| + 1)\sqrt{m}},$$

where

$$C_0 := 2\sqrt{2}\cdot\frac{\frac{1}{a} + \frac{4}{\sqrt{a}} + 1}{c - C\left(\frac{4}{\sqrt{a}} + \frac{1}{a}\right)}.$$

Based on Theorem 4, if $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in \mathbb{C}^n$ are i.i.d. complex Gaussian random vectors and $m \ge C'ak\log(en/ak)$, then with high probability the linear map $\mathcal{A}'$ defined in (12) satisfies RIP condition of order $(2, 2ak)$ with constants $c = \theta^-/12$ and $C = 3\theta^+$ under some mild conditions on $\boldsymbol{b}$. For the noiseless case where $\boldsymbol{w} = 0$, $\epsilon = 0$, taking the constant $a > (8C/c)^2$ and combining with Theorem 5, we can obtain the following result.

Corollary 2 *Suppose $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \sim 1/\sqrt{2}\cdot\mathcal{N}(0, I_n)+i/\sqrt{2}\cdot\mathcal{N}(0, I_n)$ are i.i.d. complex Gaussian random vectors and $\boldsymbol{b} \in \mathbb{C}^m$ is a independent sub-gaussian random vector (it also may be deterministic) with sub-gaussian norm $\|\boldsymbol{b}\|_{\psi_2} \le C$ and $\mathbb{E}\|\boldsymbol{b}\|_1 \ge c_1 m$, $\mathbb{E}\|\boldsymbol{b}\|_2 \le c_2\sqrt{m}$, where $C > 0$, $c_2 \ge c_1 > 0$ are universal constants. If $m \ge C'' k \log(en/k)$, then with probability at least $1 - 5\exp(-c'' m)$, then the solution to*

$$\underset{\boldsymbol{x}\in\mathbb{C}^n}{\operatorname{argmin}} \quad \|\boldsymbol{x}\|_1 \quad s.t. \quad |\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}| = |\boldsymbol{A}\boldsymbol{x}_0+\boldsymbol{b}|$$

is $\boldsymbol{x}_0$ exactly. Here, $C'', c'' > 0$ are constants depending only on c_1, c_2.

Remark 1 We give an upper bound for $\min_{\theta\in\mathbb{R}} \left(\|\widehat{\boldsymbol{x}} - e^{i\theta}\boldsymbol{x}_0\|_2 + \left|1 - e^{i\theta}\right|\right)$ in Theorem 5. However, since the affine phase retrieval can recover a signal exactly (not just up to a global phase), one may wonder: is there a stable recovery bound for $\|\widehat{\boldsymbol{x}} - \boldsymbol{x}_0\|_2$? We believe that the answer is no, especially for the case where the noise vector $\|\boldsymbol{w}\|_2 \gtrsim \sqrt{m}$. We defer the proof of it for the future work.

1.4 Notations

Throughout the paper, we denote $\boldsymbol{x} \sim \mathcal{N}(0, I_n)$ if $\boldsymbol{x} \in \mathbb{R}^n$ is a standard Gaussian random vector. A vector $\boldsymbol{x}$ is k-sparse if there are at most k nonzero entries of $\boldsymbol{x}$. For simplicity, we denote $[m] := \{1, \ldots, m\}$. For any subset $I \subseteq [m]$, let $\boldsymbol{A}_I = \left[\boldsymbol{a}_j : j \in I\right]^*$ be the submatrix whose rows are generated by $\boldsymbol{A} = \left[\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\right]^*$. Denote $\sigma_k(\boldsymbol{x}_0)_p := \min_{|\mathrm{supp}(\boldsymbol{x})|\le k} \|\boldsymbol{x} - \boldsymbol{x}_0\|_p$ as the best k-term approximation error of $\boldsymbol{x}_0$ with respect to ℓ_p norm. For a complex number b, we use $b_{\Re}$ and $b_{\Im}$ to denote the real and imaginary part of b, respectively. For any $A, B \in \mathbb{R}$, we use $A \lesssim B$ to denote $A \le C_0 B$ where $C_0 \in \mathbb{R}_+$ is an absolute constant. The notion $\gtrsim$ can be defined similarly. Throughout this paper, c, C and the subscript (superscript) forms of them denote constants whose values vary with the context.

2 Proof of Theorem 2 and Theorem 3

In this section, we consider the estimation performance of the ℓ_1-minimization program (4) for the real-valued signals. Before proceeding, we need the following lemma which shows that if $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is a real Gaussian random matrix with entries $a_{k,j} \sim \mathcal{N}(0, 1/m)$, then $\boldsymbol{A}$ satisfies the strong RIP with high probability.

Lemma 1 (Theorem 2.1 in *[34]) Suppose that $t > 1$ and that $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is a Gaussian random matrix with entries $a_{k,j} \sim \mathcal{N}(0, 1/m)$. Let $m = O(tk\log(en/k))$ where $k \in [1, d] \cap \mathbb{Z}$ and $t \ge 1$ is a constant. Then there exist constants θ_l, θ_u with $0 < \theta_l < \theta_u < 2$, independent with t, such that $\boldsymbol{A}$ satisfies SRIP of order $t \cdot k$ and constants θ_l, θ_u with probability at least $1 - \exp(-cm)$, where $c > 0$ is a universal constant.*

2.1 Proof of Theorem 2

Proof From the definition, it suffices to show there exist constants $\theta_l', \theta_u' > 0$ such that the following inequality

$$\theta_l' \|\boldsymbol{x}'\|^2 \le \min_{I \subset [m], |I| \ge m/2} \|\boldsymbol{A}_I' \boldsymbol{x}'\|^2 \le \max_{I \subset [m], |I| \ge m/2} \|\boldsymbol{A}_I' \boldsymbol{x}'\|^2 \le \theta_u' \|\boldsymbol{x}'\|^2 \tag{14}$$

holds for all $(tk+1)$-sparse signals $\boldsymbol{x}' \in \mathbb{R}^{n+1}$. To this end, we denote $\boldsymbol{x}' = (\boldsymbol{x}^{\mathrm{T}}, z)^{\mathrm{T}}$, where $\boldsymbol{x} \in \mathbb{R}^n$ and $z \in \mathbb{R}$. We first consider the case where $z = 0$. From Lemma 1, we know that if $m \gtrsim t(k+1)\log(en/(k+1))$ and $t > 1$, then there exist two positive constants $\theta_l, \theta_u \in (0, 2)$ such that

$$\theta_l \|\boldsymbol{x}\|_2^2 \le \min_{I \subseteq [m], |I| \ge m/2} \|\boldsymbol{A}_I \boldsymbol{x}\|_2^2 \le \max_{I \subseteq [m], |I| \ge m/2} \|\boldsymbol{A}_I \boldsymbol{x}\|_2^2 \le \theta_u \|\boldsymbol{x}\|_2^2 \tag{15}$$

holds for all $(tk+1)$-sparse vector $\boldsymbol{x} \in \mathbb{R}^n$ with probability at least $1 - \exp(-cm)$. Here, $c > 0$ is a universal constant. Note that $\boldsymbol{A}'\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$. We immediately obtain (14) for the case where $z = 0$.

Next, we turn to the case where $z \neq 0$. A simple calculation shows that

$$\|\boldsymbol{A}_I' \boldsymbol{x}'\|_2^2 = \|\boldsymbol{A}_I \boldsymbol{x} + z\boldsymbol{b}_I\|_2^2 = \|\boldsymbol{A}_I \boldsymbol{x}\|_2^2 + 2z\langle \boldsymbol{A}_I \boldsymbol{x}, \boldsymbol{b}_I\rangle + z^2 \|\boldsymbol{b}_I\|_2^2 \tag{16}$$

for any $I \subseteq [m]$. Denote $\boldsymbol{A} = \left[\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\right]^{\mathrm{T}}$. Note that $\sqrt{m}\boldsymbol{a}_j \sim \mathcal{N}(0, I_n)$. Taking $\zeta = \frac{\min(\theta_l, \alpha^2)}{200\beta}$ in Lemma 5, we obtain that there exists a constant $C > 0$ depending only on $\theta_l,\ \alpha,\ \beta$ such that when $m \ge Ct(k+1)\log(en/k)$, with probability at least $1 - 3\exp(-c_1 m)$, it holds

$$|\langle \boldsymbol{A}_I \boldsymbol{x}, \boldsymbol{b}_I\rangle| = |\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{b}_I\rangle| \le \frac{\min\{\theta_l, \alpha^2\}}{200\beta} \|\boldsymbol{x}\|_2 \|\boldsymbol{b}\|_2 \tag{17}$$

for all $(tk+1)$-sparse vectors $\boldsymbol{x}$ and all $I \subseteq [m]$. Here, we view $\boldsymbol{b}_I = \boldsymbol{b}\mathbb{I}_I \in \mathbb{R}^m$ ($\mathbb{I}_I(j) = 1$ if $j \in I$ and 0 if $j \notin I$), and $c_1 > 0$ is a constant depending only on $\theta_l,\ \alpha,\ \beta$. Note that the vector $\boldsymbol{b}$ satisfies

$$\alpha \le \|\boldsymbol{b}_I\|_2 \le \beta \tag{18}$$

for all $I \subseteq [m]$ with $|I| \ge m/2$. Putting (15), (17) and (18) into (16), we obtain that when $m \ge Ct(k+1)\log(en/k)$, with probability at least $1 - 4\exp(-cm)$, the following two inequalities

$$\|\boldsymbol{A}_I' \boldsymbol{x}'\|_2^2 \ge \theta_l \|\boldsymbol{x}\|_2^2 - 2\,|z|\,\frac{\min\{\theta_l, \alpha^2\}}{200\beta} \|\boldsymbol{x}\|_2 \beta + \alpha^2 z^2 \ge 0.99 \min\{\theta_l, \alpha^2\} \|\boldsymbol{x}'\|_2^2,$$

and

$$\|A_I'x'\|_2^2 \le \theta_u \|x\|_2^2 + 2\,|z|\,\frac{\min\{\theta_l, \alpha^2\}}{200\beta}\|x\|_2\beta + \beta^2 z^2 \le 1.01\max\{\theta_u, \beta^2\}\|x'\|_2^2$$

hold for all $(tk+1)$-sparse vector $x' \in \mathbb{R}^{n+1}$ and for all $I \subseteq [m]$ with $|I| \ge m/2$. Here, $c > 0$ is a constant depending only on $\theta_l,\ \alpha,\ \beta$. In other words, we have

$$\theta_l'\|x'\|_2^2 \le \min_{I\subseteq[m],|I|\ge m/2}\|A_I'x'\|_2^2 \le \max_{I\subseteq[m],|I|\ge m/2}\|A_I'x'\|_2^2 \le \theta_u'\|x'\|_2^2$$

for all $(tk+1)$-sparse vector x' with probability at least $1-4\exp(-cm)$. Here, $\theta_l' = 0.99\min\{\theta_l, \alpha^2\}$ and $\theta_u' = 1.01\max\{\theta_u, \beta^2\}$. Combining the above two cases and noting that $\theta_l, \theta_u > 0$ are universal constants, we complete the proof. □

2.2 Proof of Theorem 3

Proof Denote $A' = \left[A\ b\right]$, $\widehat{x}' = (\widehat{x}^{\mathrm{T}}, 1)^{\mathrm{T}}$ and $x_0' = (x_0^{\mathrm{T}}, 1)^{\mathrm{T}}$. Set

$$I := \{j : (\langle a_j, \widehat{x}\rangle + b_j)(\langle a_j, x_0\rangle + b_j) \ge 0\}.$$

We next divide the proof into the following two cases.

Case 1: $|I| \ge m/2$. Set $h = \widehat{x}' - x_0'$. For any $a > 1$, we decompose h into the sum of $h_{T_0}, h_{T_1}, \ldots$, where T_0 is an index set which consists the indices of the $k+1$ largest coordinates of x_0' in magnitude, T_1 is the index set corresponding to the $a(k+1)$ largest coordinates of $h_{T_0^c}$ in magnitude, T_2 is the index set corresponding to the $a(k+1)$ largest coordinates of $h_{(T_0\cup T_1)^c}$ in magnitude, and so on. For simplicity, we denote $T_{jl} := T_j \cup T_l$. To prove the theorem, we only need to give an upper bound for $\|h\|_2$. Observe that

$$\|h\|_2 \le \|h_{T_{01}}\|_2 + \|h - h_{T_{01}}\|_2. \tag{19}$$

We claim that the following holds:

$$\|h - h_{T_{01}}\|_2 \le \frac{1}{\sqrt{a}}\|h_{T_{01}}\|_2 + \frac{2\sigma_k(x_0)_1}{\sqrt{a(k+1)}} \tag{20}$$

and

$$\|h_{T_{01}}\|_2 \le \frac{2}{\sqrt{\theta_l} - \sqrt{\theta_u}/\sqrt{a}} \cdot \left(\epsilon + \frac{\sqrt{\theta_u}\sigma_k(x_0)_1}{\sqrt{a(k+1)}}\right). \tag{21}$$

Here, $C,\ c,\ \theta_l$ and θ_u are positive constants depending only on α and β. Putting (20) and (21) into (19), we obtain that

$$\|h\|_2 \le \frac{2\left(1+1/\sqrt{a}\right)}{\sqrt{\theta_l} - \sqrt{\theta_u}/\sqrt{a}}\epsilon + \left(\frac{2(1+1/\sqrt{a})\sqrt{\theta_u}}{\sqrt{\theta_l} - \sqrt{\theta_u}/\sqrt{a}} + 2\right)\frac{\sigma_k(x_0)_1}{\sqrt{a(k+1)}}.$$

It remains to prove the claim (20) and (21). Since $\widehat{\boldsymbol{x}}$ is the solution to ℓ_1 minimization program (4), we have

$$\begin{aligned}\|\boldsymbol{x}_0'\|_1 \geq \|\widehat{\boldsymbol{x}}'\|_1 = \|\boldsymbol{x}_0' + \boldsymbol{h}\|_1 &= \|(\boldsymbol{x}_0' + \boldsymbol{h})_{T_0}\|_1 + \|(\boldsymbol{x}_0' + \boldsymbol{h})_{T_0^c}\|_1 \\ &\geq \|\boldsymbol{x}_{0,T_0}'\|_1 - \|\boldsymbol{h}_{T_0}\|_1 + \|\boldsymbol{h}_{T_0^c}\|_1 - \|\boldsymbol{x}_{0,T_0^c}'\|_1.\end{aligned}$$

Therefore,

$$\|\boldsymbol{h}_{T_0^c}\|_1 \leq \|\boldsymbol{h}_{T_0}\|_1 + 2\|\boldsymbol{x}_{0,T_0^c}'\|_1. \tag{22}$$

From the definition of T_j, we obtain that, for all $j \geq 2$,

$$\|\boldsymbol{h}_{T_j}\|_2 \leq \sqrt{a(k+1)}\|\boldsymbol{h}_{T_j}\|_\infty = \frac{a(k+1)}{\sqrt{a(k+1)}}\|\boldsymbol{h}_{T_j}\|_\infty \leq \frac{\|\boldsymbol{h}_{T_{j-1}}\|_1}{\sqrt{a(k+1)}}.$$

It then gives

$$\|\boldsymbol{h}_{T_{01}^c}\|_2 \leq \sum_{j\geq 2}\|\boldsymbol{h}_{T_j}\|_2 \leq \frac{1}{\sqrt{a(k+1)}}\sum_{j\geq 2}\|\boldsymbol{h}_{T_{j-1}}\|_1 = \frac{1}{\sqrt{a(k+1)}}\|\boldsymbol{h}_{T_0^c}\|_1. \tag{23}$$

Putting (22) into (23), we obtain the conclusion of claim (20), namely,

$$\begin{aligned}\|\boldsymbol{h}_{T_{01}^c}\|_2 &\leq \frac{1}{\sqrt{a(k+1)}}\|\boldsymbol{h}_{T_0^c}\|_1 \leq \frac{\|\boldsymbol{h}_{T_0}\|_1 + 2\|\boldsymbol{x}_{0,T_0^c}'\|_1}{\sqrt{a(k+1)}} \\ &\leq \frac{1}{\sqrt{a}}\|\boldsymbol{h}_{T_0}\|_2 + \frac{2\sigma_{k+1}(\boldsymbol{x}_0')_1}{\sqrt{k}} \leq \frac{1}{\sqrt{a}}\|\boldsymbol{h}_{T_{01}}\|_2 + \frac{2\sigma_k(\boldsymbol{x}_0)_1}{\sqrt{a(k+1)}},\end{aligned} \tag{24}$$

where the third inequality follows the Cauchy-Schwarz inequality and the last inequality comes from the fact $\sigma_{k+1}(\boldsymbol{x}_0')_1 \leq \sigma_k(\boldsymbol{x}_0)_1$ by the definitions of $\widehat{\boldsymbol{x}}'$ and $\sigma_k(\cdot)_1$.

We next turn to prove the claim (21). Observe that

$$\|\boldsymbol{A}_I'\boldsymbol{h}\|_2 \geq \|\boldsymbol{A}_I'\boldsymbol{h}_{T_{01}}\|_2 - \|\boldsymbol{A}_I'\boldsymbol{h}_{T_{01}^c}\|_2. \tag{25}$$

For the left hand side of (25), by the definition of I, we have

$$\begin{aligned}\|\boldsymbol{A}_I'\boldsymbol{h}\|_2 &= \big\|\,|\boldsymbol{A}_I'\widehat{\boldsymbol{x}}'| - |\boldsymbol{A}_I'\boldsymbol{x}_0'|\,\big\|_2 \\ &\leq \big\|\,|\boldsymbol{A}'\widehat{\boldsymbol{x}}'| - |\boldsymbol{A}'\boldsymbol{x}_0'|\,\big\|_2 \\ &\leq \big\|\,|\boldsymbol{A}'\widehat{\boldsymbol{x}}'| - \boldsymbol{y}\big\|_2 + \big\|\,|\boldsymbol{A}'\boldsymbol{x}_0'| - \boldsymbol{y}\big\|_2 \\ &\leq 2\epsilon.\end{aligned} \tag{26}$$

For the first term of the right hand side of (25), since the matrix $\boldsymbol{A}'$ satisfies strong RIP of order $(a+1)(k+1)$ with constants $\theta_l, \theta_u > 0$, we immediately have

$$\|\boldsymbol{A}_I'\boldsymbol{h}_{T_{01}}\|_2 \geq \sqrt{\theta_l}\|\boldsymbol{h}_{T_{01}}\|_2. \tag{27}$$

To give an upper bound for the term $\|A_I' h_{T_{01}^c}\|_2$, note that $\|h_{T_{01}^c}\|_\infty \leq \|h_{T_1}\|_1/a(k+1)$. Let $\theta := \max\left(\|h_{T_1}\|_1/a(k+1), \|h_{T_{01}^c}\|_1/a(k+1)\right)$. Then by the Lemma 2, we could decompose the vector $h_{T_{01}^c}$ into the following form:

$$h_{T_{01}^c} = \sum_{j=1}^{N} \lambda_j u_j, \quad \text{with} \quad 0 \leq \lambda_j \leq 1, \quad \sum_{j=1}^{N} \lambda_j = 1,$$

where u_j are $a(k+1)$-sparse vectors satisfying

$$\|u_j\|_1 = \|h_{T_{01}^c}\|_1, \quad \|u_j\|_\infty \leq \theta.$$

Therefore, we have

$$\|u_j\|_2 \leq \sqrt{\theta \|h_{T_{01}^c}\|_1}.$$

We notice from (22) that

$$\|h_{T_{01}^c}\|_1 \leq \|h_{T_0^c}\|_1 \leq \|h_{T_0}\|_1 + 2\sigma_k(x_0)_1.$$

Thus, if $\theta = \|h_{T_1}\|_1/a(k+1)$, then we have

$$\begin{aligned} \|u_j\|_2 &\leq \sqrt{\frac{\|h_{T_1}\|_1 \|h_{T_{01}^c}\|_1}{a(k+1)}} \leq \sqrt{\frac{\|h_{T_0^c}\|_1 \|h_{T_{01}^c}\|_1}{a(k+1)}} \\ &\leq \frac{\|h_{T_0}\|_1 + 2\sigma_k(x_0)_1}{\sqrt{a(k+1)}} \leq \frac{\|h_{T_0}\|_2}{\sqrt{a}} + \frac{2\sigma_k(x_0)_1}{\sqrt{a(k+1)}}. \end{aligned}$$

If $\theta = \|h_{T_{01}^c}\|_1/a(k+1)$, then

$$\|u_j\|_2 \leq \frac{\|h_{T_{01}^c}\|_1}{\sqrt{a(k+1)}} \leq \frac{\|h_{T_0}\|_2}{\sqrt{a}} + \frac{2\sigma_k(x_0)_1}{\sqrt{a(k+1)}}.$$

Therefore, for the second term of the right hand side of (25), it follows from the definition of strong RIP that

$$\|A_I' h_{T_{01}^c}\|_2 = \|\sum_{j=1}^{N} \lambda_j A_I' u_j\|_2 \leq \sqrt{\theta_u} \sum_{j=1}^{N} \lambda_j \|u_j\|_2 \leq \sqrt{\theta_u} \left(\frac{\|h_{T_0}\|_2}{\sqrt{a}} + \frac{2\sigma_k(x_0)_1}{\sqrt{a(k+1)}} \right). \tag{28}$$

Putting (26), (27) and (28) into (25), we immediately obtain

$$2\epsilon \geq \sqrt{\theta_l} \|h_{T_{01}}\|_2 - \sqrt{\theta_u} \left(\frac{\|h_{T_{01}}\|_2}{\sqrt{a}} + \frac{2\sigma_k(x_0)_1}{\sqrt{a(k+1)}} \right),$$

which gives

$$\|\boldsymbol{h}_{T_{01}}\|_2 \leq \frac{2}{\sqrt{\theta_l} - \sqrt{\theta_u}/\sqrt{a}} \cdot \left(\epsilon + \frac{\sqrt{\theta_u}\sigma_k(\boldsymbol{x}_0)_1}{\sqrt{a(k+1)}}\right).$$

Case 2: $|I| < m/2$. For this case, denote $\boldsymbol{h}^+ = \widehat{\boldsymbol{x}}' + \boldsymbol{x}_0'$. Replacing $\boldsymbol{h}$ and the subset I in Case 1 by $\boldsymbol{h}^+$ and I^c respectively, and applying the same argument, we could obtain

$$\|\boldsymbol{h}_+\| \leq \frac{2\left(1+1/\sqrt{a}\right)}{\sqrt{\theta_l} - \sqrt{\theta_u}/\sqrt{a}}\epsilon + \left(\frac{2(1+1/\sqrt{a})\sqrt{\theta_u}}{\sqrt{\theta_l} - \sqrt{\theta_u}/\sqrt{a}} + 2\right)\frac{\sigma_k(\boldsymbol{x}_0)_1}{\sqrt{a(k+1)}}. \tag{29}$$

However, recall that $\widehat{\boldsymbol{x}}' = (\widehat{\boldsymbol{x}}^{\mathrm{T}}, 1)^{\mathrm{T}}$ and $\boldsymbol{x}_0' = (\boldsymbol{x}_0^{\mathrm{T}}, 1)^{\mathrm{T}}$. It means $\|\boldsymbol{h}_+\|_2 \geq 2$, which contradicts to (29) by the assumption of ϵ and $\sigma_k(\boldsymbol{x}_0)_1$, i.e., $K_1\epsilon + K_2\frac{\sigma_k(\boldsymbol{x}_0)_1}{\sqrt{a(k+1)}} < 2$. Therefore, Case 2 does not hold.

Combining the above two cases, we complete our proof. □

3 Proof of Theorems 4 and 5

3.1 Proof of Theorem 4

Proof Without loss of generality, we assume that $\|\boldsymbol{H}'\|_F = 1$. Observe that

$$\frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}')\|_1 = \frac{1}{m}\sum_{j=1}^m \left|\boldsymbol{a}_j^*\boldsymbol{H}\boldsymbol{a}_j + 2(b_j(\boldsymbol{a}_j^*\boldsymbol{h}))_{\Re}\right| := \frac{1}{m}\sum_{j=1}^m \xi_j.$$

For any fixed $\boldsymbol{H} \in \mathbb{H}^{n\times n}$ and $\boldsymbol{h} \in \mathbb{C}^n$, the terms ξ_j, $j = 1, \ldots, m$ are independent sub-exponential random variables with the maximal sub-exponential norm

$$K := \max_{1\leq j\leq m} C_1(\|\boldsymbol{H}\|_F + \|b_j\|_{\psi_2}\|\boldsymbol{h}\|) \leq C_2$$

for some universal constants $C_1,\ C_2 > 0$. Here, we use the fact $\max\left(\|\boldsymbol{H}\|_F, \|\boldsymbol{h}\|\right) \leq \|\boldsymbol{H}'\|_F = 1$. For any $0 < \epsilon \leq 1$, the Bernstein's inequality gives

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{j=1}^m\left(\xi_j - \mathbb{E}\xi_j\right)\right| \geq \epsilon\right) \leq 2\exp\left(-c\epsilon^2 m\right),$$

where $c > 0$ is a universal constant. According to Lemma 6, we obtain that

$$\frac{1}{3}\mathbb{E}\sqrt{\|\boldsymbol{H}\|_F^2 + |b_j|^2\|\boldsymbol{h}\|^2} \leq \mathbb{E}\xi_j \leq 2\mathbb{E}\sqrt{3\|\boldsymbol{H}\|_F^2 + |b_j|^2\|\boldsymbol{h}\|^2}.$$

This gives

$$\frac{1}{m}\sum_{j=1}^{m}\mathbb{E}\xi_j \le \frac{2}{m}\sum_{j=1}^{m}\mathbb{E}\left(\sqrt{3}\|\boldsymbol{H}\|_F + |b_j|\|\boldsymbol{h}\|\right) \le 2\sqrt{3}\|\boldsymbol{H}\|_F + 2c_2\|\boldsymbol{h}\| \le 2\theta^+,$$

where $\theta^+ := \max(\sqrt{6}, c_2)$. Here, we use the fact $\|\boldsymbol{H}'\|_F^2 = \|\boldsymbol{H}\|_F^2 + 2\|\boldsymbol{h}\|^2 = 1$, $\mathbb{E}\|\boldsymbol{b}\|_1 \le \sqrt{m}\mathbb{E}\|\boldsymbol{b}\| \le c_2 m$, and $\frac{a+b}{\sqrt{2}} \le \sqrt{a^2+b^2} \le a+b$ for any positive number $a, b \in \mathbb{R}$. Similarly, we could obtain

$$\frac{1}{m}\sum_{j=1}^{m}\mathbb{E}\xi_j \ge \frac{1}{3\sqrt{2}}\cdot\frac{1}{m}\sum_{j=1}^{m}\mathbb{E}\left(\|\boldsymbol{H}\|_F + |b_j|\|\boldsymbol{h}\|\right) \ge \frac{1}{3\sqrt{2}}\left(\|\boldsymbol{H}\|_F + c_1\|\boldsymbol{h}\|\right) \ge \frac{\theta^-}{6},$$

where $\theta^- := \min(1, c_1/\sqrt{2})$. Collecting the above estimators, we obtain that, with probability at least $1 - 2\exp(-c\epsilon^2 m)$, the following inequality

$$\frac{\theta^-}{6} - \epsilon \le \frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}')\|_1 \le 2\theta^+ + \epsilon \tag{30}$$

holds for a fixed $\boldsymbol{H}' \in \mathbb{H}^{(n+1)\times(n+1)}$. We next show that (30) holds for all $\boldsymbol{H}' \in \mathcal{X}$, where

$$\mathcal{X} := \left\{\boldsymbol{H}' := \begin{bmatrix}\boldsymbol{H} & \boldsymbol{h}\\ \boldsymbol{h}^* & 0\end{bmatrix} \in \mathbb{H}^{(n+1)\times(n+1)} : \|\boldsymbol{H}'\|_F = 1,\ \mathrm{rank}(\boldsymbol{H}) \le 2,\ \|\boldsymbol{H}\|_{0,2} \le k,\ \|\boldsymbol{h}\|_0 \le k\right\}.$$

To this end, we adopt a basic version of a δ-net argument. Assume that $\mathcal{N}_\delta$ is a δ-net of $\mathcal{X}$, i.e., for any $\boldsymbol{H}' = \begin{bmatrix}\boldsymbol{H} & \boldsymbol{h}\\ \boldsymbol{h}^* & 0\end{bmatrix} \in \mathcal{X}$ there exists a $\boldsymbol{H}_0' := \begin{bmatrix}\boldsymbol{H}_0 & \boldsymbol{h}_0\\ \boldsymbol{h}_0^* & 0\end{bmatrix} \in \mathcal{N}_\delta$ such that $\|\boldsymbol{H} - \boldsymbol{H}_0\|_F \le \delta$ and $\|\boldsymbol{h} - \boldsymbol{h}_0\| \le \delta$. Using the same idea of Lemma 2.1 in [36], we obtain that the covering number of $\mathcal{X}$ is

$$|\mathcal{N}_\delta| \le \left(\frac{9\sqrt{2}en}{\delta k}\right)^{4k+2}\cdot\binom{n}{k}\left(1 + \frac{2}{\delta}\right)^{2k} \le \exp\left(C_3 k \log(en/\delta k)\right),$$

where $C_3 > 0$ is a universal constant. Note that $\boldsymbol{h} - \boldsymbol{h}_0$ has at most $2k$ nonzero entries. We obtain that if $m \gtrsim k\log(en/k)$, then with probability at least $1 - 3\exp(-cm)$, it holds

$$\begin{aligned}\left|\frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}')\|_1 - \frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}_0')\|_1\right| &\le \frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}' - \boldsymbol{H}_0')\|_1\\ &\le \frac{1}{m}\|\mathcal{A}(\boldsymbol{H} - \boldsymbol{H}_0)\|_1 + \frac{2}{m}\sum_{j=1}^{m}|b_j||\boldsymbol{a}_j^*(\boldsymbol{h} - \boldsymbol{h}_0)|\\ &\le \frac{1}{m}\|\mathcal{A}(\boldsymbol{H} - \boldsymbol{H}_0)\|_1\end{aligned}$$

$$+ 2\sqrt{\frac{1}{m}\sum_{j=1}^{m}|b_j|^2}\sqrt{\frac{1}{m}\sum_{j=1}^{m}|\boldsymbol{a}_j^*(\boldsymbol{h}-\boldsymbol{h}_0)|^2}$$
$$\le 2.45\|\boldsymbol{H}-\boldsymbol{H}_0\|_F + 3(c_2+1)\|\boldsymbol{h}-\boldsymbol{h}_0\|$$
$$\le 3\,(c_2+2)\,\delta,$$

where the linear map $\mathcal{A}(\cdot)$ is defined as $\mathcal{A}(\boldsymbol{H}) := (\boldsymbol{a}_1^*\boldsymbol{H}\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m^*\boldsymbol{H}\boldsymbol{a}_m)$, and the fourth inequality follows from the combination of Lemma 3, the fact $\frac{1}{-}m\sum_{j=1}^{m}\boldsymbol{a}_j\boldsymbol{a}_j^* \le 3/2$ with probability at least $1-\exp(-cm)$, and

$$\frac{1}{m}\sum_{j=1}^{m}|b_j|^2 \le \frac{\mathbb{E}\|\boldsymbol{b}\|^2}{m} + 1 \le c_2 + 1$$

with probability at least $1-2\exp(-cm)$. Choosing $\epsilon := \frac{1}{48}$, $\delta := \frac{\theta^-}{48(c_2+2)}$, and taking the union bound, we obtain that the following inequality

$$\frac{\theta^-}{12} \le \frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}')\|_1 \le 3\theta^+ \quad \text{for all} \quad \boldsymbol{H}' \in \mathcal{X}$$

holds with probability at least

$$1 - 3\exp(-cm) - 2\exp\left(C_3 k\log(en/\delta k)\right)\cdot\exp(-c\epsilon^2 m) \ge 1 - 5\exp(-c'm),$$

provided $m \ge C'k\log(en/k)$, where C', $c' > 0$ are constants depending only on c_1 and c_2. □

3.2 Proof of Theorem 5

Proof The proof of this theorem is adapted from that of Theorem 1.3 in [36]. Note that the ℓ_1-minimization problem we consider is

$$\operatorname*{argmin}_{\boldsymbol{x}\in\mathbb{C}^n} \ \|\boldsymbol{x}\|_1 \quad \text{s.t.} \quad \|\mathcal{A}'(\boldsymbol{x}') - \boldsymbol{y}'\| \le \epsilon \quad \text{with} \quad \boldsymbol{x}' = \begin{pmatrix}\boldsymbol{x}\\ 1\end{pmatrix}. \tag{31}$$

Here, with some abuse of notation, we set

$$\mathcal{A}'(\boldsymbol{x}') := \mathcal{A}'(\boldsymbol{x}'\boldsymbol{x}') = \left(|\boldsymbol{a}_1'^*\boldsymbol{x}'|^2, \ldots, |\boldsymbol{a}_m'^*\boldsymbol{x}'|^2\right) \quad \text{with} \quad \boldsymbol{a}_j' := \begin{pmatrix}\boldsymbol{a}_j\\ b_j\end{pmatrix}, \quad j = 1, \ldots, m.$$

Let $\widehat{\boldsymbol{x}} \in \mathbb{C}^n$ be a solution to (31). Without loss of generality, we assume $\langle \widehat{\boldsymbol{x}}', \boldsymbol{x}_0' \rangle \ge 0$ (Otherwise, we can choose $e^{i\theta}\boldsymbol{x}_0'$ for an appropriate θ), where $\widehat{\boldsymbol{x}}' = \begin{pmatrix}\widehat{\boldsymbol{x}}\\ 1\end{pmatrix}$ and $\boldsymbol{x}_0' = \begin{pmatrix}\boldsymbol{x}_0\\ 1\end{pmatrix}$. Set

$$\hat{\boldsymbol{X}}' := \widehat{\boldsymbol{x}}'\widehat{\boldsymbol{x}}'^* = \begin{pmatrix} \widehat{\boldsymbol{x}}\widehat{\boldsymbol{x}}^* & \widehat{\boldsymbol{x}} \\ \widehat{\boldsymbol{x}}^* & 1 \end{pmatrix}$$

and

$$\boldsymbol{H}' := \widehat{\boldsymbol{x}}'\widehat{\boldsymbol{x}}'^* - \boldsymbol{x}_0'\boldsymbol{x}_0'^* = \begin{pmatrix} \widehat{\boldsymbol{x}}\widehat{\boldsymbol{x}}^* - \boldsymbol{x}_0\boldsymbol{x}_0^* & \widehat{\boldsymbol{x}} - \boldsymbol{x}_0 \\ \widehat{\boldsymbol{x}}^* - \boldsymbol{x}_0^* & 0 \end{pmatrix} := \begin{pmatrix} \boldsymbol{H} & \boldsymbol{h} \\ \boldsymbol{h}^* & 0 \end{pmatrix}.$$

Therefore, it suffices to give an upper bound for $\|\boldsymbol{H}'\|_F$. Denote $T_0 := \operatorname{supp}(\boldsymbol{x}_0)$ and $T_0' := T_0 \cup \{n+1\}$. Let T_1 be the index set corresponding to the indices of the ak-largest elements of $\widehat{\boldsymbol{x}}_{T_0^c}$ in magnitude, and T_2 contain the indices of the next ak largest elements, and so on. Set $T_{01} := T_0 \cup T_1$, $T_{01}' := T_0' \cup T_1$, $\bar{\boldsymbol{h}} := \boldsymbol{h}_{T_{01}}$, $\bar{\boldsymbol{H}} = \boldsymbol{H}_{T_{01},T_{01}}$, and $\bar{\boldsymbol{H}}' := \boldsymbol{H}'_{T_{01}',T_{01}'}$. Noting that

$$\|\boldsymbol{H}'\|_F \le \|\bar{\boldsymbol{H}}'\|_F + \|\boldsymbol{H}' - \bar{\boldsymbol{H}}'\|_F, \tag{32}$$

we next consider the terms $\|\bar{\boldsymbol{H}}'\|_F$ and $\|\boldsymbol{H}' - \bar{\boldsymbol{H}}'\|_F$. We claim that

$$\|\boldsymbol{H}' - \bar{\boldsymbol{H}}'\|_F \le \left(\frac{1}{a} + \frac{4}{\sqrt{a}}\right)\|\bar{\boldsymbol{H}}'\|_F \tag{33}$$

and

$$\|\bar{\boldsymbol{H}}'\|_F \le \frac{1}{c - C\left(\frac{4}{\sqrt{a}} + \frac{1}{a}\right)} \cdot \frac{2\epsilon}{\sqrt{m}}. \tag{34}$$

Combining (32), (33) and (34), we obtain that

$$\|\boldsymbol{H}'\|_F \le \frac{\frac{1}{a} + \frac{4}{\sqrt{a}} + 1}{c - C\left(\frac{4}{\sqrt{a}} + \frac{1}{a}\right)} \cdot \frac{2\epsilon}{\sqrt{m}}.$$

According to Lemma 4, we immediately have

$$\min_{\theta\in\mathbb{R}} \|\widehat{\boldsymbol{x}}' - e^{\mathrm{i}\theta}\boldsymbol{x}_0'\|_2 \le \frac{\sqrt{2}\|\boldsymbol{H}'\|}{\|\boldsymbol{x}_0\| + 1} \le \frac{\frac{1}{a} + \frac{4}{\sqrt{a}} + 1}{c - C\left(\frac{4}{\sqrt{a}} + \frac{1}{a}\right)} \cdot \frac{2\sqrt{2}\epsilon}{(\|\boldsymbol{x}_0\| + 1)\sqrt{m}}.$$

By the definition of $\widehat{\boldsymbol{x}}'$ and $\boldsymbol{x}_0'$, we arrive at the conclusion.

It remains to prove the claims (33) and (34). Note that

$$\|\boldsymbol{H}' - \bar{\boldsymbol{H}}'\|_F \le \sum_{i\ge 2, j\ge 2} \|\boldsymbol{H}_{T_i,T_j}\|_F + 2\sum_{j\ge 2} \|\boldsymbol{H}'_{T_0',T_j}\|_F + 2\sum_{j\ge 2} \|\boldsymbol{H}'_{T_1,T_j}\|_F. \tag{35}$$

We first give an upper bound for the term $\sum_{i\ge 2, j\ge 2} \|\boldsymbol{H}'_{T_i,T_j}\|_F$. Noting that $\boldsymbol{x}_0$ is a k-sparse vector and $\widehat{\boldsymbol{x}} \in \mathbb{C}^n$ is the solution to (31), we obtain that

$$\|\boldsymbol{x}_0\|_1 \ge \|\widehat{\boldsymbol{x}}\|_1 = \|\widehat{\boldsymbol{x}}_{T_0}\|_1 + \|\widehat{\boldsymbol{x}}_{T_0^c}\|_1,$$

which implies $\|\widehat{\boldsymbol{x}}_{T_0^c}\|_1 \leq \|\widehat{\boldsymbol{x}}_{T_0} - \boldsymbol{x}_0\|_1$. Moreover, by the definition of T_j, we know that for all $j \geq 2$, it holds $\|\widehat{\boldsymbol{x}}_{T_j}\|_2 \leq \frac{\|\widehat{\boldsymbol{x}}_{T_{j-1}}\|_1}{\sqrt{ak}}$. It then implies

$$\sum_{j\geq 2} \|\widehat{\boldsymbol{x}}_{T_j}\|_2 \leq \frac{1}{\sqrt{ak}} \sum_{j\geq 2} \|\widehat{\boldsymbol{x}}_{T_{j-1}}\|_1 \leq \frac{1}{\sqrt{ak}} \|\widehat{\boldsymbol{x}}_{T_0^c}\|_1 \leq \frac{1}{\sqrt{a}} \|\widehat{\boldsymbol{x}}_{T_0} - \boldsymbol{x}_0\|_2. \tag{36}$$

Therefore, the first term of (35) can be estimated as

$$\begin{aligned} \sum_{i\geq 2, j\geq 2} \|\boldsymbol{H}_{T_i,T_j}\|_F &= \sum_{i\geq 2, j\geq 2} \|\widehat{\boldsymbol{x}}_{T_i}\|_2 \|\widehat{\boldsymbol{x}}_{T_j}\|_2 = \left(\sum_{j\geq 2} \|\widehat{\boldsymbol{x}}_{T_j}\|_2\right)^2 \leq \frac{1}{ak} \|\widehat{\boldsymbol{x}}_{T_0^c}\|_1^2 \\ &= \frac{1}{ak} \|\boldsymbol{H}_{T_0^c,T_0^c}\|_1 \leq \frac{1}{ak} \|\boldsymbol{H}_{T_0,T_0}\|_1 \leq \frac{1}{a} \|\bar{\boldsymbol{H}}'\|_F, \end{aligned} \tag{37}$$

where the second inequality follows from

$$\|\boldsymbol{H} - \boldsymbol{H}_{T_0,T_0}\|_1 = \|\widehat{\boldsymbol{x}}\widehat{\boldsymbol{x}}^* - (\widehat{\boldsymbol{x}}\widehat{\boldsymbol{x}}^*)_{T_0,T_0}\|_1 \leq \|\boldsymbol{x}_0\boldsymbol{x}_0{}^*\|_1 - \|(\widehat{\boldsymbol{x}}\widehat{\boldsymbol{x}}^*)_{T_0,T_0}\|_1 \leq \|\boldsymbol{H}_{T_0,T_0}\|_1.$$

Here, the first inequality comes from $\|\widehat{\boldsymbol{x}}\|_1 \leq \|\boldsymbol{x}_0\|_1$.

For the second term and the third term of (35), we obtain that

$$\begin{aligned} \sum_{j\geq 2} \|\boldsymbol{H}'_{T'_0,T_j}\|_F + \sum_{j\geq 2} \|\boldsymbol{H}'_{T_1,T_j}\|_F &= \|\widehat{\boldsymbol{x}}'_{T'_0}\| \sum_{j\geq 2} \|\widehat{\boldsymbol{x}}'_{T_j}\| + \|\widehat{\boldsymbol{x}}'_{T_1}\| \sum_{j\geq 2} \|\widehat{\boldsymbol{x}}'_{T_j}\| \\ &\leq \frac{1}{\sqrt{a}} \|\widehat{\boldsymbol{x}}'_{T'_0} - \boldsymbol{x}'_0\|_2 \left(\|\widehat{\boldsymbol{x}}'_{T'_0}\|_2 + \|\widehat{\boldsymbol{x}}'_{T_1}\|_2\right) \\ &\leq \frac{\sqrt{2}}{\sqrt{a}} \|\widehat{\boldsymbol{x}}'_{T'_{01}} - \boldsymbol{x}'_0\|_2 \|\widehat{\boldsymbol{x}}'_{T'_{01}}\|_2 \\ &\leq \frac{2}{\sqrt{a}} \|\bar{\boldsymbol{H}}'\|_F, \end{aligned} \tag{38}$$

where the first inequality follows from (36) due to $\widehat{\boldsymbol{x}}'_{T_j} = \widehat{\boldsymbol{x}}_{T_j}$ for all $j \geq 1$, and the last inequality comes from Lemma 4. Putting (37) and (38) into (35), we obtain that

$$\|\boldsymbol{H}' - \bar{\boldsymbol{H}}'\|_F \leq \left(\frac{1}{a} + \frac{4}{\sqrt{a}}\right) \|\bar{\boldsymbol{H}}'\|_F.$$

This proves the claim (33).

Finally, we turn to prove the claim (34). Note that $\|\mathcal{A}'(\widehat{\boldsymbol{x}}') - \tilde{\boldsymbol{y}}\| \leq \epsilon$ and $\tilde{\boldsymbol{y}} := \mathcal{A}'(\boldsymbol{x}'_0) + \epsilon$, which implies

$$\|\mathcal{A}'(\boldsymbol{H}')\|_2 \leq \|\mathcal{A}'(\widehat{\boldsymbol{x}}') - \tilde{\boldsymbol{y}}\|_2 + \|\mathcal{A}'(\boldsymbol{x}'_0) - \tilde{\boldsymbol{y}}\|_2 \leq 2\epsilon.$$

Thus, we have

$$\frac{2\epsilon}{\sqrt{m}} \geq \frac{1}{\sqrt{m}}\|\mathcal{A}'(\boldsymbol{H}')\|_2 \geq \frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}')\|_1 \geq \frac{1}{m}\|\mathcal{A}'(\bar{\boldsymbol{H}}')\|_1 - \frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}' - \bar{\boldsymbol{H}}')\|_1. \tag{39}$$

Recall that $\bar{\boldsymbol{H}}' := \begin{pmatrix} \bar{\boldsymbol{H}} & \bar{\boldsymbol{h}} \\ \bar{\boldsymbol{h}}^* & 0 \end{pmatrix}$ with $\operatorname{rank}(\bar{\boldsymbol{H}}) \leq 2$, $\|\bar{\boldsymbol{H}}\|_{0,2} \leq (a+1)k$, and $\|\bar{\boldsymbol{h}}\|_0 \leq (a+1)k$. It then follows from the RIP of $\mathcal{A}'$ that

$$\|\mathcal{A}'(\bar{\boldsymbol{H}}')\|_1 \geq c\|\bar{\boldsymbol{H}}'\|_F. \tag{40}$$

To prove (34), it suffices to give an upper bound for the term $\frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}' - \bar{\boldsymbol{H}}')\|_1$. Observe that

$$\boldsymbol{H}' - \bar{\boldsymbol{H}}' = (\boldsymbol{H}'_{T_0', T_{01}'^c} + \boldsymbol{H}'_{T_{01}'^c, T_0'}) + (\boldsymbol{H}'_{T_1, T_{01}'^c} + \boldsymbol{H}'_{T_{01}'^c, T_1}) + \boldsymbol{H}'_{T_{01}'^c, T_{01}'^c}. \tag{41}$$

Since

$$\boldsymbol{H}'_{T_0', T_{01}'^c} + \boldsymbol{H}'_{T_{01}'^c, T_0'} = \sum_{j\geq 2}(\boldsymbol{H}'_{T_0', T_j} + \boldsymbol{H}'_{T_j, T_0'}) = \sum_{j \geq 2}\begin{pmatrix} \widehat{\boldsymbol{x}}_{T_0}\widehat{\boldsymbol{x}}_{T_j}^* + \widehat{\boldsymbol{x}}_{T_j}\widehat{\boldsymbol{x}}_{T_0}^* & \widehat{\boldsymbol{x}}_{T_j} \\ \widehat{\boldsymbol{x}}_{T_j}^* & 0 \end{pmatrix},$$

then the RIP of $\mathcal{A}'$ implies

$$\begin{aligned} \frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}'_{T_0', T_{01}'^c} + \boldsymbol{H}'_{T_{01}'^c, T_0'})\|_1 &\leq C\sum_{j\geq 2}\left(\|\widehat{\boldsymbol{x}}_{T_0}\widehat{\boldsymbol{x}}_{T_j}^* + \widehat{\boldsymbol{x}}_{T_j}\widehat{\boldsymbol{x}}_{T_0}^*\|_F + 2\|\widehat{\boldsymbol{x}}_{T_j}\|_2\right) \\ &\leq 2\sqrt{2}C\|\widehat{\boldsymbol{x}}'_{T_0'}\|_2\sum_{j\geq 2}\|\widehat{\boldsymbol{x}}_{T_j}\|_2 \\ &\leq \frac{2\sqrt{2}}{\sqrt{a}}C\|\widehat{\boldsymbol{x}}'_{T_0'}\|_2\|\widehat{\boldsymbol{x}}'_{T_{01}'} - \boldsymbol{x}_0'\|_2. \end{aligned} \tag{42}$$

Similarly, we could obtain

$$\frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}'_{T_1, T_{01}'^c} + \boldsymbol{H}'_{T_{01}'^c, T_1})\|_1 \leq \frac{2\sqrt{2}}{\sqrt{a}}C\|\widehat{\boldsymbol{x}}'_{T_1}\|_2\|\widehat{\boldsymbol{x}}'_{T_{01}'} - \boldsymbol{x}_0'\|_2. \tag{43}$$

Finally, observe that $\frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}'_{T_{01}'^c, T_{01}'^c})\|_1 = \frac{1}{m}\|\mathcal{A}(\boldsymbol{H}_{T_{01}^c, T_{01}^c})\|_1$. Using the same technique as [36, Eq. (3.16)], we could obtain

$$\frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}'_{T_{01}'^c, T_{01}'^c})\|_1 \leq \frac{C}{a}\|\bar{\boldsymbol{H}}'\|_F. \tag{44}$$

Putting (42), (43) and (44) into (41), we have

$$\frac{1}{m}\|\mathcal{A}'(\boldsymbol{H}'-\bar{\boldsymbol{H}}')\|_1 \le \frac{4}{\sqrt{a}}C\|\widehat{\boldsymbol{x}}'_{T'_{01}}\|_2\|\widehat{\boldsymbol{x}}'_{T'_{01}}-\boldsymbol{x}'_0\|_2+\frac{C}{a}\|\bar{\boldsymbol{H}}'\|_F \le C\left(\frac{4}{\sqrt{a}}+\frac{1}{a}\right)\|\bar{\boldsymbol{H}}'\|_F. \tag{45}$$

Combining (39), (40) and (45), we immediately obtain

$$\left(c - C\left(\frac{4}{\sqrt{a}}+\frac{1}{a}\right)\right)\|\bar{\boldsymbol{H}}'\|_F \le \frac{2\epsilon}{\sqrt{m}},$$

which implies

$$\|\bar{\boldsymbol{H}}'\|_F \le \frac{1}{c - C\left(\frac{4}{\sqrt{a}}+\frac{1}{a}\right)}\cdot\frac{2\epsilon}{\sqrt{m}}.$$

This completes the proof of claim (34). □

Acknowledgements Meng Huang was supported by NSFC grant (12201022) and the Fundamental Research Funds for the Central Universities (Grant No. YWF-22-T-204). Zhiqiang Xu was supported by the National Science Fund for Distinguished Young Scholars (12025108) and NSFC (12021001,12288201).

A Supporting Lemmas

The following lemma gives a way for how to decompose a vector $\boldsymbol{v}\in\mathbb{R}^n$ into the convex combination of several k-sparse vectors.

Lemma 2 *([8, 40]) Suppose that $\boldsymbol{v}\in\mathbb{R}^n$ satisfying $\|\boldsymbol{v}\|_\infty\le\theta$ and $\|\boldsymbol{v}\|_1\le k\theta$, where $\theta>0$ and $k\in\mathbb{Z}_+$. Then we have*

$$\boldsymbol{v}=\sum_{j=1}^{N}\lambda_j\boldsymbol{u}_j \quad \text{with} \quad 0\le\lambda_j\le 1,\quad \sum_{j=1}^{N}\lambda_j=1,$$

where $\boldsymbol{u}_j\in\mathbb{R}^n$ is k-sparse vectors and $\|\boldsymbol{u}_j\|_1\le\|\boldsymbol{v}\|_1,\ \|\boldsymbol{u}_j\|_\infty\le\theta$.

Lemma 3 *([36]) Let the linear map $\mathcal{A}(\cdot)$ be defined as*

$$\mathcal{A}(\boldsymbol{H}) := (\boldsymbol{a}_1^*\boldsymbol{H}\boldsymbol{a}_1,\ldots,\boldsymbol{a}_m^*\boldsymbol{H}\boldsymbol{a}_m),$$

where $\boldsymbol{a}_j\sim 1/\sqrt{2}\cdot\mathcal{N}(0,I_n)+i/\sqrt{2}\cdot\mathcal{N}(0,I_n)$, $j=1,\ldots,m$ are i.i.d. complex Gaussian random vectors. If $m\gtrsim k\log(en/k)$, then with probability at least $1-2\exp(-c_0m)$, $\mathcal{A}$ satisfies

$$0.12\|\boldsymbol{H}\|_F \le \frac{1}{m}\|\mathcal{A}(\boldsymbol{H})\|_1 \le 2.45\|\boldsymbol{H}\|_F$$

for all $H\in\mathbb{H}^{n\times n}$ with $\operatorname{rank}(\boldsymbol{H})\le 2$ and $\|\boldsymbol{H}\|_{0,2}\le k$. Here, $\|\boldsymbol{H}\|_{0,2}$ denotes the number of non-zero rows in $\boldsymbol{H}$.

Lemma 4 *([24, 36]) For any vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$ obeying $\langle \boldsymbol{u}, \boldsymbol{v} \rangle \geq 0$, we have*

$$\|\boldsymbol{u}\boldsymbol{u}^* - \boldsymbol{v}\boldsymbol{v}^*\|_F \geq \frac{1}{\sqrt{2}} \|\boldsymbol{u}\|_2 \|\boldsymbol{u} - \boldsymbol{v}\|_2.$$

Lemma 5 *Suppose that $\boldsymbol{a}_j \sim \mathcal{N}(0, I_n)$, $j = 1, \ldots, m$ are i.i.d. Gaussian random vectors and $\boldsymbol{b} \in \mathbb{R}^m$ is a nonzero vector. For any fixed $\zeta \in (0, 1)$, if $m \geq C\zeta^{-2}k(\log(en/k) + \log(1/\zeta))$, then with probability at least $1 - 3\exp(-c_0\zeta^2 m)$ it holds that*

$$\sum_{j=1}^{m} b_j (\boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{x}) \leq \zeta \sqrt{m} \|\boldsymbol{x}\|_2 \|\boldsymbol{b}\|_2$$

for all k-sparse vectors $\boldsymbol{x} \in \mathbb{R}^n$. Here, $c_0 > 0$ is a universal constant.

Proof Without loss of generality we assume $\|\boldsymbol{x}\|_2 = 1$. For any fixed $\boldsymbol{x}_0$, the terms $\boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{x}_0$ are independent, mean zero, sub-gaussian random variables with the maximal sub-gaussian norm being a positive universal constant. The Hoeffding's inequality implies

$$\mathbb{P}\left(\left|b_j (\boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{x}_0)\right| \geq t\right) \leq 2\exp\left(-\frac{c_1^2 t^2}{\|\boldsymbol{b}\|_2^2}\right).$$

Here, $c_1 > 0$ is a universal constant. Taking $t = \zeta \sqrt{m} \|\boldsymbol{b}\|_2 / 2$, we obtain that

$$\left|\sum_{j=1}^{m} (\boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{x}_0)\right| \leq \frac{\zeta}{2} \cdot \sqrt{m} \|\boldsymbol{b}\|_2 \tag{46}$$

holds with probability at least $1 - 2\exp(-c_1\zeta^2 m/4)$.

Next, we give a uniform bound to (46) for all k-sparse vectors $\boldsymbol{x}$. Denote

$$\mathcal{S}_{n,k} = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\|_2 = 1, \|\boldsymbol{x}\|_0 \leq k\}.$$

We assume that $\mathcal{N}$ is a δ-net of $\mathcal{S}_{n,k}$ such that for any $\boldsymbol{x} \in \mathcal{S}_{n,k}$, there exists a vector $\boldsymbol{x}_0 \in \mathcal{N}$ such that $\|\boldsymbol{x} - \boldsymbol{x}_0\|_2 \leq \delta$. The covering number $|\mathcal{N}| \leq \binom{n}{k} (1 + \frac{2}{\delta})^k$. Note that $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq 2k$. Therefore, when $m \gtrsim 2k$, with probability at least $1 - \exp(-c_2 m)$, it holds, Thus we have

$$\begin{aligned}
\left|\left|\sum_{j=1}^{m} b_j (\boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{x})\right| - \left|\sum_{j=1}^{m} b_j (\boldsymbol{a}_j^{\mathrm{T}} \boldsymbol{x}_0)\right|\right| &\leq \left|\sum_{j=1}^{m} b_j \boldsymbol{a}_j^{\mathrm{T}} (\boldsymbol{x} - \boldsymbol{x}_0)\right| \\
&\leq \|\boldsymbol{b}\|_2 \sqrt{\sum_{j=1}^{m} |\boldsymbol{a}_j^{\mathrm{T}} (\boldsymbol{x} - \boldsymbol{x}_0)|^2} \\
&\leq \|\boldsymbol{b}\|_2 \sqrt{\left\|\sum_{j=1}^{m} \boldsymbol{a}_j \boldsymbol{a}_j^{\mathrm{T}}\right\|_2 \cdot \|\boldsymbol{x} - \boldsymbol{x}_0\|_2}
\end{aligned}$$

$$\leq 2\|\boldsymbol{b}\|_2\sqrt{m}\cdot\delta,$$

where the second inequality follows from the Cauchy-Schwarz inequality and the last inequality comes from the fact $\|\sum_{j=1}^m \boldsymbol{a}_j\boldsymbol{a}_j^{\mathrm{T}}\|_2 \leq 4m$ with probability at least $1-\exp(-c_2 m)$, where $c_2>0$ is a universal constant. Choosing $\delta=\zeta/4$ and taking the union bound over $\mathcal{N}$, we obtain that

$$\left|\sum_{j=1}^m b_j(\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x}_0)\right| \leq \zeta\cdot\sqrt{m}\|\boldsymbol{b}\|_2$$

holds with probability at least

$$1-2\exp(-c_1\zeta^2 m/4)\cdot\binom{n}{k}\cdot(1+\frac{2}{\delta})^k-\exp(-c_2 m)\geq 1-3\exp(-c\zeta^2 m)$$

provided $m\geq C\zeta^{-2}k(\log(en/k)+\log(1/\zeta))$. Here, C and c are positive universal constants. This completes the proof. □

Lemma 6 *Suppose that $\boldsymbol{a}\in\mathbb{C}^n$ is a complex Gaussian random vector and $b\in\mathbb{C}$ is a complex number. For any Hermitian matrix $\boldsymbol{H}\in\mathbb{C}^{n\times n}$ with* $\operatorname{rank}(\boldsymbol{H})\leq 2$ *and any vector $\boldsymbol{h}\in\mathbb{C}^n$, we have*

$$\frac{1}{3}\sqrt{\|\boldsymbol{H}\|_F^2+b^2\|\boldsymbol{h}\|^2}\leq\mathbb{E}\left|\boldsymbol{a}^*\boldsymbol{H}\boldsymbol{a}+2(b(\boldsymbol{a}^*\boldsymbol{h}))_{\Re}\right|\leq 2\sqrt{3\|\boldsymbol{H}\|_F^2+b^2\|\boldsymbol{h}\|^2}.$$

Proof Since $\boldsymbol{H}\in\mathbb{C}^{n\times n}$ is a Hermitian matrix with $\operatorname{rank}(\boldsymbol{H})\leq 2$, we can decompose $\boldsymbol{H}$ into

$$\boldsymbol{H}=\lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^*+\lambda_2\boldsymbol{u}_2\boldsymbol{u}_2^*,$$

where $\lambda_1,\lambda_2\in\mathbb{R}$ are eigenvalues of $\boldsymbol{H}$ and $\boldsymbol{u}_1,\boldsymbol{u}_2\in\mathbb{C}^n$ are the corresponding eigenvectors with $\|\boldsymbol{u}_1\|_2=\|\boldsymbol{u}_2\|_2=1$, $\langle\boldsymbol{u}_1,\boldsymbol{u}_2\rangle=0$. For the vector $\boldsymbol{h}\in\mathbb{C}^n$, we can write it in the form of

$$\boldsymbol{h}=\sigma_1\boldsymbol{u}_1+\sigma_2\boldsymbol{u}_2+\sigma_3\boldsymbol{u}_3,$$

where $\sigma_1,\sigma_2,\sigma_3\in\mathbb{C}$, and $\boldsymbol{u}_3\in\mathbb{C}^n$ satisfying $\langle\boldsymbol{u}_3,\boldsymbol{u}_1\rangle=0$, $\langle\boldsymbol{u}_3,\boldsymbol{u}_2\rangle=0$ and $\|\boldsymbol{u}_3\|=1$. For simplicity, without loss of generality, we assume that b is a real number. Therefore, we have

$$\boldsymbol{a}^*H\boldsymbol{a}+2(b(\boldsymbol{a}^*\boldsymbol{h}))_{\Re}=\lambda_1\left|\boldsymbol{a}^*\boldsymbol{u}_1\right|^2+\lambda_2\left|\boldsymbol{a}^*\boldsymbol{u}_2\right|^2+2b\left(\sigma_1\boldsymbol{a}^*\boldsymbol{u}_1+\sigma_2\boldsymbol{a}^*\boldsymbol{u}_2+\sigma_3\boldsymbol{a}^*\boldsymbol{u}_3\right)_{\Re}.$$

Note that $\boldsymbol{a}\in\mathbb{C}^n$ is a complex Gaussian random vector and $\boldsymbol{u}_1,\ \boldsymbol{u}_2,\ \boldsymbol{u}_3$ are orthogonal vectors. Thus, we have

$$\mathbb{E}\left|\boldsymbol{a}^*H\boldsymbol{a}+2(b(\boldsymbol{a}^*\boldsymbol{h}))_{\Re}\right|=\mathbb{E}|\xi|,$$

with ξ being a random variable given by

$$\xi = \lambda_1 z_1^2 + \lambda_1 z_2^2 + \lambda_2 z_3^2 + \lambda_2 z_4^2 + 2b\left(\sigma_{1,\Re} z_1 - \sigma_{1,\Im} z_2 + \sigma_{2,\Re} z_3 - \sigma_{2,\Im} z_4 + \sigma_{3,\Re} z_5 - \sigma_{3,\Im} z_6\right).$$

Here, $z_1, z_2, z_3, z_4, z_5, z_6 \sim \mathcal{N}(0, 1/2)$ are independent. By Cauchy-Schwarz inequality, we have

$$\mathbb{E}|\xi| \leq \sqrt{\mathbb{E}\xi^2} \quad \text{and} \quad \mathbb{E}\xi^2 = \mathbb{E}(\xi^{\frac{2}{3}}\xi^{\frac{4}{3}}) \leq (\mathbb{E}\xi)^{\frac{2}{3}}(\mathbb{E}\xi_j^4)^{\frac{1}{3}}.$$

It immediately gives

$$\sqrt{\frac{(\mathbb{E}\xi^2)^3}{\mathbb{E}\xi^4}} \leq \mathbb{E}|\xi| \leq \sqrt{\mathbb{E}\xi^2} \tag{47}$$

Let $z_1 = \rho_1 \cos\theta$, $z_2 = \rho_1 \sin\theta$, $z_3 = \rho_2 \cos\phi$ and $z_4 = \rho_2 \sin\phi$, $z_5 = \rho_3 \cos\gamma$ and $z_6 = \rho_3 \sin\gamma$. Through some tedious calculations, we have

$$\begin{aligned}
\mathbb{E}\xi^2 = &\left(\frac{1}{2\pi}\right)^3 \int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi}\int_0^{\infty}\int_0^{\infty}\int_0^{\infty} \rho_1\rho_2\rho_3\Big(\lambda_1\rho_1^2 + \lambda_2\rho_2^2 + 2b(\sigma_{1,\Re}\rho_1\cos\theta \\
&- \sigma_{1,\Im}\rho_1\sin\theta + \sigma_{2,\Re}\rho_2\cos\phi - \sigma_{2,\Re}\rho_2\sin\phi + \sigma_{3,\Re}\rho_3\cos\gamma - \sigma_{3,\Im}\rho_3\sin\gamma)\Big)^2 \\
&\times e^{-\frac{\rho_1^2+\rho_2^2+\rho_3^2}{2}} \mathrm{d}\rho_1\mathrm{d}\rho_2\mathrm{d}\rho_3\mathrm{d}\theta\mathrm{d}\phi\mathrm{d}\gamma \\
= &\, 8(\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2) + 4b^2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) \\
\leq &\, 12\|\boldsymbol{H}\|_F^2 + 4b^2\|\boldsymbol{h}\|^2,
\end{aligned}$$

where the last inequality follows from the fact that $\lambda_1^2 + \lambda_2^2 = \|\boldsymbol{H}\|_F^2$ and $\sigma_1^2 + \sigma_2^2 + \sigma_3^2 = \|\boldsymbol{h}\|^2$. Similarly, we could obtain

$$\mathbb{E}\xi_j^2 \geq 4\|\boldsymbol{H}\|_F^2 + 4b^2\|\boldsymbol{h}\|^2 \tag{48}$$

and

$$\begin{aligned}
\mathbb{E}\xi^4 &= 48(8(\lambda_1^4 + \lambda_1^3\lambda_2 + \lambda_1^2\lambda_2^2 + \lambda_1\lambda_2^3 + \lambda_2^4) + b^4(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)^2 \\
&\quad + 4b^2(\lambda_1 + \lambda_2)^2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) + 8b^2(\lambda_1^2\sigma_1^2 + \lambda_2^2\sigma_2^2)) \\
&\leq 48(12\|\boldsymbol{H}\|_F^4 + b^4\|\boldsymbol{h}\|_2^4 + 16b^2\|\boldsymbol{H}\|_F^2\|\boldsymbol{h}\|_2^2) \\
&\leq 576\left(\|\boldsymbol{H}\|_F^2 + b^2\|\boldsymbol{h}\|^2\right)^2,
\end{aligned} \tag{49}$$

where the first inequality follows from the fact that

$$\lambda_1^4 + \lambda_1^3\lambda_2 + \lambda_1^2\lambda_2^2 + \lambda_1\lambda_2^3 + \lambda_2^4 \leq \lambda_1^4 + \lambda_1^2\lambda_2^2 + \lambda_2^4 + \frac{1}{2}\left(\lambda_1^2 + \lambda_2^2\right)^2 \leq \frac{2}{3}\|\boldsymbol{H}\|_F^4$$

and

$$\lambda_1^2\sigma_1^2 + \lambda_2^2\sigma_2^2 \le \left(\lambda_1^2 + \lambda_2^2\right)\left(\sigma_1^2 + \sigma_2^2 + \sigma_3^2\right) \le \|\boldsymbol{H}\|_F^2\|\boldsymbol{h}\|^2.$$

Putting (48) and (49) into (47), we obtain

$$\mathbb{E}|\xi| \ge \frac{1}{3}\sqrt{\|\boldsymbol{H}\|_F^2 + b^2\|\boldsymbol{h}\|^2}.$$

Therefore, we have

$$\frac{1}{3}\sqrt{\|\boldsymbol{H}\|_F^2 + b^2\|\boldsymbol{h}\|^2} \le \mathbb{E}|\xi| \le 2\sqrt{3\|\boldsymbol{H}\|_F^2 + b^2\|\boldsymbol{h}\|^2}.$$

This completes the proof. □

References

1. Balan, R., Casazza, P., Edidin, D.: On signal reconstruction without phase. Appl. Comput. Harmon. Anal. **20**(3), 345–356 (2006)
2. Barmherzig, D.A., Sun, J., Li, P.N., Lane, T.J., Candès, E.J.: Holographic phase retrieval and reference design. Inverse Probl. **35**(9), 094001 (2019)
3. Beinert, R., Plonka, G.: Ambiguities in one-dimensional discrete phase retrieval from Fourier magnitudes. J. Fourier Anal. Appl. **21**(6), 1169–1198 (2015)
4. Beinert, R., Plonka, G.: Enforcing uniqueness in one-dimensional phase retrieval by additional signal information in time domain. Appl. Comput. Harmon. Anal. **45**(3), 505–525 (2018)
5. Bendory, T., Beinert, R.,Eldar, Y. C.: Fourier phase retrieval: Uniqueness and algorithms. Compressed Sensing and its Applications, pp. 55–91 (2017)
6. Bandeira, A., Cahill, J., Mixon, D., Nelson, A.: Saving phase: injectivity and stability for phase retrieval. Appl. Comput. Harmon. Anal. **37**(1), 106–125 (2014)
7. Cai, J., Huang, M., Li, D., Wang, Y.: Solving phase retrieval with random initial guess is nearly as good as by spectral initialization. Appl. Comput. Harmon. Anal. **58**, 60–84 (2022)
8. Cai, T.T., Zhang, A.: Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. IEEE Trans. Inf. Theory **60**(1), 122–132 (2013)
9. Candès, E.J., Li, X., Soltanolkotabi, M.: Phase retrieval via Wirtinger flow: theory and algorithms. IEEE Trans. Inf. Theory **61**(4), 1985–2007 (2015)
10. Chen, Y., Candès, E.J.: Solving random quadratic systems of equations is nearly as easy as solving linear systems. Commun. Pure Appl. Math. **70**(5), 822–883 (2017)
11. Edidin, D.: The geometry of ambiguity in one-dimensional phase retrieval. SIAM J. Appl. Algebr. Geom. **3**(4), 644–660 (2019)
12. Eldar, Y.C., Mendelson, S.: Phase retrieval: stability and recovery guarantees. Appl. Comput. Harmon. Anal. **36**(3), 473–494 (2014)
13. Conca, A., Edidin, D., Hering, M., Vinzant, C.: An algebraic characterization of injectivity in phase retrieval. Appl. Comput. Harmon. Anal. **38**(2), 346–356 (2015)
14. Fienup, J.R.: Reconstruction of an object from the modulus of its Fourier transform. Opt. Lett. **3**(1), 27–29 (1978)
15. Fienup, J.R.: Phase retrieval algorithms: a comparison. Appl. Opt. **21**(15), 2758–2769 (1982)
16. Gabor, D.: A new microscopic principle. Nature **161**(4098), 777–778 (1948)
17. Gabor, D.: Microscopy by reconstructed wave-fronts. Proc. R. Soc. Lond. Ser. A **197**(1051), 454–487 (1949)
18. Gao, B., Wang, Y., Xu, Z.: Stable signal recovery from phaseless measurements. J. Fourier Anal. Appl. **22**(4), 787–808 (2016)

19. Gao, B., Sun, Q., Wang, Y., Xu, Z.: Phase retrieval from the magnitudes of affine linear measurements. Adv. Appl. Math. **93**, 121–141 (2018)
20. Guizar-Sicairos, M., Fienup, J.R.: Holography with extended reference by autocorrelation linear differential operation. Opt. Express **15**(26), 17592–17612 (2007)
21. Harrison, R.W.: Phase problem in crystallography. J. Opt. Soc. Am. A **10**(5) (1993)
22. Hauptman, H.A.: The phase problem of X-ray crystallography. Rep. Prog. Phys. **54**(11), 1427–1454 (1991)
23. Huang, K., Eldar, Y.C., Sidiropoulos, N.D.: Phase retrieval from 1D Fourier measurements: convexity, uniqueness, and algorithms. IEEE Trans. Signal Process. **64**(23), 6105–6117 (2016)
24. Huang, M., Xu, Z.: Performance bound of the intensity-based model for noisy phase retrieval. arXiv:2004.08764 (2020)
25. Huang, M., Xu, Z.: Strong convexity of affine phase retrieval. arXiv:2204.09412 (2022)
26. Latychevskaia, T.: Iterative phase retrieval for digital holography: tutorial. JOSA A **36**(12), 31–40 (2019)
27. Liebling, M., Blu, T., Cuche, E., Marquet, P., Depeursinge, C., Unser, M.: Local amplitude and phase retrieval method for digital holography applied to microscopy. In: European Conference on Biomedical Optics, Vol. 5143, pp. 210–214 (2003)
28. Millane, R.P.: Phase retrieval in crystallography and optics. J. Opt. Soc Am. A **7**(3), 394–411 (1990)
29. Netrapalli, P., Jain, P., Sanghavi, S.: Phase retrieval using alternating minimization. IEEE Trans. Signal Process. **63**(18), 4814–4826 (2015)
30. Rodriguez, J.A., Xu, R., Chen, C., Zou, Y., Miao, J.: Oversampling smoothness: an effective algorithm for phase retrieval of noisy diffraction intensities. J. Appl. Crystallogr. **46**(2), 312–318 (2013)
31. Sanz, J.L.C.: Mathematical considerations for the problem of Fourier transform phase retrieval from magnitude. SIAM J. Appl. Math. **45**(4), 651–664 (1985)
32. Shechtman, Y., Eldar, Y.C., Cohen, O., Chapman, H.N., Miao, J., Segev, M.: Phase retrieval with application to optical imaging: a contemporary overview. IEEE Signal Process. Mag. **32**(3), 87–109 (2015)
33. Sun, J., Qu, Q., Wright, J.: A geometric analysis of phase retrieval. Found. Comput. Math. **18**(5), 1131–1198 (2018)
34. Voroninski, V., Xu, Z.: A strong restricted isometry property, with an application to phaseless compressed sensing. Appl. Comput. Harmon. Anal. **40**(2), 386–395 (2016)
35. Walther, A.: The question of phase retrieval in optics. J. Mod. Opt. **10**(1), 41–49 (1963)
36. Xia, Y., Xu, Z.: The recovery of complex sparse signals from few phaseless measurements. Appl. Comput. Harmon. Anal. **50**, 1–15 (2021)
37. Wang, G., Giannakis, G.B., Eldar, Y.C.: Solving systems of random quadratic equations via truncated amplitude flow. IEEE Trans. Inf. Theory **64**(2), 773–794 (2018)
38. Wang, Y., Xu, Z.: Generalized phase retrieval?: measurement number, matrix recovery and beyond. Appl. Comput. Harmon. Anal. **47**(2), 423–446 (2019)
39. Wang, Y., Xu, Z.: Phase retrieval for sparse signals. Appl. Comput. Harmon. Anal. **37**(3), 531–544 (2014)
40. Xu, G., Xu, Z.: On the ℓ_1-norm invariant convex k-sparse decomposition of signals. J. Oper. Res. Soc. China **1**(4), 537–541 (2013)
41. Zhang, H., Zhou, Y., Liang, Y., Chi, Y.: A nonconvex approach for phase retrieval: reshaped Wirtinger flow and incremental algorithms. J. Mach. Learn. Res. **18**(1), 5164–5198 (2017)