Phaseless compressive sensing using partial support information

Zhiyong Zhou,* Jun Yu

Department of Mathematics and Mathematical Statistics, Umeå University,
Umeå, 901 87, Sweden

October 16, 2018

Abstract: We study the recovery conditions of weighted ℓ_1 minimization for real-valued signal reconstruction from phaseless compressive sensing measurements when partial support information is available. A strong restricted isometry property condition is provided to ensure the stable recovery. Moreover, we present the weighted null space property as the sufficient and necessary condition for the success of k-sparse phaseless recovery via weighted ℓ_1 minimization. Numerical experiments are conducted to illustrate our results.

Keywords: Phaseless compressive sensing; Partial support information; Strong restricted isometry property; Weighted null space property.

1 Introduction

Compressive sensing aims to recover an unknown signal from the underdetermined linear measurements (see [8, 9] for a comprehensive view). It is known as phase retrieval or phaseless compressive sensing when there is no phase information. The phaseless compressive sensing problem has recently attracted considerable research interests and many algorithms have been proposed to solve this problem. Existing literature include [2–4,7,12,14,16], to name a few. Specifically, the goal of phaseless compressive sensing is to recover $x \in \mathbb{R}^N$ up to a unimodular scaling constant from noisy magnitude measurements $y = |Ax| + e \in \mathbb{R}^m$ with the measurement matrix $A = (a_1, \cdots, a_m)^T \in \mathbb{R}^{m \times N}$, $|Ax| = (|\langle a_1, x \rangle|, \cdots, |\langle a_m, x \rangle|)^T$ and the noise term $e \in \mathbb{R}^m$. When x is sparse or compressible, the stable recovery can be guaranteed by solving the following ℓ_1 minimization problem

$$\min_{z \in \mathbb{R}^N} \|z\|_1 \text{ subject to } \||Az| - y\|_2 \leq \varepsilon, \tag{1}$$

provided that the measurement matrix A satisfies the strong restricted isometry property (SRIP) [11, 17]. In the noiseless case, the first sufficient and necessary condition was presented in [18] by proposing a new version of null space property for the phase retrieval problem.

In this paper, we generalize the existing theoretical framework for phaseless compressive sensing to incorporate partial support information, where we consider the case that an estimate of the support of the signal is available. We follow the similar notations and arguments in [10,20]. For an arbitrary signal $x \in \mathbb{R}^N$, let x^k be its best k-term approximation, so that x^k minimizes $\|x - f\|_1$ over all k-sparse vectors f. Let T_0 be the support of x^k, where $T_0 \subset \{1, \cdots, N\}$ and $|T_0| \leq k$. Let $\tilde{T}$, the support estimate, be a subset of $\{1, 2 \cdots, N\}$ with cardinality $|\tilde{T}| = \rho k$, where $\rho \geq 0$ and $|\tilde{T} \cap T_0| = \alpha \rho k$ with $0 \leq \alpha \leq 1$. Here the parameter ρ determines the ratio of the size of the estimated support to the size of the actual support of x^k

---

*Corresponding author, zhiyong.zhou@umu.se.

(or the support of x if x is k-sparse), while the parameter α determines the ratio of the number of indices in the support of x^k that are accurately estimated in $\tilde{T}$ to the size of $\tilde{T}$, i.e., $\alpha = \frac{|\tilde{T} \cap T_0|}{|\tilde{T}|}$. To incorporate prior support information $\tilde{T}$, we adopt the weighted ℓ_1 minimization

$$\min_{z \in \mathbb{R}^N} \sum_{i=1}^N \mathrm{w}_i |z_i|, \text{ subject to } \||Az| - y\|_2 \leq \varepsilon, \text{ where } \mathrm{w}_i = \begin{cases} \omega \in [0, 1] & i \in \tilde{T}, \\ 1 & i \in \tilde{T}^c. \end{cases} \tag{2}$$

We present the SRIP condition and weighted null space property condition to guarantee the success of the recovery via the weighted ℓ_1 minimization problem above.

The paper is organized as follows. In Section 2, we introduce the definition of SRIP and present the stable recovery condition with this tool. In Section 3, the sufficient and necessary weighted null space property condition for the real sparse noise free phase retrieval is given. In Section 4, some numerical experiments are presented to illustrate our theoretical results. Finally, Section 5 is devoted to the conclusion.

Throughout the paper, for any vector $x \in \mathbb{R}^N$, we denote the ℓ_p norm by $\|x\|_p = (\sum_{i=1}^p |x_i|^p)^{1/p}$ for $p > 0$ and the weighted ℓ_1 norm as $\|x\|_{1,\mathrm{w}} = \sum_{i=1}^N \mathrm{w}_i |x_i|$. For any matrix X, $\|X\|_1$ denotes the entry-wise ℓ_1 norm. For any set T, we denote its cardinality as $|T|$. The vector $x \in \mathbb{R}^N$ is called k-sparse if at most k of its entries are nonzero, i.e., if $\|x\|_0 = |\mathrm{supp}(x)| \leq k$, where $\mathrm{supp}(x)$ denotes the index set of the nonzero entries. We denote the index set $[N] := \{1, 2, \cdots, N\}$. For a matrix $A = (a_1, \cdots, a_m)^T \in \mathbb{R}^{m \times N}$ and an index set $I \subset [m]$, we denote A_I the sub-matrix of A where only rows with indices in I are kept, i.e., $A_I = (a_j, j \in I)^T$.

2 SRIP

To recover sparse signals via ℓ_1 minimization in the classical compressive sensing setting, [5] introduced the notion of restricted isometry property (RIP) and established a sufficient condition. We say a matrix A satisfies the RIP of order k if there exists a constant $\delta_k \in [0, 1)$ such that for all k-sparse vectors x we have

$$(1 - \delta_k)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k)\|x\|_2^2. \tag{3}$$

Cai and Zhang [1] proved that the RIP of order tk with $\delta_{tk} < \sqrt{\frac{t-1}{t}}$ where $t > 1$ can guarantee the exact recovery in the noiseless case and stable recovery in the noisy case via ℓ_1 minimization. This condition is sharp when $t \geq \frac{4}{3}$, see [1] for details. Very recently, Chen and Li [6] generalized this sharp RIP condition to the weighted ℓ_1 minimization problem when partial support information was incorporated. We first present the following useful lemma, which is an extension of the result in [6].

Lemma 1 Let $x \in \mathbb{R}^N, y = Ax + e \in \mathbb{R}^m$ with $\|e\|_2 \leq \zeta$, and $\eta \geq 0$. Suppose that A satisfies RIP of order tk with $\delta_{tk} < \sqrt{\frac{t-d}{t-d+\gamma^2}}$ for some $t > d$, where $\gamma = \omega + (1 - \omega)\sqrt{1 + \rho - 2\alpha\rho}$ and

$$d = \begin{cases} 1, & \omega = 1 \\ 1 - \alpha\rho + a, & 0 \leq \omega < 1 \end{cases} \tag{4}$$

with $a = \max\{\alpha, 1 - \alpha\}\rho$. Then for any

$$\hat{x} \in \{z \in \mathbb{R}^N : \|z\|_{1,\mathrm{w}} \leq \|x\|_{1,\mathrm{w}} + \eta, \|Az - y\|_2 \leq \varepsilon\},$$

we have

$$\|\hat{x} - x\|_2 \leq C_1(\zeta + \varepsilon) + C_2 \frac{2(\omega\|x_{T_0^c}\|_1 + (1 - \omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1)}{\sqrt{k}} + C_2 \frac{\eta}{\sqrt{k}}, \tag{5}$$

2

where

$$C_1 = \frac{\sqrt{2(t-d)(t-d+\gamma^2)(1+\delta_{tk})}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})},$$

$$C_2 = \frac{\sqrt{2}\delta_{tk}\gamma + \sqrt{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})\delta_{tk}}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})} + \frac{1}{\sqrt{d}}.$$

Remark 1 Note that if x^{ℓ_2} is the solution of the weighted ℓ_1 minimization problem:

$$\min_{z \in \mathbb{R}^N} \|z\|_{1,\mathrm{w}}, \text{ subject to } \|Az - y\|_2 \leq \varepsilon,$$

then $x^{\ell_2} \in \{z \in \mathbb{R}^N : \|z\|_{1,\mathrm{w}} \leq \|x\|_{1,\mathrm{w}} + \eta, \|Az - y\|_2 \leq \varepsilon\}$ with $\eta = 0$. Therefore, this lemma is an extension of Theorem 3.1 in [6] by letting $\zeta = \varepsilon$ and $\eta = 0$. The proof follows from almost the same procedure for the proof of Theorem 3.1 in Section 4 of [6] via replacing the $P = \frac{2(\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1)}{\sqrt{k}\gamma}$ with $P' = \frac{2(\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1) + \eta}{\sqrt{k}\gamma}$, and letting $\zeta = \varepsilon$. In order not to repeat, we leave out all the details. In addition, this result also generalizes the Lemma 2.1 in [11], which is the special case with the noise term $e = 0$, $\zeta = 0$ and $\omega = 1$. This lemma will play a crucial role in establishing the stable phaseless recovery result via weighted ℓ_1 minimization later on.

To address the phaseless compressive sensing problem (2), a stronger version of RIP is needed. Its definition is provided as follows.

Definition 1 *(SRIP [11, 17]) We say a matrix $A = (a_1, \cdots, a_m)^T \in \mathbb{R}^{m \times N}$ has the Strong Restricted Isometry Property (SRIP) of order k with bounds $\theta_-, \theta_+ \in (0, 2)$ if*

$$\theta_-\|x\|_2^2 \leq \min_{I \subseteq [m], |I| \geq m/2} \|A_I x\|_2^2 \leq \max_{I \subseteq [m], |I| \geq m/2} \|A_I x\|_2^2 \leq \theta_+\|x\|_2^2 \tag{6}$$

holds for all k-sparse vectors $x \in \mathbb{R}^N$, where $[m] = \{1, \cdots, m\}$. We say A has the Strong Lower Restricted Isometry Property of order k with bound θ_- if the lower bound in (6) holds. Similarly, we say A has the Strong Upper Restricted Isometry Property of order k with bound θ_+ if the upper bound in (6) holds.

Next, we present the conditions for the stable recovery via weighted ℓ_1 minimization by using SRIP.

Theorem 1 *Let $x \in \mathbb{R}^N, y = |Ax| + e \in \mathbb{R}^m$ with $\|e\|_2 \leq \zeta$. Adopt the notations in Lemma 1 and assume that $A \in \mathbb{R}^{m \times N}$ satisfies the SRIP of order tk with bounds $\theta_-, \theta_+ \in (0, 2)$ such that*

$$t \geq \max\left\{d + \frac{\gamma^2(1 - \theta_-)^2}{2\theta_- - \theta_-^2}, d + \frac{\gamma^2(1 - \theta_+)^2}{2\theta_+ - \theta_+^2}\right\}. \tag{7}$$

Then any solution $x^\sharp$ of (2) satisfies

$$\min\{\|x^\sharp - x\|_2, \|x^\sharp + x\|_2\} \leq C_1(\zeta + \varepsilon) + C_2\frac{2(\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1)}{\sqrt{k}}. \tag{8}$$

where C_1 and C_2 are constants defined in Lemma 1.

Remark 2 As it has been proved in [17] that Gaussian matrices with $m = O(tk\log(N/k))$ satisfy SRIP of order tk with high probability, thus the stable recovery result (8) can be achieved by using Gaussian measurement matrix with appropriate number of measurements m.

Remark 3 Note that when the weight $\omega = 1$, we have $\gamma = d = 1$. Then, by assuming $\zeta = \varepsilon = 0$ and x is exactly k-sparse, our theorem reduces to Theorem 2.2 in [17]. That is, if A satisfies the SRIP of order tk with bounds θ_-, θ_+ and $t \geq \max\{\frac{1}{2\theta_- - \theta_-^2}, \frac{1}{2\theta_+ - \theta_+^2}\}$, then for any k-sparse signal $x \in \mathbb{R}^N$ we have $\arg\min_{z \in \mathbb{R}^N}\{\|z\|_1 : |Az| = |Ax|\} = \{\pm x\}$. Similarly, if we let the noise term $e = 0$, $\zeta = 0$ and $\omega = 1$, this theorem goes to Theorem 3.1 in [11].

Remark 4 If $\alpha = \frac{1}{2}$, we have $\gamma = d = 1$. The sufficient condition (7) of Theorem 1 is identical to that of Theorem 2.2 in [17] and that of Theorem 3.1 in [11]. And the constants $C_1 = c_1 = \frac{\sqrt{2(1+\delta_{tk})}}{1-\sqrt{t/(t-1)}\delta_{tk}}, C_2 = c_2 = \frac{\sqrt{2}\delta_{tk} + \sqrt{(\sqrt{t(t-1)} - \delta_{tk}t)\delta_{tk}}}{\sqrt{t(t-1)} - \delta_{tk}t}$ (see Theorem 3.1 in [11]). In addition, if $0 \leq \omega < 1$ and $\alpha > \frac{1}{2}$, then $d = 1$ and $\gamma < 1$. The sufficient condition (7) in Theorem 1 is weaker than that of Theorem 2.2 in [17] and that of Theorem 3.1 in [11]. In this case, the constants $C_1 < c_1, C_2 < c_2$.

Set $t^\omega = \max\left\{d + \frac{\gamma^2(1-\theta_-)^2}{2\theta_- - \theta_-^2}, d + \frac{\gamma^2(1-\theta_+)^2}{2\theta_+ - \theta_+^2}\right\}$. We illustrate how the constants t^ω, C_1 and C_2 change with ω for different values of α in Figure 1. In all the plots, we set $\rho = 1$. In the plot of t^ω, we set $\theta_- = \frac{1}{2}$ and $\theta_+ = \frac{3}{2}$, then $t^\omega = d + \frac{\gamma^2}{3}$. In the plots of C_1 and C_2, we fix $t = 4$ and $\delta_{tk} = 0.3$. Note that if $\omega = 1$ or $\alpha = 0.5$, then $t^\omega \equiv 1 + \frac{1}{3} = \frac{4}{3}$, $C_1 \equiv c_1$ and $C_1 \equiv c_2$. And it shows that t^ω decreases as α increases, which means that the sufficient condition (7) becomes weaker as α increases. For each $\alpha > 0.5$, the sufficient condition becomes stronger (t^ω increases) as ω increases. For instance, if 90% of the support estimate is accurate ($\alpha = 0.9$) and $\omega = 0.6$, we have $t^\omega = 1.2022$, while $t^\omega = 1.3333$ for standard ℓ_1 minimization ($\omega = 1$). The opposite conclusion holds for the case $\alpha < 0.5$. In addition, as α increases, the constant C_1 decreases with $t = 4$ and $\delta_{tk} = 0.3$. Meanwhile, the constant C_2 with $\alpha \neq 0.5$ is smaller than that with $\alpha = 0.5$.

Proof of Theorem 1. For any solution $x^\sharp$ of (2), we have

$$\|x^\sharp\|_{1,w} \leq \|x\|_{1,w}$$

and

$$\||Ax^\sharp| - |Ax| - e\|_2 \leq \varepsilon.$$

If we divide the index set $\{1, 2, \cdots, m\}$ into two subsets

$$T = \{j : \text{sign}(\langle a_j, x^\sharp\rangle) = \text{sign}(\langle a_j, x\rangle)\} \text{ and } T^c = \{j : \text{sign}(\langle a_j, x^\sharp\rangle) = -\text{sign}(\langle a_j, x\rangle)\},$$

then it implies that

$$\|A_T x^\sharp - A_T x - e\|_2 + \|A_{T^c} x^\sharp + A_{T^c} x - e\|_2 \leq \varepsilon. \tag{9}$$

Here either $|T| \geq m/2$ or $|T^c| \geq m/2$. If $|T| \geq m/2$, we use the fact that

$$\|A_T x^\sharp - A_T x - e\|_2 \leq \varepsilon. \tag{10}$$

Then, we obtain

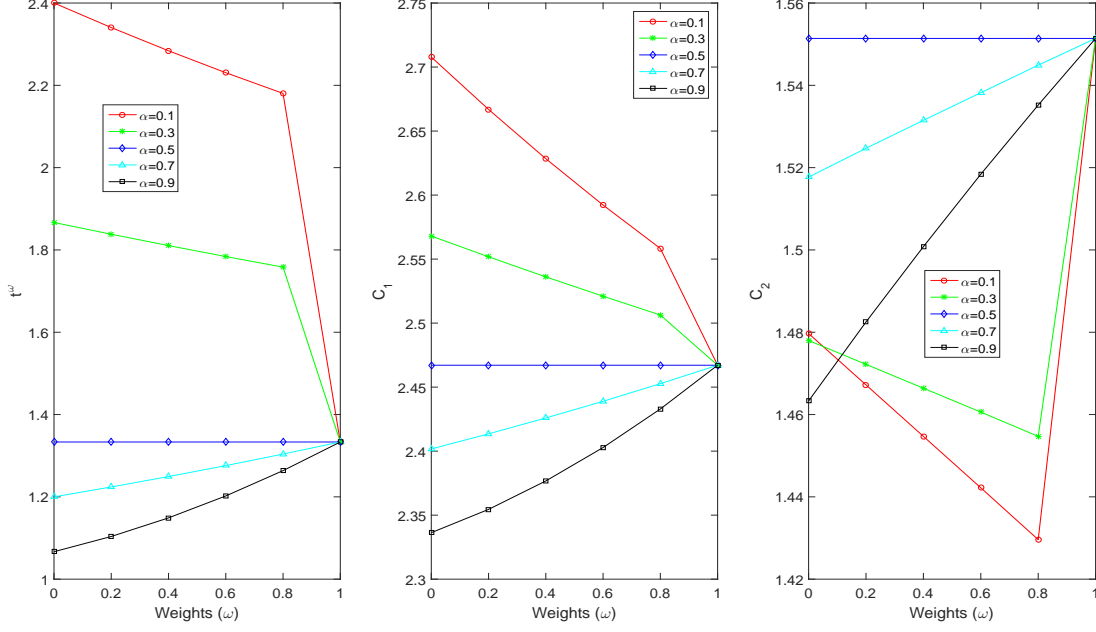$$x^\sharp \in \{z \in \mathbb{R}^N : \|z\|_{1,w} \leq \|x\|_{1,w}, \|A_T z - A_T x - e\|_2 \leq \varepsilon\}.$$

4

Figure 1: Comparison of the constants t^ω, C_1 and C_2 for various of α. In all the plots, we set $\rho = 1$. In the plot of t^ω, we set $\theta_- = \frac{1}{2}$ and $\theta_+ = \frac{3}{2}$. In the plots of C_1 and C_2, we fix $t = 4$ and $\delta_{tk} = 0.3$.

Since A satisfies SRIP of order tk with bounds θ_-, θ_+ and

$$t \geq \max \left\{ d + \frac{\gamma^2(1-\theta_-)^2}{2\theta_- - \theta_-^2}, d + \frac{\gamma^2(1-\theta_+)^2}{2\theta_+ - \theta_+^2} \right\} > d,$$

therefore, the definition of SRIP implies that A_T satisfies the RIP of order tk with

$$\delta_{tk} \leq \max\{1 - \theta_-, \theta_+ - 1\} \leq \sqrt{\frac{t-d}{t-d+\gamma^2}}. \tag{11}$$

Thus, by using Lemma 1 with $\eta = 0$, we have

$$\|x^\sharp - x\|_2 \leq C_1(\zeta + \varepsilon) + C_2 \frac{2(\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1)}{\sqrt{k}}.$$

Similarly, if $|T^c| \geq m/2$, we obtain the other corresponding result

$$\|x^\sharp + x\|_2 \leq C_1(\zeta + \varepsilon) + C_2 \frac{2(\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1)}{\sqrt{k}}.$$

The proof of Theorem 1 is now completed.

3 Weighted Null Space Property

In this section, we consider the noiseless weighted ℓ_1 minimization problem, i.e.,

$$\min_{z \in \mathbb{R}^N} \|z\|_{1,\mathrm{w}}, \quad \text{subject to } |Az| = |Ax|, \quad \text{where } \mathrm{w}_i = \begin{cases} \omega \in [0,1], & i \in \tilde{T} \\ 1, & i \in \tilde{T}^c \end{cases}. \tag{12}$$

5

We denote the kernel space of A by $\mathcal{N}(A) := \{h \in \mathbb{R}^N : Ah = 0\}$ and denote the k-sparse vector space $\Sigma_k^N := \{x \in \mathbb{R}^N : \|x\|_0 \leq k\}$.

Definition 2 *The matrix A satisfies the* w-*weighted null space property of order k if for any nonzero $h \in \mathcal{N}(A)$ and any $T \subset [N]$ with $|T| \leq k$ it holds that*

$$\|h_T\|_{1,\mathrm{w}} < \|h_{T^c}\|_{1,\mathrm{w}}, \tag{13}$$

where T^c is the complementary index set of T and h_T is the restriction of h to T.

Remark 5 Obviously, when the weight $\omega = 1$, the weighted null space property reduces to the classical null space property. And according to the specific setting of w_i, the expression (13) is equivalent to

$$\omega\|h_{T\cap\tilde{T}}\|_1 + \|h_{T\cap\tilde{T}^c}\|_1 < \omega\|h_{T^c\cap\tilde{T}}\|_1 + \|h_{T^c\cap\tilde{T}^c}\|_1 \Leftrightarrow \omega\|h_T\|_1 + (1-\omega)\|h_G\|_1 < \|h_{T^c}\|_1,$$

where $G = (T \cap \tilde{T}^c) \cup (T^c \cap \tilde{T})$ (see [13] for more arguments).

It is known that a signal $x \in \Sigma_k^N$ can be recovered via the weighted ℓ_1 minimization problem if and only if the measurement matrix A has the weighted null space property of order k. We state it as follows (see [19]):

Lemma 2 *Given $A \in \mathbb{R}^{m \times N}$, for every k-sparse vector $x \in \mathbb{R}^N$ it holds that*

$$\underset{z \in \mathbb{R}^N}{\arg\min} \ \{\|z\|_{1,\mathrm{w}} : Az = Ax\} = x$$

if and only if A satisfies the w-*weighted null space property of order k.*

Next, we extend Lemma 2 to the following theorem on phaseless compressive sensing for the real-valued signal reconstruction.

Theorem 2 *The following statements are equivalent:*
(a) For any k-sparse $x \in \mathbb{R}^N$, we have

$$\underset{z \in \mathbb{R}^N}{\arg\min}\{\|z\|_{1,\mathrm{w}} : |Az| = |Ax|\} = \{\pm x\}. \tag{14}$$

(b) For every $S \subseteq [m]$, it holds

$$\|u + v\|_{1,\mathrm{w}} < \|u - v\|_{1,\mathrm{w}} \tag{15}$$

for all nonzero $u \in \mathcal{N}(A_S)$ and $v \in \mathcal{N}(A_{S^c})$ satisfying $\|u + v\|_0 \leq k$.

Remark 6 If $\omega = 1$, then Theorem 2 reduces to Theorem 3.2 in [18]. Since $\mathrm{w}_i = \omega$ when $i \in \tilde{T}$, and $\mathrm{w}_i = 1$ otherwise, the expression (15) is equivalent to

$$\omega\|u + v\|_1 + (1-\omega)\|(u+v)_{\tilde{T}^c}\|_1 < \omega\|u - v\|_1 + (1-\omega)\|(u-v)_{\tilde{T}^c}\|_1.$$

Proof of Theorem 2. The proof follows from the proof of Theorem 3.2 in [18] with minor modifications. First we show $(a) \Rightarrow (b)$. Assume (b) is false, that is, there exist nonzero $u \in \mathcal{N}(A_S)$ and $v \in \mathcal{N}(A_{S^c})$ such that

$$\|u + v\|_{1,\mathrm{w}} \geq \|u - v\|_{1,\mathrm{w}}$$

and $u + v \in \Sigma_k^N$. Now set $x = u + v \in \Sigma_k^N$, obviously for $i = 1, \cdots, m$, we have

$$|\langle a_i, x \rangle| = |\langle a_i, u + v \rangle| = |\langle a_i, u - v \rangle|,$$

since either $\langle a_i, u \rangle = 0$ or $\langle a_i, v \rangle = 0$. In other words $|Ax| = |A(u-v)|$. Note that $u - v \neq -x$, for otherwise we would have $u = 0$, which is a contradiction. Then, it follows from (a) that we obtain

$$\|x\|_{1,\mathrm{w}} = \|u + v\|_{1,\mathrm{w}} < \|u - v\|_{1,\mathrm{w}},$$

This is a contradiction. Thus, (b) holds.

Next we prove $(b) \Rightarrow (a)$. Let $b = (b_1, \cdots, b_m)^T = |Ax|$ where $x \in \Sigma_k^N$. For a fixed $\sigma = (\sigma_1, \cdots, \sigma_m)^T \in \{-1, 1\}^m$, we set $b^\sigma = (\sigma_1 b_1, \cdots, \sigma_m b_m)^T$. We now consider the following weighted ℓ_1 minimization problem:

$$\min_{z \in \mathbb{R}^N} \|z\|_{1,\mathrm{w}} \text{ subject to } Az = b^\sigma. \tag{16}$$

Its solution is denoted as x^σ. Then, we claim that for any $\sigma \in \{1, -1\}^m$, if x^σ exists (it may not exist), we have

$$\|x^\sigma\|_{1,\mathrm{w}} \geq \|x\|_{1,\mathrm{w}}$$

and the equality holds if and only if $x^\sigma = \pm x$.

To prove the claim, we assume $\sigma^\star \in \{1, -1\}^m$ such that $b^{\sigma^\star} = Ax$. First note that the statement (b) implies the classical weighted null space property of order k. To see this, for any nonzero $h \in \mathcal{N}(A)$ and $T \subseteq [N]$ with $|T| \leq k$, we set $u = h$, $v = h_T - h_{T^c}$ and $S = [m]$. Then, we have $u \in \mathcal{N}(A_S)$ and $v \in \mathcal{N}(A_{S^c})$. Therefore, the statement (b) now implies

$$2\|h_T\|_{1,\mathrm{w}} = \|u + v\|_{1,\mathrm{w}} < \|u - v\|_{1,\mathrm{w}} = 2\|h_{T^c}\|_{1,\mathrm{w}}.$$

As a consequence, we have $x^{\sigma^\star} = x$ by Lemma 2. And, similarly we have $x^{-\sigma^\star} = -x$. Next, for any $\sigma \in \{-1, 1\}^m \neq \pm\sigma^\star$, if x^σ doesn't exist then we have nothing to prove. Assume it does exist, set $S_\star = \{i : \sigma_i = \sigma_i^\star\}$. Then

$$\langle a_i, x^\sigma \rangle = \begin{cases} \langle a_i, x \rangle & i \in S_\star, \\ -\langle a_i, x \rangle & i \in S_\star^c. \end{cases}$$

Set $u = x - x^\sigma$ and $v = x + x^\sigma$. Obviously, $u \in \mathcal{N}(A_{S_\star})$ and $v \in \mathcal{N}(A_{S_\star^c})$. Furthermore, $u + v = 2x \in \Sigma_k^N$. Then, by the statement (b), we have

$$2\|x\|_{1,\mathrm{w}} = \|u + v\|_{1,\mathrm{w}} < \|u - v\|_{1,\mathrm{w}} = 2\|x^\sigma\|_{1,\mathrm{w}}.$$

This proves (a) and the proof is completed.

4   Simulations

In this section, we present some simple numerical experiments to illustrate the benefits of using weighted ℓ_1 minimization to recover sparse and compressible signals when partial prior support information is available in the phaseless compressive sensing case. In order to facilitate the computation, we follow a non-standard noise model:

$$b = |Ax|^2 + e = \{a_i^T x x^T a_i\}_{1 \leq i \leq m} + e, \tag{17}$$

where $e \in \mathbb{R}^m$ is a noise term with $\|e\|_2 \leq \varepsilon$. Then the weighted ℓ_1 minimization goes to

$$\min_{z \in \mathbb{R}^N} \sum_{i=1}^N \mathrm{w}_i |z_i|, \ \text{ subject to } \||Az|^2 - b\|_2 \leq \varepsilon, \ \text{ where } \mathrm{w}_i = \begin{cases} \omega \in [0,1] & i \in \tilde{T}, \\ 1 & i \in \tilde{T}^c. \end{cases} \tag{18}$$

Here we adopt the compressive phase retrieval via lifting (CPRL) algorithm developed in [15] to solve this phaseless recovery problem. By using a lifting technique, this problem can be rewritten as a semidefinite program (SDP). More specifically, given the ground truth signal $x \in \mathbb{R}^N$, let $X = xx^T \in \mathbb{R}^{N \times N}$ be an induced rank-1 semidefinite matrix. We further denote $\Phi_i = a_i a_i^T$, a linear operator B of $Z = zz^T \in \mathbb{R}^{N \times N}$ as

$$B : Z \mapsto \{\mathrm{Tr}(\Phi_i Z)\}_{1 \leq i \leq m} \in \mathbb{R}^m$$

and the weight matrix $W = \mathrm{diag}\{\mathrm{w}_i, 1 \leq i \leq N\} \in \mathbb{R}^{N \times N}$. Then the phaseless vector recovery problem (18) can be cast as the following rank-1 matrix recovery problem:

$$\min_{Z \in \mathbb{R}^{N \times N}} \|WZW^T\|_1,$$
$$\text{subject to } \|B(Z) - b\|_2 \leq \varepsilon,$$
$$\mathrm{rank}(WZW^T) = 1, Z \succeq 0.$$

This is of course still a non-convex problem due to the rank constraint. The lifting approach addresses this issue by replacing $\mathrm{rank}(WZW^T)$ with $\mathrm{Tr}(WZW^T)$. This leads to an SDP:

$$\min_{Z \in \mathbb{R}^{N \times N}} \mathrm{Tr}(WZW^T) + \lambda\|WZW^T\|_1,$$
$$\text{subject to } \|B(Z) - b\|_2 \leq \varepsilon,$$
$$Z \succeq 0, \tag{19}$$

where $\lambda > 0$ is a design parameter. Then the estimate of x can be finally be found by computing the rank-1 decomposition of the recovered matrix via singular value decomposition.

The recovery performance is assessed by the average reconstruction signal to noise ratio (SNR) over 10 experiments. The SNR is measured in dB and it is given by

$$\mathrm{SNR}(x, x^\sharp) = 20 \log_{10} \left(\frac{\|x\|_2}{\min\{\|x^\sharp - x\|_2, \|x^\sharp + x\|_2\}} \right), \tag{20}$$

where x is the true signal and $x^\sharp$ is the recovered signal. For all the experiments, we fix the parameter $\lambda = 1$. In the experiments where the measurements are noisy, we set the noise $\{e_i, 1 \leq i \leq m\} \overset{i.i.d}{\sim} N(0, \sigma^2)$ with $\sigma = 0.1$ and $\varepsilon = \|e\|_2$.

4.1 Sparse Case

We first consider the case that x is exactly sparse with an ambient dimension $N = 32$ and fixed sparsity $k = 4$. The sparse signals are generated by choosing k nonzero positions uniformly at random, and then choosing the nonzero values from the standard normal distribution for these k nonzero positions. The recovery is done via (19) using a support estimate of size $|\tilde{T}| = 4$ (i.e., $\rho = 1$).

Figure 2 shows the recovery performances for different α and ω with an increasing number of measurements m, both in the noise free and noisy cases. It can be observed that when $\alpha = 0.75 > 0.5$, the best recovery is achieved for very small ω whereas a $\omega = 1$ results in the lowest SNR for both cases. On the other hand, when $\alpha = 0.25 < 0.5$, the performance of the recovery algorithms is better for large ω than that for small ω. The case $\omega = 0$ results in the lowest SNR. When $\alpha = 0.5$, the performance gaps for different ω are not particularly large and it seems that a medium ω ($\omega = 0.5$) achieves the best recovery. In the noise free case, a perfect recovery can be achieved as long as the number of measurements m is large enough. As is also expected that in all settings, comparing to the noise free case, we have a lower SNR in the noisy case. These findings are largely consistent with the theoretical results provided in Section 2.
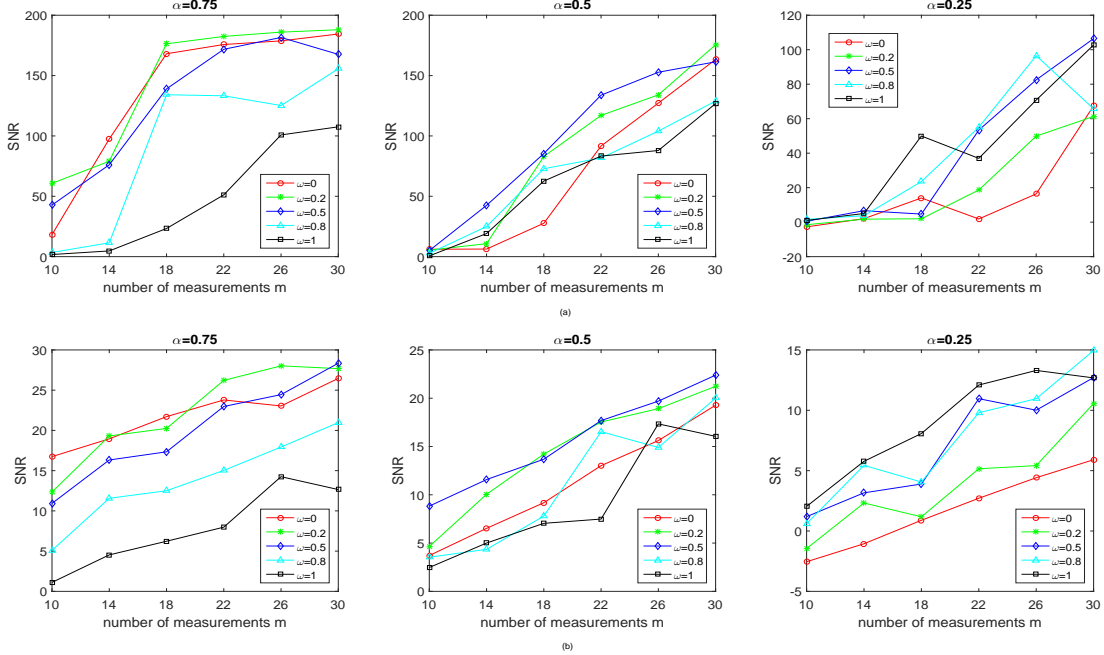

Figure 2: Performance of weighted ℓ_1 recovery in terms of SNR averaged over 10 experiments for sparse signals x with $N = 32$, $k = 4$, while varying the number of measurements m. From left to right, $\alpha = 0.75$, $\alpha = 0.5$ and $\alpha = 0.25$. (a) Noise Free. (b) $\sigma = 0.1$.

4.2 Compressible Case

Here we generate a signal x whose coefficients decay like $j^{-\theta}$ where $j \in \{1, \cdots, N\}$ and $\theta = 4.5$. This kind of signal itself is not sparse, but can be well approximated by an exactly sparse signal. For this experiment, we set $k = 4$, i.e., we use the best 4-term approximation. We fix $\rho = 1$ as in the sparse case. The phaseless recovery results are presented in Figure 3. It shows that on average a mediate value of ω ($\omega = 0.5$) results in the best recovery. In general, when $\alpha > 0.5$, smaller ω favours better reconstruction results. The opposite conclusion holds for the case that $\alpha < 0.5$. Therefore, as is expected that the behaviors that occur in the exactly sparse case also occur in the compressible case.
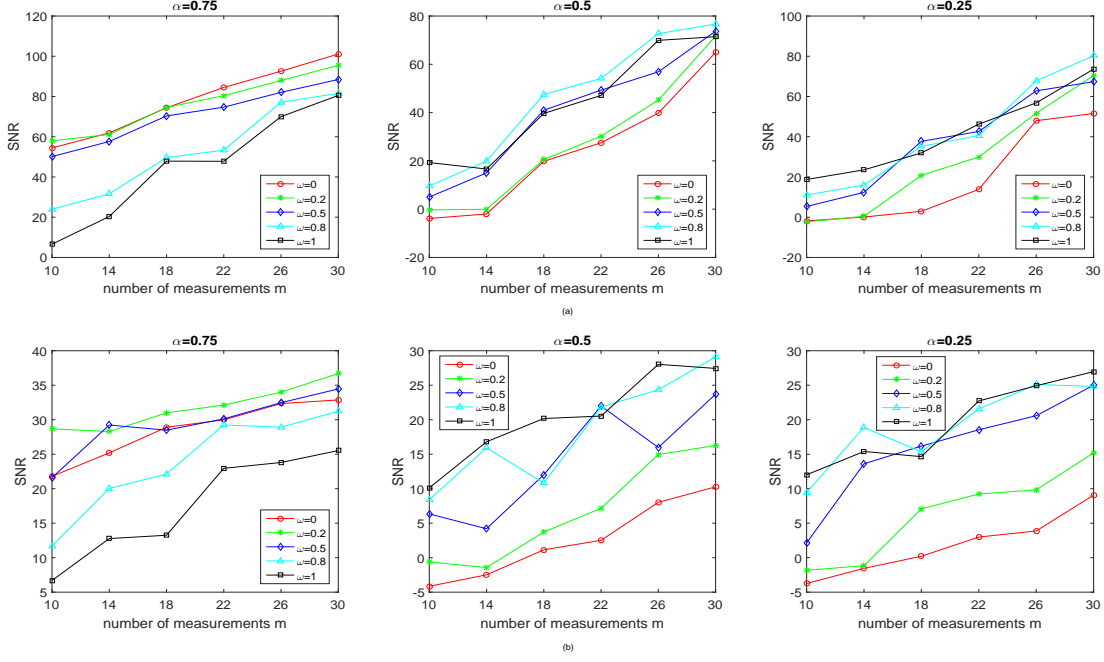
Figure 3: Performance of weighted ℓ_1 recovery in terms of SNR averaged over 10 experiments for compressible signals x with $N = 32$, $\theta = 4.5$, while varying the number of measurements m. From left to right, $\alpha = 0.75$, $\alpha = 0.5$ and $\alpha = 0.25$. (a) Noise Free. (b) $\sigma = 0.1$.

5　Conclusion

In this paper, we established the sufficient SRIP condition and the sufficient and necessary weighted null space property condition for phaseless compressive sensing using partial support information via weighted ℓ_1 minimization, and we conducted some numerical experiments to illustrate the theoretical results.

Some further problems are left for future work. As we only consider the real-valued signal reconstruction case, it will be challenging to generalize the present results to the complex-valued signal case. Besides it will be very interesting to construct the measurement matrix $A \in \mathbb{R}^{m \times N}$ satisfying the weighted null space property given in (15) directly.

Acknowledgements

References

[1] Cai, T. T. and Zhang, A. (2014). Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. *IEEE Transactions on Information Theory* **60**(1) 122–132.

[2] Candes, E. J., Eldar, Y. C., Strohmer, T. and Voroninski, V. (2015). Phase retrieval via matrix completion. *SIAM review* **57**(2) 225–251.

[3] Candes, E. J., Li, X. and Soltanolkotabi, M. (2015). Phase retrieval via Wirtinger flow: Theory and algorithms. *IEEE Transactions on Information Theory* **61**(4) 1985–2007.

[4] Candes, E. J., Strohmer, T. and Voroninski, V. (2013). Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics* **66**(8) 1241–1274.

[5] Candes, E. J. and Tao, T. (2005). Decoding by linear programming. *IEEE Transactions on Information Theory* **51**(12) 4203–4215.

[6] Chen, W. and Li, Y. (2016). Recovery of signals under the high order RIP condition via prior support information. *arXiv preprint arXiv:1603.03464.*

[7] Chen, Y. and Candes, E. (2015). Solving random quadratic systems of equations is nearly as easy as solving linear systems. In *Advances in Neural Information Processing Systems* (pp. 739–747).

[8] Eldar, Y. C. and Kutyniok, G. (2012). Compressed Sensing: Theory and Applications. Cambridge University Press.

[9] Foucart, S. and Rauhut, H. (2013). A Mathematical Introduction to Compressive Sensing. New York, NY, USA: Springer-Verlag.

[10] Friedlander, M. P., Mansour, H., Saab, R. and Yilmaz, O. (2012). Recovering compressively sampled signals using partial support information. *IEEE Transactions on Information Theory* **58**(2) 1122–1134.

[11] Gao, B., Wang, Y. and Xu, Z. (2016). Stable signal recovery from phaseless measurements. *Journal of Fourier Analysis and Applications* **22**(4) 787–808.

[12] Gao, B. and Xu, Z. (2017). Phaseless recovery using the GaussNewton method. *IEEE Transactions on Signal Processing* **65**(22) 5885–5896.

[13] Mansour, H. and Saab, R. (2017). Recovery analysis for weighted ℓ_1-minimization using the null space property. *Applied and Computational Harmonic Analysis* **43**(1) 23–38.

[14] Netrapalli, P., Jain, P. and Sanghavi, S. (2015). Phase retrieval using alternating minimization. *IEEE Transactions on Signal Processing* **18**(63) 4814–4826.

[15] Ohlsson, H., Yang, A. Y., Dong, R. and Sastry, S. S. (2011). Compressive phase retrieval from squared output measurements via semidefinite programming. *arXiv preprint arXiv:1111.6323.*

[16] Shechtman, Y., Beck, A. and Eldar, Y. C. (2014). GESPAR: Efficient phase retrieval of sparse signals. *IEEE Transactions on Signal Processing* **62**(4) 928–938.

[17] Voroninski, V. and Xu, Z. (2016). A strong restricted isometry property, with an application to phaseless compressed sensing. *Applied and Computational Harmonic Analysis* **40**(2) 386–395.

[18] Wang, Y. and Xu, Z. (2014). Phase retrieval for sparse signals. *Applied and Computational Harmonic Analysis* **37**(3) 531–544.

[19] Zhou, S., Xiu, N., Wang, Y. and Kong, L. (2013). Exact recovery for sparse signal via weighted ℓ_1 minimization. *arXiv preprint arXiv:1312.2358.*

[20] Zhou, Z. and Yu, J. (2017). Recovery analysis for weighted mixed ℓ_2/ℓ_p minimization with $0 < p \leq 1$. *arXiv preprint arXiv:1709.00257.*