Weighted-ℓ_1 minimization with multiple weighting sets

Hassan Mansour[a,b] and Özgür Yılmaz[a]

[a]Mathematics Department, University of British Columbia, Vancouver - BC, Canada;
[b]Computer Science Department, University of British Columbia, Vancouver - BC, Canada

ABSTRACT

In this paper, we study the support recovery conditions of weighted ℓ_1 minimization for signal reconstruction from compressed sensing measurements when multiple support estimate sets with different accuracy are available. We identify a class of signals for which the recovered vector from ℓ_1 minimization provides an accurate support estimate. We then derive stability and robustness guarantees for the weighted ℓ_1 minimization problem with more than one support estimate. We show that applying a smaller weight to support estimate that enjoy higher accuracy improves the recovery conditions compared with the case of a single support estimate and the case with standard, i.e., non-weighted, ℓ_1 minimization. Our theoretical results are supported by numerical simulations on synthetic signals and real audio signals.

Keywords: Compressed sensing, weighted ℓ_1 minimization, partial support recovery

1. INTRODUCTION

A wide range of signal processing applications rely on the ability to realize a signal from linear and sometimes noisy measurements. These applications include the acquisition and storage of audio, natural and seismic images, and video, which all admit sparse or approximately sparse representations in appropriate transform domains.

Compressed sensing has emerged as an effective paradigm for the acquisition of sparse signals from significantly fewer linear measurements than their ambient dimension.[1–3] Consider an arbitrary signal $x \in \mathbb{R}^N$ and let $y \in \mathbb{R}^n$ be a set of measurements given by

$$y = Ax + e,$$

where A is a known $n \times N$ measurement matrix, and e denotes additive noise that satisfies $\|e\|_2 \leq \epsilon$ for some known $\epsilon \geq 0$. Compressed sensing theory states that it is possible to recover x from y (given A) even when $n \ll N$, i.e., using very few measurements.

When x is strictly sparse, i.e. when there are only $k < n$ nonzero entries in x, and when $e = 0$, one may recover an estimate x^* of the signal x as the solution of the constrained ℓ_0 minimization problem

$$\underset{u \in \mathbb{R}^N}{\text{minimize}} \ \|u\|_0 \text{ subject to } \ Au = y. \tag{1}$$

In fact, using (1), the recovery is exact when $n \geq 2k$ and A is in general position.[4] However, ℓ_0 minimization is a combinatorial problem and quickly becomes intractable as the dimensions increase. Instead, the convex relaxation

$$\underset{u \in \mathbb{R}^N}{\text{minimize}} \ \|u\|_1 \text{ subject to } \|Au - y\|_2 \leq \epsilon \tag{2}$$

can be used to recover the estimate x^*. Candés, Romberg and Tao[2] and Donoho[1] show that if $n \gtrsim k \log(N/k)$, then ℓ_1 minimization (2) can stably and robustly recover x from inaccurate and what appears to be "incomplete" measurements $y = Ax+e$, where, as before, A is an appropriately chosen $n \times N$ measurement matrix and $\|e\|_2 \leq \epsilon$. Contrary to ℓ_0 minimization, (2), which is a convex program, can be solved efficiently. Consequently, it is possible to recover a stable and robust approximation of x by solving (2) instead of (1) at the cost of increasing the number of measurements taken.

Further author information: (Send correspondence to Hassan Mansour)
Hassan Mansour: E-mail: hassanm@cs.ubc.ca;
Özgür Yılmaz: E-mail: oyilmaz@math.ubc.ca

Several works in the literature have proposed alternate algorithms that attempt to bridge the gap between ℓ_0 and ℓ_1 minimization. For example, the recovery from compressed sensing measurements using ℓ_p minimization with $0 < p < 1$ has been shown to be stable and robust under weaker conditions that those of ℓ_1 minimization.[5–9] However, the problem is non-convex and even though various simple and efficient algorithms were proposed and observed to perform well empirically,[7,10] so far only local convergence can be proved. Another approach for improving the recovery performance of ℓ_1 minimization is to incorporate prior knowledge regarding the support of the signal to-be-recovered. One way to accomplish this is to replace ℓ_1 minimization in (2) with *weighted* ℓ_1 *minimization*

$$\underset{u}{\text{minimize}} \ \|u\|_{1,\mathrm{w}} \text{ subject to } \|Au - y\|_2 \leq \epsilon, \tag{3}$$

where $\mathrm{w} \in [0,1]^N$ and $\|u\|_{1,\mathrm{w}} := \sum_i \mathrm{w}_i |u_i|$ is the weighted ℓ_1 norm. This approach has been studied by several groups[11–14] and most recently, by the authors, together with Saab and Friedlander [15]. In this work, we proved that conditioned on the accuracy and relative size of the support estimate, weighted ℓ_1 minimization is stable and robust under weaker conditions than those of standard ℓ_1 minimization.

The works mentioned above mainly focus on a "two-weight" scenario: for $x \in \mathbb{R}^N$, one is given a partition of $\{1, \ldots, N\}$ into two sets, say $\widetilde{T}$ and $\widetilde{T}^c$. Here $\widetilde{T}$ denotes the estimated support of the entries of x that are largest in magnitude. In this paper, we consider the more general case and study recovery conditions of weighted ℓ_1 minimization when multiple support estimates with different accuracies are available. We first give a brief overview of compressed sensing and review our previous result on weighted ℓ_1 minimization in Section 2. In Section 3, we prove that for a certain class of signals it is possible to estimate the support of its best k-term approximation using standard ℓ_1 minimization. We then derive stability and robustness guarantees for weighted ℓ_1 minimization which generalizes our previous work to the case of two or more weighting sets. Finally, we present numerical experiments in Section 4 that verify our theoretical results.

2. COMPRESSED SENSING WITH PARTIAL SUPPORT INFORMATION

Consider an arbitrary signal $x \in \mathbb{R}^N$ and let x_k be its best k-term approximation, given by keeping the k largest-in-magnitude components of x and setting the remaining components to zero. Let $T_0 = \mathrm{supp}(x_k)$, where $T_0 \subseteq \{1, \ldots, N\}$ and $|T_0| \leq k$. We wish to reconstruct the signal x from $y = Ax + e$, where A is a known $n \times N$ measurement matrix with $n \ll N$, and e denotes the (unknown) measurement error that satisfies $\|e\|_2 \leq \epsilon$ for some known margin $\epsilon > 0$. Also let the set $\widetilde{T} \subset \{1, \ldots, N\}$ be an estimate of the support T_0 of x_k.

2.1 Compressed sensing overview

It was shown in [2] that x can be stably and robustly recovered from the measurements y by solving the optimization problem (1) if the measurement matrix A has the *restricted isometry property*[16] (RIP).

Definition 1. *The restricted isometry constant δ_k of a matrix A is the smallest number such that for all k-sparse vectors u,*

$$(1 - \delta_k)\|u\|_2^2 \leq \|Au\|_2^2 \leq (1 + \delta_k)\|u\|_2^2. \tag{4}$$

The following theorem uses the RIP to provide conditions and bounds for stable and robust recovery of x by solving (2).

Theorem 2.1 (Candès, Romberg, Tao[2]). *Suppose that x is an arbitrary vector in $\mathbb{R}^N$, and let x_k be the best k-term approximation of x. Suppose that there exists an $a \in \frac{1}{k}\mathbb{Z}$ with $a > 1$ and*

$$\delta_{ak} + a\delta_{(1+a)k} < a - 1. \tag{5}$$

Then the solution x^* to (2) obeys

$$\|x^* - x\|_2 \leq C_0 \epsilon + C_1 k^{-1/2} \|x - x_k\|_1. \tag{6}$$

Remark 1. *The constants in Theorem 2.1 are explicitly given by*

$$C_0 = \frac{2\left(1+a^{-1/2}\right)}{\sqrt{1-\delta_{(a+1)k}}-a^{-1/2}\sqrt{1+\delta_{ak}}}, \quad C_1 = \frac{2a^{-1/2}\left(\sqrt{1-\delta_{(a+1)k}}+\sqrt{1+\delta_{ak}}\right)}{\sqrt{1-\delta_{(a+1)k}}-a^{-1/2}\sqrt{1+\delta_{ak}}}. \tag{7}$$

Theorem 2.1 shows that the constrained ℓ_1 minimization problem in (2) recovers an approximation to x with an error that scales well with noise and the "compressibility" of x, provided (5) is satisfied. Moreover, if x is sufficiently sparse (i.e., $x = x_k$), and if the measurement process is noise-free, then Theorem 2.1 guarantees exact recovery of x from y. At this point, we note that a slightly stronger sufficient condition compared to (5)—that is easier to compare with conditions we obtain in the next section—is given by

$$\delta_{(a+1)k} < \frac{a-1}{a+1}. \tag{8}$$

2.2 Weighted ℓ_1 minimization

The ℓ_1 minimization problem (2) does not incorporate any prior information about the support of x. However, in many applications it may be possible to draw an estimate of the support of the signal or an estimate of the indices of its largest coefficients.

In our previous work,[15] we considered the case where we are given a support estimate $\widetilde{T} \subset \{1, \ldots, N\}$ for x with a certain accuracy. We investigated the performance of weighted ℓ_1 minimization, as described in (3), where the weights are assigned such that $\mathrm{w}_j = \omega \in [0,1]$ whenever $j \in \widetilde{T}$, and $\mathrm{w}_j = 1$ otherwise. In particular, we proved that if the (partial) support estimate is *at least* 50% accurate, then weighted ℓ_1 minimization with $\omega < 1$ outperforms standard ℓ_1 minimization in terms of accuracy, stability, and robustness.

Suppose that $\widetilde{T}$ has cardinality $|\widetilde{T}| = \rho k$, where $0 \le \rho \le N/k$ is the *relative size* of the support estimate $\widetilde{T}$. Furthermore, define the *accuracy* of $\widetilde{T}$ via $\alpha := \frac{\widetilde{T} \cap T_0}{|\widetilde{T}|}$, i.e., α is the fraction of $\widetilde{T}$ inside T_0. As before, we wish to recover an arbitrary vector $x \in \mathbb{R}^N$ from noisy compressive measurements $y = Ax + e$, where e satisfies $\|e\|_2 \le \epsilon$. To that end, we consider the weighted ℓ_1 minimization problem with the following choice of weights:

$$\underset{z}{\text{minimize}} \ \ \|z\|_{1,\mathrm{w}} \text{ subject to } \|Az - y\|_2 \le \epsilon \quad \text{with} \quad \mathrm{w}_i = \begin{cases} 1, & i \in \widetilde{T}^c, \\ \omega, & i \in \widetilde{T}. \end{cases} \tag{9}$$

Here, $0 \le \omega \le 1$ and $\|z\|_{1,\mathrm{w}}$ is as defined in (3). Figure 1 illustrates the relationship between the support T_0, support estimate $\widetilde{T}$ and the weight vector w.

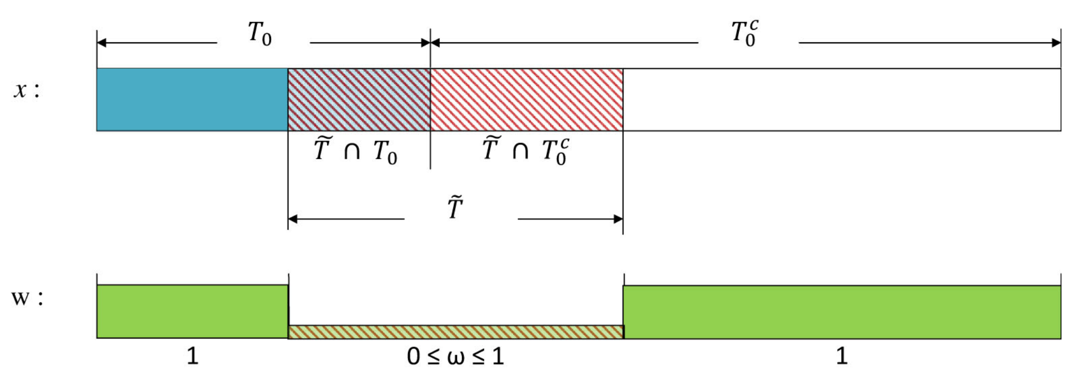

Figure 1. Illustration of the signal x and weight vector w emphasizing the relationship between the sets T_0 and $\widetilde{T}$.

Theorem 2.2 (FMSY[15]). *Let x be in $\mathbb{R}^N$ and let x_k be its best k-term approximation, supported on T_0. Let $\widetilde{T} \subset \{1, \ldots, N\}$ be an arbitrary set and define ρ and α as before such that $|\widetilde{T}| = \rho k$ and $|\widetilde{T} \cap T_0| = \alpha\rho k$. Suppose that there exists an $a \in \frac{1}{k}\mathbb{Z}$, with $a \geq (1-\alpha)\rho$, $a > 1$, and the measurement matrix A has RIP with*

$$\delta_{ak} + \frac{a}{\left(\omega + (1-\omega)\sqrt{1+\rho-2\alpha\rho}\right)^2}\delta_{(a+1)k} < \frac{a}{\left(\omega + (1-\omega)\sqrt{1+\rho-2\alpha\rho}\right)^2} - 1, \tag{10}$$

for some given $0 \leq \omega \leq 1$. Then the solution x^* to (9) obeys

$$\|x^* - x\|_2 \leq C_0'\epsilon + C_1' k^{-1/2}\left(\omega\|x - x_k\|_1 + (1-\omega)\|x_{\widetilde{T}^c \cap T_0^c}\|_1\right), \tag{11}$$

where C_0' and C_1' are well-behaved constants that depend on the measurement matrix A, the weight ω, and the parameters α and ρ.

Remark 2. *The constants C_0' and C_1' are explicitly given by the expressions*

$$C_0' = \frac{2\left(1 + \frac{\omega+(1-\omega)\sqrt{1+\rho-2\alpha\rho}}{\sqrt{a}}\right)}{\sqrt{1-\delta_{(a+1)k}} - \frac{\omega+(1-\omega)\sqrt{1+\rho-2\alpha\rho}}{\sqrt{a}}\sqrt{1+\delta_{ak}}}, \quad C_1' = \frac{2a^{-1/2}\left(\sqrt{1-\delta_{(a+1)k}} + \sqrt{1+\delta_{ak}}\right)}{\sqrt{1-\delta_{(a+1)k}} - \frac{\omega+(1-\omega)\sqrt{1+\rho-2\alpha\rho}}{\sqrt{a}}\sqrt{1+\delta_{ak}}}. \tag{12}$$

Consequently, Theorem 2.2, with $\omega = 1$, reduces to the stable and robust recovery theorem of [2], which we stated above—see Theorem 2.1.

Remark 3. *It is sufficient that A satisfies*

$$\delta_{(a+1)k} < \hat{\delta}^{(\omega)} := \frac{a - \left(\omega + (1-\omega)\sqrt{1+\rho-2\alpha\rho}\right)^2}{a + \left(\omega + (1-\omega)\sqrt{1+\rho-2\alpha\rho}\right)^2} \tag{13}$$

for Theorem 2.2 to hold, i.e., to guarantee stable and robust recovery of the signal x from measurements $y = Ax + e$.

It is easy to see that the sufficient conditions of Theorem 2.2, given in (10) or (13), are weaker than their counterparts for the standard ℓ_1 recovery, as given in (5) or (8) respectively, if and only if $\alpha > 0.5$. A similar statement holds for the constants. In words, if the support estimate is more than 50% accurate, weighted ℓ_1 is more favorable than ℓ_1, at least in terms of sufficient conditions and error bounds.

The theoretical results presented above suggest that the weight ω should be set equal to zero when $\alpha \geq 0.5$ and to one when $\alpha < 0.5$ as these values of ω give the best sufficient conditions and error bound constants. However, we conducted extensive numerical simulations in [15] which suggest that a choice of $\omega \approx 0.5$ results in the best recovery when there is little confidence in the support estimate accuracy. An heuristic explanation of this observation is given in [15].

3. WEIGHTED ℓ_1 MINIMIZATION WITH MULTIPLE SUPPORT ESTIMATES

The result in the previous section relies on the availability of a support estimate set $\widetilde{T}$ on which to apply the weights ω. In this section, we first show that it is possible to draw support estimates from the solution of (2). We then present the main theorem for stable and robust recovery of an arbitrary vector $x \in \mathbb{R}^N$ from measurements $y = Ax + e$, $y \in \mathbb{R}^n$ and $n \ll N$, with multiple support estimates having different accuracies.

3.1 Partial support recovery from ℓ_1 minimization

For signals x that belong to certain signal classes, the solution to the ℓ_1 minimization problem can carry significant information on the support T_0 of the best k-term approximation x_k of x. We start by recalling the *null space property* (NSP) of a matrix A as defined in [17]. Necessary conditions as well as sufficient conditions for the existence of some algorithm that recovers x from measurements $y = Ax$ with an error related to the best k-term approximation of x can be formulated in terms of an appropriate NSP. We state below a particular form of the NSP pertaining to the ℓ_1-ℓ_1 instance optimality.

Definition 2. *A matrix $A \in \mathbb{R}^{n\times N}$, $n < N$, is said to have the null space property of order k and constant c_0 if for any vector $h \in \mathcal{N}(A)$, $Ah = 0$, and for every index set $T \subset \{1 \dots N\}$ of cardinality $|T| = k$*

$$\|h\|_1 \le c_0 \|h_{T^c}\|_1.$$

Among the various important conclusions of [17], the following (in a slightly more general form) will be instrumental for our results.

Lemma 3.1 ([17]). *If A has the restricted isometry property with $\delta_{(a+1)k} < \frac{a-1}{a+1}$ for some $a > 1$, then it has the NSP of order k and constant c_0 given explicitly by*

$$c_0 = 1 + \frac{\sqrt{1+\delta_{ak}}}{\sqrt{a}\sqrt{1-\delta_{(a+1)k}}}.$$

In what follows, let x^* be the solution to (2) and define the sets $S = \text{supp}(x_s)$, $T_0 = \text{supp}(x_k)$, and $\widetilde{T} = \text{supp}(x^*_k)$ for some integers $k \ge s > 0$.

Proposition 3.2. *Suppose that A has the null space property (NSP) of order k with constant c_0 and*

$$\min_{j\in S} |x(j)| \ge (\eta + 1)\|x_{T_0^c}\|_1, \tag{14}$$

where $\eta = \frac{2c_0}{2-c_0}$. Then $S \subseteq \widetilde{T}$.

The proof is presented in section A of the appendix.

Remark 4. *Note that if A has RIP so that $\delta_{(a+1)k} < \frac{a-1}{a+1}$ for some $a > 1$, then η is given explicitly by*

$$\eta = \frac{2(\sqrt{a}\sqrt{1-\delta_{(a+1)k}} + \sqrt{1+\delta_{ak}})}{\sqrt{a}\sqrt{1-\delta_{(a+1)k}} - \sqrt{1+\delta_{ak}}}. \tag{15}$$

Proposition 3.2 states that if x belongs to the class of signals that satisfy (14), then the support S of x_s—i.e., the set of indices of the s largest-in-magnitude coefficients of x—is guaranteed to be contained in the set of indices of the k largest-in-magnitude coefficients of x^*. Consequently, if we consider $\widetilde{T}$ to be a support estimate for x_k, then it has an accuracy $\alpha \ge \frac{s}{k}$.

Note here that Proposition 3.2 specifies a class of signals, defined via (14), for which partial support information can be obtained by using the standard ℓ_1 recovery method. Though this class is quite restrictive and does not include various signals of practical interest, experiments suggest that highly accurate support estimates can still be obtained via ℓ_1 minimization for signals that only satisfy significantly milder decay conditions than (14). A theoretical investigation of this observation is an open problem.

3.2 Multiple support estimates with varying accuracy: an idealized motivating example

Suppose that the entries of x decay according to a power law such that $|x(j)| = cj^{-p}$ for some scaling constant c, $p > 1$ and $j \in \{1, \dots, N\}$. Consider the two support sets $T_1 = \text{supp}(x_{k_1})$ and $T_2 = \text{supp}(x_{k_2})$ for $k_1 > k_2$, $T_2 \subset T_1$. Suppose also that we can find entries $|x(s_1)| = cs_1^{-p} \approx c(\eta+1)\frac{k_1^{1-p}}{p-1}$ and $x(s_2) = cs_2^{-p} \approx c(\eta+1)\frac{k_2^{1-p}}{p-1}$ that satisfy (14) for the sets T_1 and T_2, respectively, where $s_1 \le k_1$ and $s_2 \le k_2$. Then

$$\begin{aligned} s_1 - s_2 &= \left(\tfrac{p-1}{\eta+1}\right)^{1/p}\left(k_1^{1-1/p} - k_2^{1-1/p}\right) \\ &\le \left(\tfrac{p-1}{\eta+1}\right)^{1/p}(k_1 - k_2). \end{aligned}$$

which follows because $0 < 1 - 1/p < 1$ and $k_1 - k_2 \ge 1$.

Consequently, if we define the support estimate sets $\widetilde{T}_1 = \text{supp}(x^*_{k_1})$ and $\widetilde{T}_2 = \text{supp}(x^*_{k_2})$, clearly the corresponding accuracies $\alpha_1 = \frac{s_1}{k_1}$ and $\alpha_2 = \frac{s_2}{k_2}$ are not necessarily equal. Moreover, if

$$\left(\frac{p-1}{\eta+1}\right)^{1/p} < \alpha_1, \tag{16}$$

$s_1 - s_2 < \alpha_1(k_1 - k2)$, and thus $\alpha_1 < \alpha_2$. For example, if we have $p = 1.3$ and $\eta = 5$, we get $\left(\frac{p-1}{\eta+1}\right)^{1/p} \approx 0.1$. Therefore, in this particular case, if $\alpha_1 > 0.1$, choosing some $k_2 < k_1$ results in $\alpha_2 > \alpha_1$, i.e., we identify two different support estimates with different accuracies. This observation raises the question, "How should we deal with the recovery of signals from CS measurements when multiple support estimates with different accuracies are available?" We propose an answer to this question in the next section.

3.3 Stability and robustness conditions

In this section we present our main theorem for stable and robust recovery of an arbitrary vector $x \in \mathbb{R}^N$ from measurements $y = Ax + e$, $y \in \mathbb{R}^n$ and $n \ll N$, with multiple support estimates having different accuracies. Figure 2 illustrates an example of the particular case when only two disjoint support estimate sets are available.

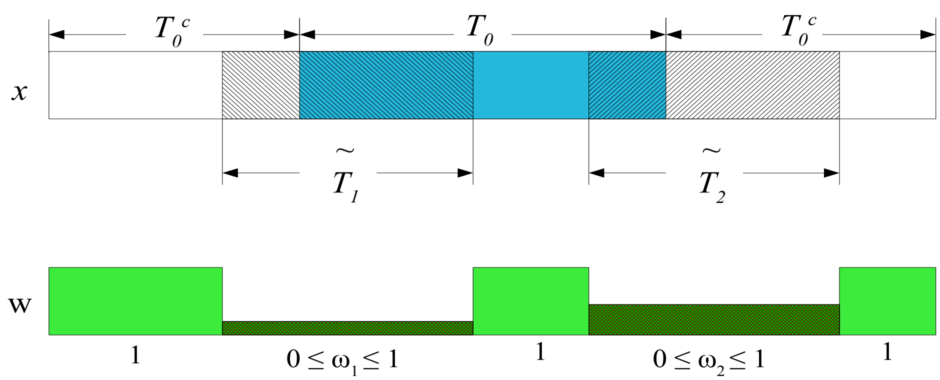

Figure 2. Example of a sparse vector x with support set T_0 and two support estimate sets $\widetilde{T}_1$ and $\widetilde{T}_2$. The weight vector is chosen so that weights ω_1 and ω_2 are applied to the sets $\widetilde{T}_1$ and $\widetilde{T}_2$, respectively, and a weight equal to one elsewhere.

Let T_0 be the support of the best k-term approximation x_k of the signal x. Suppose that we have a support estimate $\widetilde{T}$ that can be written as the union of m disjoint subsets $\widetilde{T}_j$, $j \in \{1, \ldots, m\}$, each of which has cardinality $|\widetilde{T}_j| = \rho_j k$, $0 \le \rho_j \le a$ for some $a > 1$ and accuracy $\alpha_j = \frac{|\widetilde{T}_j \cap T_0|}{|\widetilde{T}_j|}$.

Again, we wish to recover x from measurements $y = Ax + e$ with $\|e\|_2 \le \epsilon$. To do this, we consider the general weighted ℓ_1 minimization problem

$$\min_{u \in \mathbb{R}^N} \|u\|_{1,\mathrm{w}} \quad \text{subject to} \quad \|Au - y\| \le \epsilon \tag{17}$$

where $\|u\|_{1,\mathrm{w}} = \sum\limits_{i=1}^{N} \mathrm{w}_i |u_i|$, and $\mathrm{w}_i = \begin{cases} \omega_1, & i \in \widetilde{T}_1 \\ \vdots & \\ \omega_m, & i \in \widetilde{T}_m \\ 1, & i \in \widetilde{T}^c \end{cases}$ for $0 \le \omega_j \le 1$, for all $j \in \{1, \dots, m\}$ and $\widetilde{T} = \bigcup\limits_{j=1}^{m} \widetilde{T}_j$.

Theorem 3.3. *Let $x \in \mathbb{R}^n$ and $y = Ax + e$, where A is an $n \times N$ matrix and e is additive noise with $\|e\|_2 \le \epsilon$ for some known $\epsilon > 0$. Denote by x_k the best k-term approximation of x, supported on T_0 and let $\widetilde{T}_1, \dots, \widetilde{T}_m \subset \{1, ..., N\}$ be as defined above with cardinality $|\widetilde{T}_j| = \rho_j k$ and accuracy $\alpha_j = \frac{|\widetilde{T}_j \cap T_0|}{|\widetilde{T}_j|}$, $j \in \{1, \dots, m\}$. For some given $0 \le \omega_1, \dots, \omega_m \le 1$, define $\gamma := \sum\limits_{j=1}^{m} \omega_j - (m-1) + \sum\limits_{j=1}^{m} (1-\omega_j)\sqrt{1 + \rho_j - 2\alpha_j \rho_j}$. If the RIP constants of A are such that there exists an $a \in \frac{1}{k}\mathbb{Z}$, with $a > 1$, and*

$$\delta_{ak} + \frac{a}{\gamma^2}\delta_{(a+1)k} < \frac{a}{\gamma^2} - 1, \tag{18}$$

then the solution $x^{\#}$ to (17) obeys

$$\|x^{\#} - x\|_2 \le C_0(\gamma)\epsilon + C_1(\gamma)k^{-1/2} \left(\sum_{j=1}^{m} \omega_j \|x_{\widetilde{T}_j \cap T_0^c}\|_1 + \|x_{\widetilde{T}^c \cap T_0^c}\|_1 \right). \tag{19}$$

The proof is presented in section B of the appendix.

Remark 5. *The constants $C_0(\gamma)$ and $C_1(\gamma)$ are well-behaved and given explicitly by the expressions*

$$C_0(\gamma) = \frac{2\left(1 + \frac{\gamma}{\sqrt{a}}\right)}{\sqrt{1 - \delta_{(a+1)k}} - \frac{\gamma}{\sqrt{a}}\sqrt{1 + \delta_{ak}}}, \quad C_1(\gamma) = \frac{2a^{-1/2}\left(\sqrt{1 - \delta_{(a+1)k}} + \sqrt{1 + \delta_{ak}}\right)}{\sqrt{1 - \delta_{(a+1)k}} - \frac{\gamma}{\sqrt{a}}\sqrt{1 + \delta_{ak}}}. \tag{20}$$

Remark 6. *Theorem 3.3 is a generalization of Theorem 2.2 for $m \ge 1$ support estimates. It is easy to see that when the number of support estimates $m = 1$, Theorem 3.3 reduces to the recovery conditions of Theorem 2.2. Moreover, setting $\omega_j = 1$ for all $j \in \{1, \dots, m\}$ reduces the result to that in Theorem 2.1.*

Remark 7. *The sufficient recovery condition* (13) *becomes in the case of multiple support estimates*

$$\delta_{(a+1)k} < \hat{\delta}^{(\gamma)} := \frac{a - \gamma^2}{a + \gamma^2}, \tag{21}$$

where γ is as defined in Theorem 3.3. It can be shown that when $m = 1$, γ reduces to the expression in (13).

Remark 8. *The value of γ controls the recovery guarantees of the multiple-set weighted ℓ_1 minimization problem. For instance, as γ approaches 0, condition* (21) *becomes weaker and the error bound constants $C_0(\gamma)$ and $C_1(\gamma)$ become smaller. Therefore, given a set of support estimate accuracies α_j for all $j \in \{1 \dots m\}$, it is useful to find the corresponding weights ω_j that minimize γ. Notice that for all j, γ is a sum of linear functions of ω_j with α_j controlling the slope. When $\alpha_j > 0.5$, the slope is positive and the optimal value of $\omega_j = 0$. Otherwise, when $\alpha_j \le 0.5$, the slope is negative and the optimal value of $\omega_j = 1$. Hence, as in the single support estimate case, the theoretical conditions indicate that when the α_j are known a choice of ω_j equal to zero or one should be optimal. However, when the knowledge of α_j is not reliable, experimental results indicate that intermediate values of ω_j produce the best recovery results.*

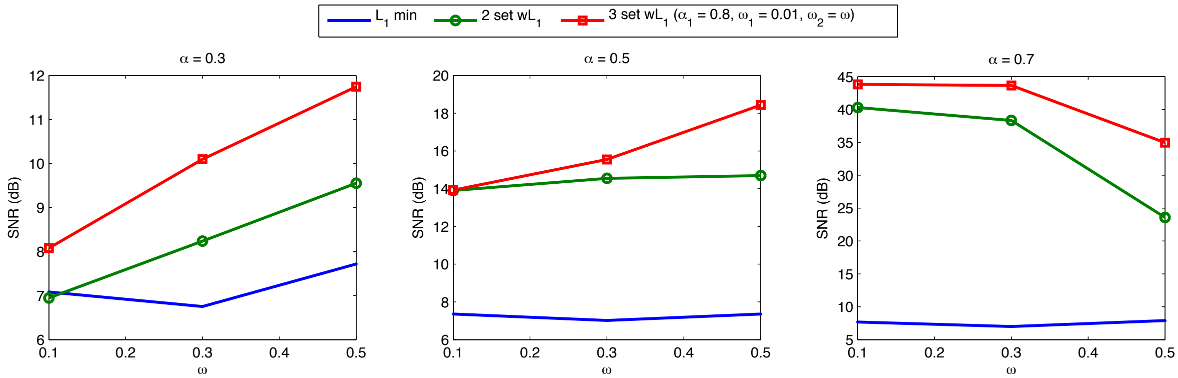

Figure 3. Comparison between the recovered SNR (averaged over 100 experiments) using two-set weighted ℓ_1 with support estimate $\widetilde{T}$ and accuracy α, three-set weighted ℓ_1 minimization with support estimates $\widetilde{T}_1 \cup \widetilde{T}_2 = \widetilde{T}$ and accuracy $\alpha_1 = 0.8$ and $\alpha_2 < \alpha$, and non-weighted ℓ_1 minimization.

4. NUMERICAL EXPERIMENTS

In what follows, we consider the particular case of $m = 2$, i.e. where there exists prior information on two disjoint support estimates $\widetilde{T}_1$ and $\widetilde{T}_2$ with respective accuracies α_1 and α_2. We present numerical experiments that illustrate the benefits of using three-set weighted ℓ_1 minimization over two-set weighted ℓ_1 and non-weighted ℓ_1 minimization when additional prior support information is available.

To that end, we compare the recovery capabilities of these algorithms for a suite of synthetically generated sparse signals. We also present the recovery results for a practical application of recovering audio signals using the proposed weighting. In all of our experiments, we use SPGL1[18,19] to solve the standard and weighted ℓ_1 minimization problems.

4.1 Recovery of synthetic signals

We generate signals x with an ambient dimension $N = 500$ and fixed sparsity $k = 35$. We compute the (noisy) compressed measurements of x using a Gaussian random measurement matrix A with dimensions $n \times N$ where $n = 100$. To quantify the reconstruction quality, we use the reconstruction signal to noise ratio (SNR) average over 100 realizations of the same experimental conditions. The SNR is measured in dB and is given by

$$\text{SNR}(x, \tilde{x}) = 10 \log_{10} \left(\frac{\|x\|_2^2}{\|x - \tilde{x}\|_2^2} \right), \tag{22}$$

where x is the true signal and $\tilde{x}$ is the recovered signal.

The recovery via two-set weighted ℓ_1 minimization uses a support estimate $\widetilde{T}$ of size $|\widetilde{T}| = 40$ (i.e., $\rho = 1$) where the accuracy α of the support estimate takes on the values $\{0.3, 0.5, 0.7\}$, and the weight ω is chosen from $\{0.1, 0.3, 0.5\}$.

Recovery via three-set weighted ℓ_1 minimization assumes the existence of two support estimates $\widetilde{T}_1$ and $\widetilde{T}_2$, which are disjoint subsets of $\widetilde{T}$ described above. The set $\widetilde{T}_1$ is chosen such that it always has an accuracy $\alpha_1 = 0.8$ while $\widetilde{T}_2 = \widetilde{T} \setminus \widetilde{T}_1$. In all experiments, we fix $\omega_1 = 0.01$ and set $\omega_2 = \omega$.

Figure 4.1 illustrates the recovery performance of three-set weighted ℓ_1 minimization compared to two-set weighted ℓ_1 using the setup described above and non-weighted ℓ_1 minimization. The figure shows that utilizing the extra accuracy of $\widetilde{T}_1$ by setting a smaller weight ω_1 results in better signal recovery from the same measurements.

4.2 Recovery of audio signals

Next, we examine the performance of three-set weighted ℓ_1 minimization for the recovery of compressed sensing measurements of speech signals. In particular, the original signals are sampled at 44.1 kHz, but only 1/4th of the samples are retained (with their indices chosen randomly from the uniform distribution). This yields the measurements $y = Rs$, where s is the speech signal and R is a restriction (of the identity) operator. Consequently, by dividing the measurements into blocks of size N, we can write $y = [y_1^T, y_2^T, ...]^T$. Here each $y_j = R_j s_j$ is the measurement vector corresponding to the jth block of the signal, and $R_j \in \mathbb{R}^{n_j \times N}$ is the associated restriction matrix. The signals we use in our experiments consist of 21 such blocks.

We make the following assumptions about speech signals:

1. The signal blocks are compressible in the DCT domain (for example, the MP3 compression standard uses a version of the DCT to compress audio signals.)

2. The support set corresponding to the largest coefficients in adjacent blocks does not change much from block to block.

3. Speech signals have large low-frequency coefficients.

Thus, for the reconstruction of the jth block, we identify the support estimates $\widetilde{T}_1$ is the set corresponding to the largest $n_j/16$ recovered coefficients of the previous block (for the first block $\widetilde{T}_1$ is empty) and $\widetilde{T}_2$ is the set corresponding to frequencies up to 4kHz. For recovery using two-set weighted ℓ_1 minimization, we define $\widetilde{T} = \widetilde{T}_1 \cup \widetilde{T}_2$ and assign it a weight of ω. In the three-set weighted ℓ_1 case, we assign weights $\omega_1 = \omega/2$ on the set $\widetilde{T}_1$ and $\omega_2 = \omega$ on the set $\widetilde{T} \setminus \widetilde{T}_1$. The results of experiments on an example speech signal with $N = 2048$, and $\omega \in \{0, 1/6, 2/6, \ldots, 1\}$ are illustrated in Figure 4. It is clear from the figure that three-set weighted ℓ_1 minimization has better recovery performance over all 10 values of ω spanning the interval $[0, 1]$.

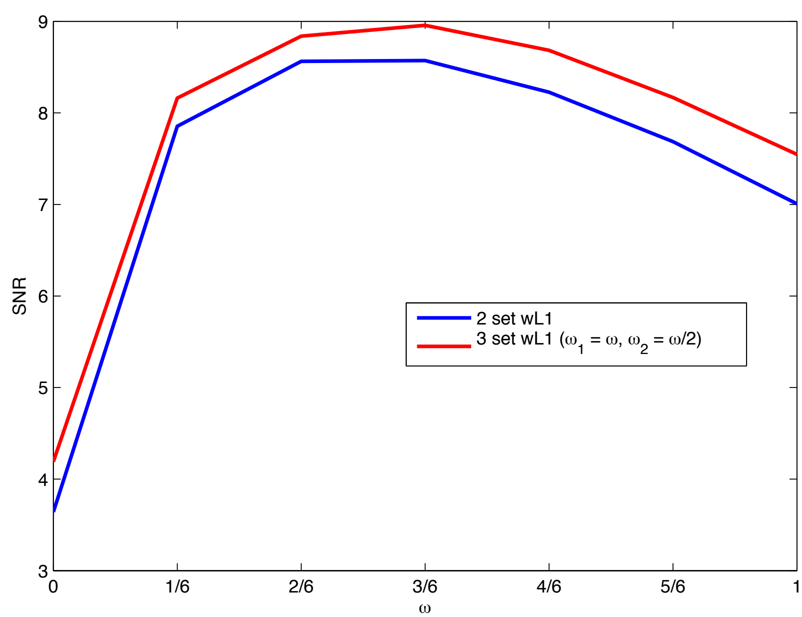

Figure 4. SNRs of the two reconstruction algorithms two-set and three-set weighted ℓ_1 minimization for a speech signal from compressed sensing measurements plotted against ω.

5. CONCLUSION

In conclusion, we derived stability and robustness guarantees for the weighted ℓ_1 minimization problem with multiple support estimates with varying accuracy. We showed that incorporating additional support information by applying a smaller weight to the estimated subsets of the support with higher accuracy improves the recovery conditions compared with the case of a single support estimate and the case of (non-weighted) ℓ_1 minimization. We also showed that for a certain class of signals—the coefficients of which decay in a particular way—it is possible to draw a support estimate from the solution of the ℓ_1 minimization problem. These results raise the question of whether it is possible to improve on the support estimate by solving a subsequent weighted ℓ_1 minimization problem. Moreover, it raises an interest in defining a new iterative weighted ℓ_1 algorithm which depends on the support accuracy instead of the coefficient magnitude as is the case of the Candès, Wakin, and Boyd[20] (IRL1) algorithm. We shall consider these problems elsewhere.

APPENDIX A. PROOF OF PROPOSITION 3.2

We want to find the conditions on the signal x and the matrix A which guarantee that the solution x^* to the ℓ_1 minimization problem (2) has the following property

$$\min_{j\in S}|x^*(j)| \geq \max_{j\in \widetilde{T}^c}|x^*(j)| = |x^*(k+1)|.$$

Suppose that the matrix A has the Null Space property (NSP)[17] of order k, i.e., for any $h \in \mathcal{N}(A)$, $Ah = 0$, then

$$\|h\|_1 \leq c_0\|h_{T_0^c}\|_1,$$

where $T_0 \subset \{1, 2, \dots N\}$ with $|T_0| = k$, and $\mathcal{N}(A)$ denotes the Null-Space of A.

If A has RIP with $\delta_{(a+1)k} < \frac{a-1}{a+1}$ for some constant $a > 1$, then it has the NSP of order k with constant c_0 which can be written explicitly in terms of the RIP constant of A as follows

$$c_0 = 1 + \frac{\sqrt{1+\delta_{ak}}}{\sqrt{a}\sqrt{1-\delta_{(a+1)k}}}.$$

Define $h = x^* - x$, then $h \in \mathcal{N}(A)$ and we can write the ℓ_1-ℓ_1 instance optimality as follows

$$\|h\|_1 \leq \frac{2c_0}{2-c_0}\|x_{T_0^c}\|_1,$$

with $c_0 < 2$. Let $\eta = \frac{2c_0}{2-c_0}$, the bound on $\|h_{T_0}\|_1$ is then given by

$$\|h_{T_0}\|_1 \leq (\eta+1)\|x_{T_0^c}\|_1 - \|x^*_{T_0^c}\|_1. \tag{23}$$

The next step is to bound $\|x^*_{T_0^c}\|_1$. Noting that $\widetilde{T} = \text{supp}(x^*_k)$, then $\|x^*_{\widetilde{T}}\|_1 \leq \|x^*_{T_0^c}\|_1$, and

$$\|x^*_{T_0^c}\|_1 \geq \|x^*_{\widetilde{T}^c}\|_1 \geq |x^*(k+1)|.$$

Using the reverse triangle inequality, we have $\forall j$, $|x(j) - x^*(j)| \geq |x(j)| - |x^*(j)|$ which leads to

$$\min_{j\in S}|x^*(j)| \geq \min_{j\in S}|x(j)| - \max_{j\in S}|x(j) - x^*(j)|.$$

But $\max_{j\in S}|x(j) - x^*(j)| = \|h_S\|_\infty \leq \|h_S\|_1 \leq \|h_{T_0}\|_1$, so combining the above three equations we get

$$\min_{j\in S}|x^*(j)| \geq |x^*(k+1)| + \min_{j\in S}|x(j)| - (\eta+1)\|x_{T_0^c}\|_1. \tag{24}$$

Equation (24) says that if the matrix A has $\delta_{(a+1)k}$-RIP and the signal x obeys

$$\min_{j\in S}|x(j)| \geq (\eta+1)\|x_{T_0^c}\|_1,$$

then the support $\widetilde{T}$ of the largest k entries of the solution x^* to (2) contains the support S of the largest s entries of the signal x.

APPENDIX B. PROOF OF THEOREM 3.3

The proof of Theorem 3.3 follows in the same line as our previous work in [15] with some modifications. Recall that the sets $\widetilde{T}_j$ and disjoint and $\widetilde{T} = \bigcup_{j=1}^{m} \widetilde{T}_j$, and define the sets $\widetilde{T}_{j\alpha} = T_0 \cap \widetilde{T}_j$, for all $j \in \{1,\dots,m\}$, where $|\widetilde{T}_{j\alpha}| = \alpha_j \rho_j k$.

Let $x^{\#} = x + h$ be the minimizer of the weighted ℓ_1 problem (17). Then

$$\|x+h\|_{1,\mathrm{w}} \leq \|x\|_{1,\mathrm{w}}.$$

Moreover, by the choice of weights in (17), we have

$$\omega_1 \|x_{\widetilde{T}_1} + h_{\widetilde{T}_1}\|_1 + \dots \omega_m \|x_{\widetilde{T}_m} + h_{\widetilde{T}_m}\|_1 + \|x_{\widetilde{T}^c} + h_{\widetilde{T}^c}\|_1 \leq \omega_1 \|x_{\widetilde{T}_1}\|_1 \dots + \omega_m \|x_{\widetilde{T}_m}\|_1 + \|x_{\widetilde{T}^c}\|_1.$$

Consequently,

$$\begin{aligned}
\|x_{\widetilde{T}^c\cap T_0} + h_{\widetilde{T}^c\cap T_0}\|_1 + \|x_{\widetilde{T}^c\cap T_0^c} + h_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\left(\omega_j \|x_{\widetilde{T}_j\cap T_0} + h_{\widetilde{T}_j\cap T_0}\|_1 + \omega_j \|x_{\widetilde{T}_j\cap T_0^c} + h_{\widetilde{T}_j\cap T_0^c}\|_1\right) \\
\leq \quad \|x_{\widetilde{T}^c\cap T_0}\|_1 + \|x_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\left(\omega_j \|x_{\widetilde{T}_j\cap T_0}\|_1 + \omega_j \|x_{\widetilde{T}_j\cap T_0^c}\|_1\right).
\end{aligned}$$

Next, we use the forward and reverse triangle inequalities to get

$$\sum_{j=1}^{m}\left(\omega_j \|h_{\widetilde{T}_j\cap T_0^c}\|_1\right) + \|h_{\widetilde{T}^c\cap T_0^c}\|_1 \leq \|h_{\widetilde{T}^c\cap T_0}\|_1 + \sum_{j=1}^{m}\omega_j \|h_{\widetilde{T}_j\cap T_0}\|_1 + 2\left(\|x_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\omega_j \|x_{\widetilde{T}_j\cap T_0^c}\|_1\right).$$

Adding $\sum_{j=1}^{m}(1-\omega_j)\|h_{\widetilde{T}_j^c\cap T_0^c}\|_1$ on both sides of the inequality above we obtain

$$\begin{aligned}
\sum_{j=1}^{m}\|h_{\widetilde{T}_j\cap T_0^c}\|_1 + \|h_{\widetilde{T}^c\cap T_0^c}\|_1 \quad &\leq \sum_{j=1}^{m}\omega_j \|h_{\widetilde{T}_j\cap T_0}\|_1 + \sum_{j=1}^{m}(1-\omega_j)\|h_{\widetilde{T}_j\cap T_0^c}\|_1 + \|h_{\widetilde{T}^c\cap T_0}\|_1 \\
&+ \quad 2\left(\|x_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\omega_j \|x_{\widetilde{T}_j\cap T_0^c}\|_1\right).
\end{aligned}$$

Since $\|h_{T_0^c}\|_1 = \|h_{\widetilde{T}\cap T_0^c}\|_1 + \|h_{\widetilde{T}^c\cap T_0^c}\|_1$ and $\|h_{\widetilde{T}\cap T_0^c}\|_1 = \sum_{j=1}^{m}\|h_{\widetilde{T}_j\cap T_0^c}\|_1$, this easily reduces to

$$\|h_{T_0^c}\|_1 \leq \sum_{j=1}^{m}\omega_j \|h_{\widetilde{T}_j\cap T_0}\|_1 + \sum_{j=1}^{m}(1-\omega_j)\|h_{\widetilde{T}_j\cap T_0^c}\|_1 + \|h_{\widetilde{T}^c\cap T_0}\|_1 + 2\left(\|x_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\omega_j \|x_{\widetilde{T}_j\cap T_0^c}\|_1\right). \tag{25}$$

Now consider the following term from the left hand side of (25)

$$\sum_{j=1}^{m}\omega_j \|h_{\widetilde{T}_j\cap T_0}\|_1 + \sum_{j=1}^{m}(1-\omega_j)\|h_{\widetilde{T}_j\cap T_0^c}\|_1 + \|h_{\widetilde{T}^c\cap T_0}\|_1$$

Add and subtract $\sum_{j=1}^{m}(1-\omega_j)\|h_{\widetilde{T}_j^c\cap T_0}\|_1$, and since the set $\widetilde{T}_{j\alpha} = T_0 \cap \widetilde{T}_j$, we can write $\|h_{\widetilde{T}_j^c\cap T_0}\|_1 + \|h_{\widetilde{T}_j\cap T_0^c}\|_1 = \|h_{T_0\cup\widetilde{T}\setminus\widetilde{T}_{j\alpha}}\|_1$ to get

$$\begin{aligned}
&\sum_{j=1}^{m}\omega_j\left(\|h_{\widetilde{T}_j\cap T_0}\|_1 + \|h_{\widetilde{T}_j^c\cap T_0}\|_1\right) + \sum_{j=1}^{m}(1-\omega_j)\left(\|h_{\widetilde{T}_j\cap T_0^c}\|_1 + \|h_{\widetilde{T}_j^c\cap T_0}\|_1\right) + \|h_{\widetilde{T}^c\cap T_0}\|_1 - \sum_{j=1}^{m}\|h_{\widetilde{T}_j^c\cap T_0}\|_1 \\
= \quad &\left(\sum_{j=1}^{m}\omega_j\right)\|h_{T_0}\|_1 + \|h_{\widetilde{T}^c\cap T_0}\|_1 - \sum_{j=1}^{m}\|h_{\widetilde{T}_j^c\cap T_0}\|_1 + \sum_{j=1}^{m}(1-\omega_j)\|h_{T_0\cup\widetilde{T}\setminus\widetilde{T}_{j\alpha}}\|_1 \\
= \quad &\left(\sum_{j=1}^{m}\omega_j - m + 1\right)\|h_{T_0}\|_1 + \sum_{j=1}^{m}(1-\omega_j)\|h_{T_0\cup\widetilde{T}\setminus\widetilde{T}_{j\alpha}}\|_1.
\end{aligned}$$

The last equality comes from $\|h_{T_0\cap\widetilde{T}_j^c}\|_1 = \|h_{\widetilde{T}^c\cap T_0}\|_1 + \|h_{T_0\cap(\widetilde{T}\setminus\widetilde{T}_j)}\|_1$ and $\sum_{j=1}^{m}\|h_{T_0\cap(\widetilde{T}\setminus\widetilde{T}_j)}\|_1 = (m-1)\|h_{T_0\cap\widetilde{T}}\|_1$.

Consequently, we can reduce the bound on $\|h_{T_0^c}\|_1$ to the following expression:

$$\|h_{T_0^c}\|_1 \le \left(\sum_{j=1}^{m}\omega_j - m + 1\right)\|h_{T_0}\|_1 + \sum_{j=1}^{m}(1-\omega_j)\|h_{T_0\cup\widetilde{T}\setminus\widetilde{T}_{j\alpha}}\|_1 + 2\left(\|x_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\omega_j\|x_{\widetilde{T}_j\cap T_0^c}\|_1\right). \quad (26)$$

Next we follow the technique of Candès et al.[2] and sort the coefficients of $h_{T_0^c}$ partitioning T_0^c it into disjoint sets $T_j, j\in\{1,2,\ldots\}$ each of size ak, where $a>1$. That is, T_1 indexes the ak largest in magnitude coefficients of $h_{T_0^c}$, T_2 indexes the second ak largest in magnitude coefficients of $h_{T_0^c}$, and so on. Note that this gives $h_{T_0^c} = \sum_{j\ge1} h_{T_j}$, with

$$\|h_{T_j}\|_2 \le \sqrt{ak}\|h_{T_j}\|_\infty \le (ak)^{-1/2}\|h_{T_{j-1}}\|_1. \quad (27)$$

Let $T_{01} = T_0\cup T_1$, then using (27) and the triangle inequality we have

$$\begin{aligned}\|h_{T_{01}^c}\|_2 &\le \sum_{j\ge2}\|h_{T_j}\|_2 \le (ak)^{-1/2}\sum_{j\ge1}\|h_{T_j}\|_1\\ &\le (ak)^{-1/2}\|h_{T_0^c}\|_1.\end{aligned} \quad (28)$$

Next, consider the feasibility of $x^{\#}$ and x. Both vectors are feasible, so we have $\|Ah\|_2 \le 2\epsilon$ and

$$\begin{aligned}\|Ah_{T_{01}}\|_2 &\le 2\epsilon + \|Ah_{T_{01}^c}\|_2 \le 2\epsilon + \sum_{j\ge2}\|Ah_{T_j}\|_2\\ &\le 2\epsilon + \sqrt{1+\delta_{ak}}\sum_{j\ge2}\|h_{T_j}\|_2.\end{aligned}$$

From (26) and (28) we get

$$\begin{aligned}\|Ah_{T_{01}}\|_2 &\le 2\epsilon + 2\tfrac{\sqrt{1+\delta_{ak}}}{\sqrt{ak}}\left(\|x_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\omega_j\|x_{\widetilde{T}_j\cap T_0^c}\|_1\right)\\ &+ \tfrac{\sqrt{1+\delta_{ak}}}{\sqrt{ak}}\left((\sum_{j=1}^{m}\omega_j - m+1)\|h_{T_0}\|_1 + \sum_{j=1}^{m}(1-\omega_j)\|h_{T_0\cup\widetilde{T}\setminus\widetilde{T}_{j\alpha}}\|_1\right).\end{aligned}$$

Noting that $|T_0\cup\widetilde{T}\setminus\widetilde{T}_{j\alpha}| = (1+\rho_j-2\alpha_j\rho_j)k$,

$$\begin{aligned}\sqrt{1-\delta_{(a+1)k}}\|h_{T_{01}}\|_2 &\le 2\epsilon + 2\tfrac{\sqrt{1+\delta_{ak}}}{\sqrt{ak}}\left(\|x_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\omega_j\|x_{\widetilde{T}_j\cap T_0^c}\|_1\right)\\ &+ \tfrac{\sqrt{1+\delta_{ak}}}{\sqrt{a}}\left((\sum_{j=1}^{m}\omega_j - m+1)\|h_{T_0}\|_2 + \sum_{j=1}^{m}(1-\omega_j)\sqrt{1+\rho_j-2\alpha_j\rho_j}\|h_{T_0\cup\widetilde{T}\setminus\widetilde{T}_{j\alpha}}\|_2\right).\end{aligned}$$

Since for every j we have $\|h_{T_0\cup\widetilde{T}_j\setminus\widetilde{T}_{j\alpha}}\|_2 \le \|h_{T_{01}}\|_2$ and $\|h_{T_0}\|_2 \le \|h_{T_{01}}\|_2$, thus

$$\|h_{T_{01}}\|_2 \le \frac{2\epsilon + 2\frac{\sqrt{1+\delta_{ak}}}{\sqrt{ak}}\left(\|x_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\omega_j\|x_{\widetilde{T}_j\cap T_0^c}\|_1\right)}{\sqrt{1-\delta_{(a+1)k}} - \frac{\sum_{j=1}^{m}\omega_j - m+1+\sum_{j=1}^{m}(1-\omega_j)\sqrt{1+\rho_j-2\alpha_j\rho_j}}{\sqrt{a}}\sqrt{1+\delta_{ak}}}. \quad (29)$$

Finally, using $\|h\|_2 \le \|h_{T_{01}}\|_2 + \|h_{T_{01}^c}\|_2$ and let $\gamma = \sum_{j=1}^{m}\omega_j - m+1+\sum_{j=1}^{m}(1-\omega_j)\sqrt{1+\rho_j-2\alpha_j\rho_j}$, we combine (26), (28) and (29) to get

$$\|h\|_2 \le \frac{2\left(1+\frac{\gamma}{\sqrt{a}}\right)\epsilon + 2\frac{\sqrt{1-\delta_{(a+1)k}}+\sqrt{1+\delta_{ak}}}{\sqrt{ak}}\left(\|x_{\widetilde{T}^c\cap T_0^c}\|_1 + \sum_{j=1}^{m}\omega_j\|x_{\widetilde{T}_j\cap T_0^c}\|_1\right)}{\sqrt{1-\delta_{(a+1)k}} - \frac{\gamma}{\sqrt{a}}\sqrt{1+\delta_{ak}}}, \quad (30)$$

with the condition that the denominator is positive, equivalently $\delta_{ak} + \frac{a}{\gamma^2}\delta_{(a+1)k} < \frac{a}{\gamma^2} - 1$.

ACKNOWLEDGMENTS

The authors would like to thank Rayan Saab for helpful discussions and for sharing his code, which we used for conducting our audio experiments. Both authors were supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development Grant DNOISE II (375142-08). Ö. Yılmaz was also supported in part by an NSERC Discovery Grant.

REFERENCES

[1] Donoho, D., "Compressed sensing.," *IEEE Transactions on Information Theory* **52**(4), 1289–1306 (2006).

[2] Candès, E. J., Romberg, J., and Tao, T., "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics* **59**, 1207–1223 (2006).

[3] Candès, E. J., Romberg, J., and Tao, T., "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory* **52**, 489–509 (2006).

[4] Donoho, D. and Elad, M., "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proceedings of the National Academy of Sciences of the United States of America* **100**(5), 2197–2202 (2003).

[5] Gribonval, R. and Nielsen, M., "Highly sparse representations from dictionaries are unique and independent of the sparseness measure," *Applied and Computational Harmonic Analysis* **22**, 335–355 (May 2007).

[6] Foucart, S. and Lai, M., "Sparsest solutions of underdetermined linear systems via ℓ^q-minimization for $0 < q \leq 1$," *Applied and Computational Harmonic Analysis* **26**(3), 395–407 (2009).

[7] Saab, R., Chartrand, R., and Yilmaz, O., "Stable sparse approximations via nonconvex optimization," in [*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*], 3885–3888 (2008).

[8] Chartrand, R. and Staneva, V., "Restricted isometry properties and nonconvex compressive sensing," *Inverse Problems* **24**(035020) (2008).

[9] Saab, R. and Yilmaz, O., "Sparse recovery by non-convex optimization – instance optimality," *Applied and Computational Harmonic Analysis* **29**, 30–48 (July 2010).

[10] Chartrand, R., "Exact reconstruction of sparse signals via nonconvex minimization," *Signal Processing Letters, IEEE* **14**(10), 707 –710 (2007).

[11] von Borries, R., Miosso, C., and Potes, C., "Compressed sensing using prior information," in [*2nd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMPSAP 2007.*], 121 – 124 (12-14 2007).

[12] Vaswani, N. and Lu, W., "Modified-CS: Modifying compressive sensing for problems with partially known support," *arXiv:0903.5066v4* (2009).

[13] Jacques, L., "A short note on compressed sensing with partially known signal support," *Signal Processing* **90**, 3308 – 3312 (December 2010).

[14] Amin Khajehnejad, M., Xu, W., Salman Avestimehr, A., and Hassibi, B., "Weighted l1 minimization for sparse recovery with prior information," in [*IEEE International Symposium on Information Theory, ISIT 2009*], 483 – 487 (June 2009).

[15] Friedlander, M. P., Mansour, H., Saab, R., and Özgür Yılmaz, "Recovering compressively sampled signals using partial support information," *to appear in the IEEE Transactions on Information Theory* .

[16] Candès, E. J. and Tao, T., "Decoding by linear programming.," *IEEE Transactions on Information Theory* **51**(12), 489–509 (2005).

[17] Cohen, A., Dahmen, W., and DeVore, R., "Compressed sensing and best k-term approximation," *Journal of the American Mathematical Society* **22**(1), 211–231 (2009).

[18] van den Berg, E. and Friedlander, M. P., "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing* **31**(2), 890–912 (2008).

[19] van den Berg, E. and Friedlander, M. P., "SPGL1: A solver for large-scale sparse reconstruction," (June 2007). http://www.cs.ubc.ca/labs/scl/spgl1.

[20] Candès, E. J., Wakin, M. B., and Boyd, S. P., "Enhancing sparsity by reweighted ℓ_1 minimization," *The Journal of Fourier Analysis and Applications* **14**(5), 877–905 (2008).