

Recovery of k-Sparse Signals via Truncated ℓ_1 - ℓ_2 Minimization with Prior Support Information

Shaohua Xie

xsh1209490052@163.com

Shangrao Normal University

Kaihao Liang

Zhongkai University of Agriculture and Engineering

Research Article

Keywords: compressed sensing, ℓ_1 - ℓ_2 minimization, truncated ℓ_1 - ℓ_2 minimization, restricted isometry property, DCA algorithm, k-sparse signal, null space property

Posted Date: October 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11258202/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Recovery of k -Sparse Signals via Truncated ℓ_{1-2} Minimization with Prior Support Information

Shaohua Xie^{*} Kaihao Liang[†]

Abstract

In this paper we study a *truncated* ℓ_{1-2} minimization model for the recovery of sparse signals from a small number of linear measurements, in the situation where partial information about the support of the target signal is available beforehand. We first establish a necessary and sufficient condition under which a *fixed* sparse signal can be recovered by the truncated ℓ_{1-2} *local* minimization model. Building on this result, we derive necessary and sufficient conditions guaranteeing the recovery of *every* k -sparse signal by the truncated ℓ_{1-2} local minimization model and by the corresponding *global* minimization model, respectively. The obtained conditions are shown to be strictly weaker than, and to include as a special case, the existing conditions for the conventional ℓ_{1-2} model; they also refine the conditions recently derived for a related truncated model. Since null-space conditions are not verifiable, we further derive a *restricted isometry* (RIP) condition, $\delta_{2k} < (t+1)/(2k-t-1)$ with $t = |T^C|$, which is sufficient for local recovery and becomes strictly more permissive as the amount of prior information grows. We also treat the extreme case $t = k-1$, where the conditions reduce to a simple rank condition, and we exhibit an explicit matrix for which the truncated model recovers every admissible signal whereas the conventional ℓ_{1-2} model fails, thereby proving that the improvement is strict rather than merely numerical. In addition, we design a difference-of-convex (DCA) based numerical algorithm, complemented by an efficient primal-dual subroutine, to solve the resulting nonconvex optimization problem. Numerical experiments on Gaussian random measurements confirm that the proposed model recovers sparse signals reliably and outperforms the conventional ℓ_{1-2} minimization model, and that the recovery performance improves as the amount of available prior support information increases: with $n = 256$ and $|T^C| = 8$ known active indices, the success rate at sparsity $k = 22$ rises from 38% to 89%, and the number of measurements required for $k = 18$ drops by about 14 (roughly 20% of the sampling budget). The advantage is shown to persist under inaccurate prior information and under measurement noise.

Keywords: compressed sensing, ℓ_{1-2} minimization, truncated ℓ_{1-2} minimization, restricted isometry property, DCA algorithm, k -sparse signal, null space property.

AMS subject classifications: 90C26, 65K10, 49M29.

1 Introduction

Compressed sensing has become a standard paradigm for the efficient acquisition, storage and transmission of signals that admit a sparse representation. Given a measurement matrix $A \in \mathbb{R}^{m \times n}$ with $m \ll n$ and a measurement vector $y = Ax \in \mathbb{R}^m$, the ideal reconstruction problem consists in finding the sparsest solution of the underdetermined linear system $Ax = y$, that is,

$$\arg \min_{x \in \mathbb{R}^n} \|x\|_0 \quad \text{subject to} \quad Ax = y, \quad (1)$$

^{*}School of Mathematics and Physics, Shangrao Normal College, No.401 Zhimin Avenue, Xinzhou District, Shangrao 334001, China. E-mail: xsh1209490052@163.com.

[†]Department of Mathematics, Zhongkai University of Agriculture and Engineering, No.501 Zhongkai Road, Haizhu District, Guangzhou 510225, China. E-mail: karman03@126.com.

where $\|x\|_0$ counts the nonzero entries of x . Since problem (1) is NP-hard [1], a large body of research has been devoted to tractable surrogate models.

Relaxed models. The best known surrogate is the ℓ_1 minimization (basis pursuit) model $\min \|x\|_1$ s.t. $Ax = y$. It is by now classical that the null space property (NSP) of A is necessary and sufficient for ℓ_1 minimization to recover all k -sparse signals [1], while the restricted isometry property (RIP) and coherence conditions provide more tractable sufficient conditions [2, 3, 4, 5]. The sharp RIP bound $\delta_{tk} < \sqrt{(t-1)/t}$ guarantees exact recovery by ℓ_1 minimization.

Because ℓ_1 minimization is not equivalent to ℓ_0 minimization [6, 7], numerous nonconvex relaxations have been proposed. The ℓ_p ($0 < p < 1$) model

$$\arg \min_{x \in \mathbb{R}^n} \|x\|_p \quad \text{subject to} \quad Ax = y, \quad (2)$$

with $\|x\|_p = (\sum_i |x_i|^p)^{1/p}$, admits an NSP characterization [8] and sharp RIP bounds [9]. Another popular scale-invariant alternative is the ratio model $\min \|x\|_1 / \|x\|_2$ s.t. $Ax = y$, for which a strong null space property ensures that every sparse signal is a local minimizer [10, 11].

A particularly successful relaxation, and the starting point of the present work, is the ℓ_{1-2} minimization model

$$\arg \min_{x \in \mathbb{R}^n} \|x\|_1 - \|x\|_2 \quad \text{subject to} \quad Ax = y. \quad (3)$$

Model (3) has attracted considerable attention [12, 13]; it has been proved to recover sparse signals under weaker conditions than ℓ_1 minimization [14, 15]. In particular, [15] provides null-space characterizations of the recovery of a fixed vector by the local minimization model, of all k -sparse vectors by the local minimization model, and of all k -sparse vectors by the global minimization model; these are the results that the present paper extends to the truncated setting. Several variants have been investigated, including the truncated $\ell_1 - \ell_2$ model [17], the $\ell_1 - a\ell_2$ model [18], the $\ell_p - a\ell_q$ model [19], the ℓ_{1-2} analysis model [20], the $\ell_p - w\ell_q$ model [21], as well as extensions to one-bit compressed sensing [22, 23].

Prior support information. In many applications one possesses *a priori* knowledge that a certain subset of indices is guaranteed to be active in the target signal. Writing $T \subseteq [n] := \{1, \dots, n\}$ for the complementary index set, this prior information is encoded as

$$T^C \subseteq \text{supp}(x), \quad (4)$$

i.e., every entry of x outside T is known to be nonzero. Such knowledge can be exploited to weaken the penalty and hence to improve recoverability. A well-known instance is the *truncated basis pursuit* (Truncated BP) model

$$\arg \min_{x \in \mathbb{R}^n} \|x_T\|_1 \quad \text{subject to} \quad Ax = y, \quad (5)$$

where x_T denotes the vector obtained from x by keeping the entries indexed by T and setting all remaining entries to zero. It was shown in [16] that (5) recovers sparse signals better than plain ℓ_1 minimization.

The proposed model. Combining the idea of truncation with the ℓ_{1-2} penalty, we propose the *truncated ℓ_{1-2} minimization* model

$$\arg \min_{x \in \mathbb{R}^n} \|x_T\|_1 - \|x_T\|_2 \quad \text{subject to} \quad Ax = y, \quad (6)$$

where the target signal is assumed to satisfy the prior information (4). Two boundary cases are worth emphasizing:

- if $T^C = \emptyset$ (no prior information), then $T = [n]$ and (6) reduces exactly to the conventional ℓ_{1-2} model (3);
- the larger the known active set T^C , the fewer indices remain penalized, so that the model becomes progressively less restrictive.

A natural question is therefore whether the truncated model (6) inherits, and indeed improves upon, the recovery guarantees of (3). The present paper gives an affirmative answer, both theoretically and numerically.

Contributions and outline. The main contributions of this paper are fourfold.

- (1) *Sharp theory.* We establish a necessary and sufficient condition for a fixed sparse signal to be recovered by the truncated ℓ_{1-2} *local* minimization model (Theorem 2.4), and then derive necessary and sufficient conditions for *all* k -sparse signals satisfying (4) to be recovered by the local model (Theorem 2.10) and by the *global* model (Theorem 2.18). We carefully compare our conditions with those of [17] and [15] and show that ours are weaker and more general (Remarks 2.12, 2.13 and 2.7).
- (2) *Verifiable and strict guarantees.* Because null-space conditions are not checkable in practice, we derive a sufficient *RIP* condition $\delta_{2k} < (t+1)/(2k-t-1)$, $t = |T^C|$, which improves monotonically with the amount of prior information (Theorem 2.23). We settle the boundary case $t = k-1$ (Corollary 2.25) and, by an explicit matrix example, prove that the improvement over the conventional ℓ_{1-2} model is *strict* (Theorem 2.26).
- (3) *Algorithm.* We design a DCA-based algorithm for solving (6), together with an efficient primal–dual subroutine for the convex subproblem, and we provide closed-form expressions for all proximity operators involved.
- (4) *Numerics.* We validate the model on Gaussian random measurements. The results show that (6) recovers sparse signals reliably and consistently outperforms (3), and that its advantage grows with the amount of available prior support information.

The remainder of the paper is organized as follows. Section 2 contains the recovery theory: Subsection 2.1 treats local minimization, Subsection 2.2 treats global minimization, and Subsection 2.3 derives the verifiable conditions and the strict-improvement result. Section 3 presents the numerical algorithm. Section 4 reports the numerical experiments, and Section 5 concludes the paper with a discussion of future work.

Notation. For $x \in \mathbb{R}^n$, $\text{supp}(x) = \{i : x_i \neq 0\}$ is the support of x , and x is called k -sparse if $|\text{supp}(x)| \leq k$. For $\Omega \subseteq [n]$, $x_\Omega \in \mathbb{R}^n$ denotes the vector with $(x_\Omega)_i = x_i$ for $i \in \Omega$ and $(x_\Omega)_i = 0$ otherwise; Ω^C is the complement of Ω in $[n]$. We write $\ker A = \{h \in \mathbb{R}^n : Ah = 0\}$. Given $h \in \mathbb{R}^n$ and an integer s , $h_{\max(s)}$ denotes the vector obtained from h by retaining its s largest entries in magnitude and setting all other entries to zero, so that $\text{supp}(h_{\max(s)})$ is the set of the s largest-magnitude indices. The standard inner product is written $\langle \cdot, \cdot \rangle$, and $\text{sign}(\cdot)$ is the sign function. We write $\subset$ for $\subseteq$, i.e. set inclusion is not required to be proper.

Throughout, $T \subseteq [n]$ is a fixed index set encoding the prior information (4), and $S = \text{supp}(x)$. We always assume $|T^C| \leq k - 1$, so that the index set I introduced below is nonempty; the excluded case $|T^C| = k$ is discussed in Remark 2.11.

2 Main Results

In this section we characterize when a sparse signal satisfying the prior information (4) can be recovered by model (6). We distinguish between *local* recovery, i.e., x is a strict local minimizer of the objective along the feasible affine subspace, and *global* recovery, i.e., x is the unique global minimizer.

Definition 2.1 (Truncated ℓ_{1-2} local minimization). *A vector $x \in \mathbb{R}^n$ with $T^C \subset \text{supp}(x)$ can be recovered from $Ax = y$ via truncated ℓ_{1-2} local minimization if there exists a constant $r > 0$ such that*

$$f_T(x, th) := \|(x + th)_T\|_1 - \|(x + th)_T\|_2 - (\|x_T\|_1 - \|x_T\|_2) > 0 \quad (7)$$

holds for all $h \in \ker A$ with $\|h\|_2 = 1$ and all $0 < t < r$.

Definition 2.2 (Truncated ℓ_{1-2} global minimization). *A vector $x \in \mathbb{R}^n$ with $T^C \subset \text{supp}(x)$ can be recovered from $Ax = y$ via truncated ℓ_{1-2} global minimization if*

$$f_T(x, h) := \|(x + h)_T\|_1 - \|(x + h)_T\|_2 - (\|x_T\|_1 - \|x_T\|_2) > 0 \quad (8)$$

holds for all $h \in \ker A \setminus \{0\}$.

Remark 2.3. *It follows directly from Definitions 2.1 and 2.2 that global recovery implies local recovery, whereas the converse implication fails in general. Consequently, the conditions for global recovery are expected to be stronger than those for local recovery; this will be made precise in Theorems 2.10 and 2.18.*

2.1 Local Minimization

We begin with a necessary and sufficient condition for a *fixed* sparse signal to be recovered by the truncated ℓ_{1-2} local minimization model.

Theorem 2.4. *Let $A \in \mathbb{R}^{m \times n}$ be a measurement matrix and let $T \subseteq [n]$ be a fixed index set. A vector $x \in \mathbb{R}^n$ with $T^C \subset \text{supp}(x)$ and $x_T \neq 0$ can be recovered from $Ax = y$ via truncated ℓ_{1-2} local minimization if and only if*

$$\sum_{i \in T \cap S} \text{sign}(x_i) h_i + \|h_{S^C}\|_1 - \left\langle h_T, \frac{x_T}{\|x_T\|_2} \right\rangle > 0 \quad (9)$$

holds for all $h \in \ker A$ with $\|h\|_2 = 1$, where $S = \text{supp}(x)$.

Proof. Necessity. Suppose that x can be recovered via truncated ℓ_{1-2} local minimization, so that (7) holds for all $0 < t < r$. Choose any $t \in (0, \min\{r, \min_{i \in S} |x_i|\})$. Such a choice guarantees $\text{sign}(x_i + th_i) = \text{sign}(x_i)$ for every $i \in S$. Since $T^C \subset S$ implies $S^C \subset T$, we have

$T \cap S^C = S^C$, and therefore

$$\begin{aligned}
f_T(x, th) &= \|(x + th)_{T \cap S}\|_1 + \|(x + th)_{T \cap S^C}\|_1 - \|x_T\|_1 + (\|x_T\|_2 - \|(x + th)_T\|_2) \\
&= \sum_{i \in S \cap T} \text{sign}(x_i)(x_i + th_i) + t \|h_{S^C}\|_1 - \|x_T\|_1 + \frac{\|x_T\|_2^2 - \|(x + th)_T\|_2^2}{\|x_T\|_2 + \|(x + th)_T\|_2} \\
&= \sum_{i \in S \cap T} \text{sign}(x_i) th_i + t \|h_{S^C}\|_1 - \frac{t^2 \|h_T\|_2^2 + 2t \langle h_T, x_T \rangle}{\|x_T\|_2 + \|(x + th)_T\|_2} \\
&= t \left(\sum_{i \in S \cap T} \text{sign}(x_i) h_i + \|h_{S^C}\|_1 - \left\langle h_T, \frac{x_T}{\|x_T\|_2} \right\rangle - g(t) \right), \tag{10}
\end{aligned}$$

where

$$g(t) = \frac{t \|h_T\|_2^2 + 2 \langle h_T, x_T \rangle}{\|x_T\|_2 + \|(x + th)_T\|_2} - \left\langle h_T, \frac{x_T}{\|x_T\|_2} \right\rangle.$$

Rearranging $g(t)$ gives

$$\begin{aligned}
g(t) &= \frac{\|x_T\|_2 (t \|h_T\|_2^2 + 2 \langle h_T, x_T \rangle) - (\|x_T\|_2 + \|(x + th)_T\|_2) \langle h_T, x_T \rangle}{\|x_T\|_2 (\|x_T\|_2 + \|(x + th)_T\|_2)} \\
&= \frac{t \|x_T\|_2 \|h_T\|_2^2 + \langle h_T, x_T \rangle (\|x_T\|_2 - \|(x + th)_T\|_2)}{\|x_T\|_2 (\|x_T\|_2 + \|(x + th)_T\|_2)} \\
&\geq \frac{t \|x_T\|_2 \|h_T\|_2^2 - |\langle h_T, x_T \rangle| |\|x_T\|_2 - \|(x + th)_T\|_2|}{\|x_T\|_2 (\|x_T\|_2 + \|(x + th)_T\|_2)} \\
&\geq \frac{t \|x_T\|_2 \|h_T\|_2^2 - t \|x_T\|_2 \|h_T\|_2^2}{\|x_T\|_2 (\|x_T\|_2 + \|(x + th)_T\|_2)} = 0, \tag{11}
\end{aligned}$$

where we used $|\|x_T\|_2 - \|(x + th)_T\|_2| \leq t \|h_T\|_2$ and $|\langle h_T, x_T \rangle| \leq \|h_T\|_2 \|x_T\|_2$ together with $\|h_T\|_2 \leq \|h\|_2 = 1$. Combining (7), (10) and (11) and dividing by $t > 0$ yields

$$\sum_{i \in S \cap T} \text{sign}(x_i) h_i + \|h_{S^C}\|_1 - \left\langle h_T, \frac{x_T}{\|x_T\|_2} \right\rangle > g(t) \geq 0,$$

which is (9).

Sufficiency. Assume that (9) holds. Since the unit sphere of $\ker A$ is compact and the left-hand side of (9) is continuous in h , we may define

$$\beta := \min_{\substack{h \in \ker A \\ \|h\|_2 = 1}} \left(\sum_{i \in T \cap S} \text{sign}(x_i) h_i + \|h_{S^C}\|_1 - \left\langle h_T, \frac{x_T}{\|x_T\|_2} \right\rangle \right) > 0. \tag{12}$$

On the other hand, estimating $g(t)$ from above in the same way as in (11) gives

$$g(t) \leq \frac{2t \|h_T\|_2^2}{\|x_T\|_2 + \|(x + th)_T\|_2} \leq \frac{2t}{\|x_T\|_2}. \tag{13}$$

Hence, whenever $0 < t < \min\{\min_{i \in S} |x_i|, \beta \|x_T\|_2 / 2\}$, it follows from (10), (12) and (13) that

$$f_T(x, th) = t \left(\sum_{i \in S \cap T} \text{sign}(x_i) h_i + \|h_{S^C}\|_1 - \left\langle h_T, \frac{x_T}{\|x_T\|_2} \right\rangle - g(t) \right) \geq t \left(\beta - \frac{2t}{\|x_T\|_2} \right) > 0.$$

By Definition 2.1, x can be recovered via truncated ℓ_{1-2} local minimization. $\square$

Remark 2.5 (Two boundary aspects of Theorem 2.4). (a) Reduction to [15]. When $T = [n]$ the criterion (9) is exactly condition (3) of [15]. Our proof of necessity is, however, shorter than the one given there: instead of introducing the auxiliary quantity $\theta_x = \max_{h \in \ker A, \|h\|_2=1} \langle x, h \rangle / \|x\|_2 < 1$ (cf. [15, Lemma 11]) and estimating the remainder τ from below, we bound the remainder $g(t)$ directly by 0 in (11).

(b) The degenerate case $x_T = 0$. The expression (9) presupposes $x_T \neq 0$. If $x_T = 0$ then, by the prior information (4), $\text{supp}(x) = T^C$, and (7) gives directly, for every $h \in \ker A$ and $t > 0$,

$$f_T(x, th) = \|(th)_T\|_1 - \|(th)_T\|_2 = t(\|h_T\|_1 - \|h_T\|_2).$$

Hence such an x is recoverable if and only if $\|h_T\|_1 > \|h_T\|_2$, i.e. $|\text{supp}(h_T)| \geq 2$, for every $h \in \ker A \setminus \{0\}$. This condition is implied by condition (ii) of Theorem 2.10: if some $h \neq 0$ had $|\text{supp}(h_T)| \leq 1$, then $\text{supp}(h) \subseteq T^C \cup \{j_0\}$ for some $j_0 \in T$, and the columns of $A_{T^C \cup \{j_0\}}$ (or of A_{T^C} , if $h_T = 0$) would be linearly dependent.

Condition (9) still depends on the unknown signal x . In order to obtain a condition that guarantees recovery of *all* k -sparse signals satisfying the prior information (4), we need the following equivalence, which removes the dependence on the signs and on the magnitude distribution of x .

Lemma 2.6. Let $A \in \mathbb{R}^{m \times n}$ be a measurement matrix and let $T \subseteq [n]$ be a fixed index set. Then the following four statements are equivalent:

- (i) $\sum_{i \in T \cap S} \text{sign}(x_i)h_i + \|h_{S^C}\|_1 - \langle h_T, \frac{x_T}{\|x_T\|_2} \rangle > 0$ for all $0 \neq x \in \mathbb{R}^n$, $h \in \ker A$ with $\|h\|_2 = 1$, where $S = \text{supp}(x)$ satisfies $|S| \leq k$ and $T^C \subset S$;
- (ii) $\|h_{S^C}\|_1 - \|h_{T \cap S}\|_1 + \sum_{i \in T \cap S} \frac{|x_i|}{\|x_T\|_2} |h_i| > 0$ for all $0 \neq x \in \mathbb{R}^n$, $h \in \ker A$ with $\|h\|_2 = 1$, where $S = \text{supp}(x)$ satisfies $|S| \leq k$ and $T^C \subset S$;
- (iii) the same inequality as in (ii) but required only for $|S| = k$;
- (iv) $\left\| h_{(I \cup T^C)^C} \right\|_1 - \|h_I\|_1 + \sum_{i \in I} \frac{|x_i|}{\|x_T\|_2} |h_i| > 0$ for all $0 \neq x \in \mathbb{R}^n$, $h \in \ker A$ with $\|h\|_2 = 1$ and $\text{supp}(x) = I \cup T^C$, where $I \subseteq T$ satisfies $\text{supp}((h_T)_{\max(k-|T^C|)}) \subseteq I$ and $|I| = k - |T^C|$.

Proof. (i) $\Rightarrow$ (ii). Fix admissible x, h and set $S = \text{supp}(x)$. Define $x^* \in \mathbb{R}^n$ by

$$x_i^* := \begin{cases} -\text{sign}(h_i)|x_i|, & i \in S \cap T \cap \text{supp}(h), \\ x_i, & \text{otherwise.} \end{cases} \quad (14)$$

Then $\text{supp}(x^*) = \text{supp}(x) = S$, $|x_i^*| = |x_i|$, and

$$\begin{aligned} \|h_{S^C}\|_1 - \|h_{T \cap S}\|_1 + \sum_{i \in T \cap S} \frac{|x_i|}{\|x_T\|_2} |h_i| &= \|h_{S^C}\|_1 - \sum_{i \in S \cap T} \left(1 - \frac{|x_i|}{\|x_T\|_2}\right) |h_i| \\ &= \|h_{S^C}\|_1 + \sum_{i \in S \cap T} \left(1 - \frac{|x_i^*|}{\|x_T^*\|_2}\right) \text{sign}(x_i^*) h_i \\ &= \sum_{i \in S \cap T} \text{sign}(x_i^*) h_i + \|h_{S^C}\|_1 - \left\langle h_T, \frac{x_T^*}{\|x_T^*\|_2} \right\rangle > 0, \end{aligned}$$

where the last inequality is exactly statement (i) applied to x^* .

(ii) $\Rightarrow$ (iii). Immediate, since (iii) is the special case $|S| = k$ of (ii).

(iii) $\Rightarrow$ (i). Let x, h be as in (i) with $|S| < k$. Choose $S_1 \supset S$ with $|S_1| = k$ and define x^* by

$$x_i^* := \begin{cases} x_i, & i \in S, \\ 1, & i \in S_1 \setminus S, \\ 0, & \text{otherwise.} \end{cases}$$

Then $\text{supp}(x^*) = S_1$, $\|h_{S^C}\|_1 \geq \|h_{S_1^C}\|_1$ (because $S \subset S_1$), and, since $\|x_T^*\|_2^2 = \|x_T\|_2^2 + |T \cap (S_1 \setminus S)| \geq \|x_T\|_2^2$ while $|x_i^*| = |x_i|$ for $i \in S$,

$$1 - \frac{|x_i|}{\|x_T\|_2} \leq 1 - \frac{|x_i^*|}{\|x_T^*\|_2} \quad (i \in T \cap S).$$

Note that $1 - \frac{|x_i|}{\|x_T\|_2} \geq 0$, since $|x_i| \leq \|x_T\|_2$. Consequently,

$$\begin{aligned} \sum_{i \in S \cap T} \text{sign}(x_i) h_i + \|h_{S^C}\|_1 - \left\langle h_T, \frac{x_T}{\|x_T\|_2} \right\rangle &= \|h_{S^C}\|_1 + \sum_{i \in S \cap T} \left(1 - \frac{|x_i|}{\|x_T\|_2}\right) \text{sign}(x_i) h_i \\ &\geq \|h_{S^C}\|_1 - \sum_{i \in S \cap T} \left(1 - \frac{|x_i|}{\|x_T\|_2}\right) |h_i| \\ &\geq \|h_{S_1^C}\|_1 - \sum_{i \in S \cap T} \left(1 - \frac{|x_i^*|}{\|x_T^*\|_2}\right) |h_i| \\ &\geq \|h_{S_1^C}\|_1 - \|h_{T \cap S_1}\|_1 + \sum_{i \in T \cap S_1} \frac{|x_i^*|}{\|x_T^*\|_2} |h_i| > 0, \end{aligned}$$

where the last inequality follows from (iii).

(iii) $\Rightarrow$ (iv). Take $S = I \cup T^C$, so that $|S| = k$ and $T \cap S = I$. Then (iv) is precisely (iii).

(iv) $\Rightarrow$ (iii). Let x, h be as in (iii) with $|S| = k$. Write $\text{supp}(x_T) = \{i_1, \dots, i_{k-|T^C|}\}$ with $0 < |x_{i_1}| \leq \dots \leq |x_{i_{k-|T^C|}}|$, and let $I = \{j_1, \dots, j_{k-|T^C|}\} \subseteq T$ be an index set such that $|(h_T)_{j_1}| \geq \dots \geq |(h_T)_{j_{k-|T^C|}}|$ and $\text{supp}((h_T)_{\max(k-|T^C|)}) \subseteq I$. Define x^* by

$$x_l^* := \begin{cases} x_l, & l \in T^C, \\ x_{i_t}, & l = j_t \in I, \\ 0, & \text{otherwise.} \end{cases}$$

Then $\text{supp}(x^*) = I \cup T^C$ and $\|x_T^*\|_2 = \|x_T\|_2$. Next, $T^C \subset S$ gives $S^C \subseteq T$ with $|S^C| = n - k$, while $(I \cup T^C)^C = T \setminus I$ also has cardinality $n - k$ and consists of the $n - k$ entries of h_T of *smallest* magnitude; hence $\|h_{S^C}\|_1 \geq \|h_{(I \cup T^C)^C}\|_1$. Moreover, by the rearrangement inequality (which pairs the increasing sequence $|x_{i_1}| \leq \dots \leq |x_{i_{k-t}}|$ with the decreasing sequence $|(h_T)_{j_1}| \geq \dots \geq |(h_T)_{j_{k-t}}|$ so as to *minimize* the sum of products),

$$\sum_{i \in S \cap T} \left(1 - \frac{|x_i|}{\|x_T\|_2}\right) |h_i| \leq \sum_{i \in I} \left(1 - \frac{|x_i^*|}{\|x_T^*\|_2}\right) |h_i|.$$

Therefore,

$$\begin{aligned}
\|h_{S^C}\|_1 - \|h_{T \cap S}\|_1 + \sum_{i \in T \cap S} \frac{|x_i|}{\|x_T\|_2} |h_i| &= \|h_{S^C}\|_1 - \sum_{i \in S \cap T} \left(1 - \frac{|x_i|}{\|x_T\|_2}\right) |h_i| \\
&\geq \left\| h_{(I \cup T^C)^C} \right\|_1 - \sum_{i \in I} \left(1 - \frac{|x_i^*|}{\|x_T^*\|_2}\right) |h_i| \\
&= \left\| h_{(I \cup T^C)^C} \right\|_1 - \|h_I\|_1 + \sum_{i \in I} \frac{|x_i^*|}{\|x_T^*\|_2} |h_i| > 0,
\end{aligned}$$

which is (iii). The proof is complete. $\square$

Remark 2.7 (Relation to [15, Lemma 12]). *Lemma 2.6 is the truncated counterpart of [15, Lemma 12], which establishes the equivalence of four conditions a)–d) for the conventional model. The correspondence is (i)–(iv) $\leftrightarrow$ a)–d): when $T = [n]$ we have $t = 0$, $I = \mathcal{T} := \text{supp}(h_{\max(k)})$ and $(I \cup T^C)^C = \mathcal{T}^C$, so that statements (i) and (iv) reduce exactly to conditions a) and d) of [15]. The elementary identity (15), which is what allows one to pass from (iv) to the x -free condition of Theorem 2.10, plays in the truncated setting the same role as [15, Remark 13] plays in the conventional one.*

Remark 2.8. *If $\text{supp}(x) = I \cup T^C$ with I as in Lemma 2.6, then $\text{supp}(x_T) = I$, hence $\|x_T\|_2 = \|x_I\|_2$ and $\sum_{i \in I} \frac{|x_i|^2}{\|x_T\|_2^2} = 1$.*

Remark 2.9. *In general the index set I in Lemma 2.6(iv) is not unique (ties in the magnitudes of h_T may occur). It suffices that one admissible set I satisfies the stated inequality; all statements below are understood in this sense.*

We shall repeatedly use the following elementary fact: for every $y \in \mathbb{R}^p$ with $y \geq 0$,

$$\inf \{ \langle u, y \rangle : u \in \mathbb{R}^p, u \geq 0, \|u\|_2 = 1 \} = \min_{i \in [p]} y_i, \quad (15)$$

the infimum being attained at a unit vector supported on an index where y is minimal.

Combining Theorem 2.4 with Lemma 2.6, we now characterize uniform recovery of *all* k -sparse signals.

Theorem 2.10 (Local recovery: necessary and sufficient condition). *Let $A \in \mathbb{R}^{m \times n}$. Every nonzero k -sparse vector $x \in \mathbb{R}^n$ with $T^C \subset \text{supp}(x)$ can be recovered from $Ax = y$ via truncated ℓ_{1-2} local minimization if and only if the following two conditions hold:*

- (i) $\left\| h_{(I \cup T^C)^C} \right\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| \geq 0$ for all $h \in \ker A$ with $\|h\|_2 = 1$, where $I \subseteq T$ satisfies $\text{supp}((h_T)_{\max(k-|T^C|)}) \subseteq I$ and $|I| = k - |T^C|$ (see Remark 2.9 regarding non-uniqueness of I);
- (ii) for every $j_0 \in [n] \setminus T^C$, the columns of $A_{T^C \cup \{j_0\}}$ are linearly independent.

Proof. Necessity. Assume that every nonzero k -sparse vector satisfying the prior information can be recovered.

- (i) Let $h \in \ker A$ with $\|h\|_2 = 1$ and let I be as in the statement. By Remark 2.8 and (15),

$$\inf_{x: \text{supp}(x) = I \cup T^C} \sum_{i \in I} \frac{|x_i|}{\|x_T\|_2} |h_i| = \inf_{x: \text{supp}(x) = I \cup T^C} \left\langle \frac{|x_I|}{\|x_I\|_2}, |h_I| \right\rangle = \min_{i \in I} |h_i|.$$

Hence, by Theorem 2.4 and Lemma 2.6,

$$\left\| h_{(I \cup T^C)^C} \right\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| = \inf_{x: \text{supp}(x) = I \cup T^C} \left(\left\| h_{(I \cup T^C)^C} \right\|_1 - \|h_I\|_1 + \sum_{i \in I} \frac{|x_i|}{\|x_T\|_2} |h_i| \right) \geq 0,$$

which is (i).

(ii) Suppose there exists $j_0 \in [n] \setminus T^C$ such that the columns of $A_{T^C \cup \{j_0\}}$ are linearly dependent. Then there is $h \neq 0$ with $\text{supp}(h) \subseteq T^C \cup \{j_0\}$ and $Ah = 0$. For any x with $\text{supp}(x) = T^C \cup \{j_0\}$ we then have $(x + h)_T = x_T$, whence

$$\|(x + h)_T\|_1 - \|(x + h)_T\|_2 - (\|x_T\|_1 - \|x_T\|_2) = 0,$$

so that such an x cannot be recovered, a contradiction.

Sufficiency. Assume (i) and (ii). By Theorem 2.4 and Lemma 2.6 it suffices to prove statement (iv) of Lemma 2.6. Let $0 \neq x \in \mathbb{R}^n$ and $h \in \ker A$, $\|h\|_2 = 1$, with $\text{supp}(x) = I \cup T^C$, where $I = \text{supp}((h_T)_{\max(k-|T^C|)}) \subseteq T$. By Remark 2.8, $\text{supp}(x_T) = I$ and $\|x_T\|_2 = \|x_I\|_2$. We distinguish two cases.

Case 1: $\|x_I\|_0 > 1$. Then $\|x_I\|_1 / \|x_T\|_2 > 1$. We split further according to whether h_I vanishes.

- If $h_I \neq 0$ and $\min_{i \in I} |h_i| > 0$, then

$$\begin{aligned} \left\| h_{(I \cup T^C)^c} \right\|_1 - \|h_I\|_1 + \sum_{i \in I} \frac{|x_i|}{\|x_T\|_2} |h_i| &\geq \left\| h_{(I \cup T^C)^c} \right\|_1 - \|h_I\|_1 + \frac{\|x_I\|_1}{\|x_T\|_2} \min_{i \in I} |h_i| \\ &> \left\| h_{(I \cup T^C)^c} \right\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| \geq 0. \end{aligned}$$

- If $h_I \neq 0$ and $\min_{i \in I} |h_i| = 0$, then $\max_{i \in I} |h_i| > 0$ and the sum $\sum_{i \in I} \frac{|x_i|}{\|x_T\|_2} |h_i|$ is strictly positive, hence

$$\left\| h_{(I \cup T^C)^c} \right\|_1 - \|h_I\|_1 + \sum_{i \in I} \frac{|x_i|}{\|x_T\|_2} |h_i| > \left\| h_{(I \cup T^C)^c} \right\|_1 - \|h_I\|_1 = \left\| h_{(I \cup T^C)^c} \right\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| \geq 0.$$

- If $h_I = 0$, then condition (ii) implies that the columns of A_{T^C} are linearly independent, so h cannot be supported in T^C ; since $h_I = 0$, this forces $h_{(I \cup T^C)^c} \neq 0$ and therefore

$$\left\| h_{(I \cup T^C)^c} \right\|_1 - \|h_I\|_1 + \sum_{i \in I} \frac{|x_i|}{\|x_T\|_2} |h_i| = \left\| h_{(I \cup T^C)^c} \right\|_1 > 0.$$

Case 2: $\|x_I\|_0 = 1$. Write $\text{supp}(x) = \{j_0\} \cup T^C =: S$. Since the columns of $A_{T^C \cup \{j_0\}}$ are linearly independent and $h \neq 0$, we have $\left\| h_{(T^C \cup \{j_0\})^c} \right\|_1 > 0$. Moreover x_T is supported on the single index j_0 , so that $\frac{x_T}{\|x_T\|_2} = \text{sign}(x_{j_0})e_{j_0}$ and therefore

$$\sum_{i \in T \cap S} \text{sign}(x_i)h_i + \|h_{S^c}\|_1 - \left\langle h_T, \frac{x_T}{\|x_T\|_2} \right\rangle = \text{sign}(x_{j_0})h_{j_0} + \|h_{S^c}\|_1 - \text{sign}(x_{j_0})h_{j_0} = \|h_{S^c}\|_1 > 0.$$

By Theorem 2.4, the conclusion follows. $\square$

Remark 2.11 (The degenerate case $k = |T^C|$). If $k = |T^C| =: t$ then $k - t = 0$ and the index set I occurring in Theorems 2.10 and 2.18 is empty; condition (i) is then vacuous. Every admissible signal satisfies $\text{supp}(x) = T^C$, i.e. $x_T = 0$, so that Remark 2.5(b) applies and local recovery is equivalent to condition (ii) alone. To avoid this degenerate situation we assume throughout that $|T^C| \leq k - 1$.

Remark 2.12 (Relation to the truncated $\ell_1 - \ell_2$ model). When $|T^C| = t$, model (6) coincides in form with model (1) of [17], and Theorem 2.10 is the counterpart of Theorem 4 therein; our hypotheses are weaker. To avoid a clash of notation, we denote the index sets used in [17] by calligraphic letters: $\mathcal{T} := \text{supp}(h_{\max(k)})$ and $\mathcal{T}_t := \text{supp}((h_{\max(k)})_{\min(t)})$. Then $|\mathcal{T}^C| = n - k$, $|(I \cup T^C)^C| = n - k$ and $|I| = |\mathcal{T}_t \cap \mathcal{T}| = k - t$. By the definition of I , and because h_I consists of the $k - t$ largest-magnitude entries of $h_{\mathcal{T}}$,

$$\|h_I\|_1 - \min_{i \in I} |h_i| \leq \|h_{\mathcal{T}_t \cap \mathcal{T}}\|_1 - \min_{i \in \mathcal{T}_t \cap \mathcal{T}} |h_i|,$$

while $|\mathcal{T}^C| = |(I \cup T^C)^C| = n - k$, together with the maximality of $\mathcal{T}$, gives $\|h_{(I \cup T^C)^C}\|_1 \geq \|h_{\mathcal{T}^C}\|_1$. Therefore

$$\|h_{(I \cup T^C)^C}\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| \geq \|h_{\mathcal{T}^C}\|_1 - \|h_{\mathcal{T}_t \cap \mathcal{T}}\|_1 + \min_{i \in \mathcal{T}_t \cap \mathcal{T}} |h_i|,$$

i.e., condition (i) above is weaker than condition (i) of Theorem 4 in [17]. Condition (ii) above is likewise weaker than condition (ii) of Theorem 4 in [17]. Hence the hypotheses of Theorem 2.10 are weaker than those of the corresponding result in [17].

Remark 2.13 (Relation to the conventional ℓ_{1-2} model). The results of this section contain the characterization of the conventional ℓ_{1-2} model (3) obtained in [15] as a special case. To avoid a clash of notation we again write $\mathcal{T} := \text{supp}(h_{\max(k)})$ for the index set used in [15] (cf. Remark 2.12), so that $|\mathcal{T}| = k$ and conditions (4) and (5) of [15] read

$$G_2(\mathcal{T}, h) = \|h_{\mathcal{T}^C}\|_1 - \|h_{\mathcal{T}}\|_1 + \min_{i \in \mathcal{T}} |h_i| \geq 0, \quad G_1(\mathcal{T}, h) = \|h_{\mathcal{T}^C}\|_1 - \|h_{\mathcal{T}^C}\|_2 - (\|h_{\mathcal{T}}\|_1 - \|h_{\mathcal{T}}\|_2) > 0.$$

Exact reduction. If $T^C = \emptyset$, then $T = [n]$, the set I appearing in Theorems 2.10 and 2.18 is exactly $\mathcal{T}$, and $(I \cup T^C)^C = \mathcal{T}^C$. Hence condition (i) of Theorem 2.10 is precisely condition (4) of [15], and conditions (i)–(ii) of Theorem 2.18 below are precisely conditions (4)–(5) of [15]. Therefore Theorem 2.10 and Theorem 2.18 reduce to Theorem 5 and Theorem 6 of [15] respectively. The only difference is that condition (ii) of Theorem 2.10 then becomes the requirement that A have no zero column; this is equivalent to the standing assumption $\|h\|_0 > 1$ ($h \in \ker A \setminus \{0\}$) imposed in [15], and we prefer to state it explicitly.

Strict weakening for $T^C \neq \emptyset$. Let $|T^C| = t > 0$. Then condition (i) of Theorem 2.10 is strictly weaker than condition (4) of [15]. Indeed, $|(I \cup T^C)^C| = |\mathcal{T}^C| = n - k$, and since $\mathcal{T}^C$ consists of the $n - k$ entries of h of smallest absolute value, any other index set of cardinality $n - k$ has at least as large an ℓ_1 -norm, i.e. $\|h_{(I \cup T^C)^C}\|_1 \geq \|h_{\mathcal{T}^C}\|_1$. On the other hand, $\|h_I\|_1 - \min_{i \in I} |h_i|$ is the sum of the $|I| - 1 = k - t - 1$ largest entries of h_I , whereas $\|h_{\mathcal{T}}\|_1 - \min_{i \in \mathcal{T}} |h_i|$ is the sum of the $k - 1$ largest entries of h ; since $I \subseteq [n]$ and $k - t - 1 \leq k - 1$, the former does not exceed the latter. Combining the two displays,

$$\|h_{(I \cup T^C)^C}\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| \geq \|h_{\mathcal{T}^C}\|_1 - \|h_{\mathcal{T}}\|_1 + \min_{i \in \mathcal{T}} |h_i| = G_2(\mathcal{T}, h).$$

This quantifies precisely the benefit of prior support information: enlarging T^C relaxes the recovery condition, and the relaxation is monotone in $t = |T^C|$.

2.2 Global Minimization

We now turn to the global minimization model. The analysis is more delicate because arbitrarily large perturbations $h \in \ker A$ must be handled. For fixed nonzero $x \in \mathbb{R}^n$ and $h \in \mathbb{R}^n$ with $T^C \subset \text{supp}(x)$, set $S = \text{supp}(x)$, $S_1 = \{i \in S : |x_i| > |h_i|\}$, $S_2 = S \setminus S_1$, and define

$$\bar{x} = x_{S_1} - h_{S_2}, \quad x_i^* := \begin{cases} -\text{sign}(h_i)|x_i|, & h_i \neq 0, i \in S_1, \\ \bar{x}_i, & \text{otherwise.} \end{cases} \quad (16)$$

Thus $|x_i^*| = |\bar{x}_i|$ for every i (both sides equal $|x_i|$ when $i \in S_1$), and hence

$$\|x_T^*\|_2 = \|\bar{x}_T\|_2. \quad (17)$$

Moreover $T^C \subset \text{supp}(\bar{x})$, $T^C \subset \text{supp}(x^*)$ and $\text{supp}(\bar{x}) = \text{supp}(x^*) = S$; indeed, for $i \in S_2$ we have $|x_i| \leq |h_i|$ with $x_i \neq 0$, hence $h_i \neq 0$ and $\bar{x}_i = x_i^* = -h_i \neq 0$.

We need three auxiliary lemmas. The first one is taken from [15].

Lemma 2.14 ([15]). *For fixed $\alpha \geq 0$, $\beta \geq 0$, $a \geq 0$, consider the function*

$$f(t) = \sqrt{t^2 - a^2 + \beta} - \sqrt{(t-a)^2 + \alpha}, \quad t \in [a, +\infty).$$

Then $f(t) \geq \min\{a, \sqrt{\beta} - \sqrt{\alpha}\}$ for all $t \geq a$. Moreover, if $\sqrt{\beta} - \sqrt{\alpha} > a$, then $f(t) > a$ for all $t \geq a$.

Lemma 2.15 (Truncated analogue of [15, Lemma 10]). *For any given $x \in \mathbb{R}^n$ and $0 \neq h \in \mathbb{R}^n$, with $S = \text{supp}(x)$ such that $S \cap T \neq \emptyset$,*

$$f_T(x, h) \geq \|h_{T \cap S^C}\|_1 - \|h_{T \cap S}\|_1 + \min\left\{\|h_{T \cap S}\|_2 - \|h_{T \cap S^C}\|_2, \min_{i \in S \cap T} |h_i|\right\}, \quad (18)$$

where f_T is defined by (8). Moreover, if

$$\|h_{T \cap S}\|_2 - \|h_{T \cap S^C}\|_2 > \min_{i \in S \cap T} |h_i|, \quad (19)$$

then the strict inequality

$$f_T(x, h) > \|h_{T \cap S^C}\|_1 - \|h_{T \cap S}\|_1 + \min_{i \in S \cap T} |h_i| \quad (20)$$

holds, i.e. (18) is strict once the minimum in (18) is resolved by the second entry; see Remark 2.16 below. When $T = [n]$ the lemma reduces exactly to [15, Lemma 10].

Remark 2.16 (On the form of (20)). *It is essential that the ℓ_1 terms $\|h_{T \cap S^C}\|_1 - \|h_{T \cap S}\|_1$ appear on the right-hand side of (20): the apparently stronger assertion $f_T(x, h) > \min_{i \in S \cap T} |h_i|$ is false in general. Indeed, take $T = [n]$, $n = 203$, $S = \text{supp}(x) = \{1, 2, 3\}$, $x_S = (-10^3, -10^3, -10^3)^\top$, $h_S = (10, 10, 10)^\top$ and $h_i = 0.1$ for $i \geq 4$. Then $\|h_{T \cap S}\|_2 - \|h_{T \cap S^C}\|_2 = \sqrt{300} - \sqrt{4} \approx 15.91 > 10 = \min_{i \in S \cap T} |h_i|$, so (19) holds, while a direct computation gives $f_T(x, h) \approx 7.32 < 10$. The correct conclusion is the one stated in (20), namely $f_T(x, h) > -10 + 10 = 0$. Note that the counterexample already works for $T = [n]$, i.e. in the setting of [15]: what fails is only the tempting simplification in which the ℓ_1 terms are dropped—the estimate of [15, Lemma 10] itself is stated with these terms in place and is correct. Inequality (20) is their strict part carried over to the truncated setting.*

Proof. From the definition (16) of $\bar{x}$ we obtain

$$\|\bar{x}_T\|_2 - \|x_T\|_2 = \sum_{i \in S_2 \cap T} (|h_i| - |x_i|) \frac{|h_i| + |x_i|}{\|\bar{x}_T\|_2 + \|x_T\|_2} \leq \sum_{i \in S_2 \cap T} (|h_i| - |x_i|), \quad (21)$$

and, similarly,

$$\|(x+h)_T\|_2 - \|(\bar{x}+h)_T\|_2 = \sum_{i \in S_2 \cap T} |h_i + x_i| \frac{|h_i + x_i|}{\|(x+h)_T\|_2 + \|(\bar{x}+h)_T\|_2} \leq \sum_{i \in S_2 \cap T} |h_i + x_i|. \quad (22)$$

Adding (21) and (22) yields

$$\|\bar{x}_T\|_2 - \|x_T\|_2 + \|(x+h)_T\|_2 - \|(\bar{x}+h)_T\|_2 \leq \|(x+h)_T\|_1 - \|x_T\|_1 + \|\bar{x}_T\|_1 - \|(\bar{x}+h)_T\|_1,$$

which, by the definition of f_T , is equivalent to

$$f_T(x, h) \geq f_T(\bar{x}, h). \quad (23)$$

Next we show $f_T(\bar{x}, h) \geq f_T(x^*, h)$. By (16) and (17),

$$\begin{aligned} \|(\bar{x} + h)_T\|_2 - \|(x^* + h)_T\|_2 &= \frac{2 \langle \bar{x} - x^*, h_T \rangle}{\|(\bar{x} + h)_T\|_2 + \|(x^* + h)_T\|_2} \\ &= \sum_{i \in S_1 \cap T} (1 + \text{sign}(x_i h_i)) \frac{2|x_i|}{\|(\bar{x} + h)_T\|_2 + \|(x^* + h)_T\|_2} |h_i|. \end{aligned} \quad (24)$$

For every $i \in S_1 \cap T$ one has $x_i \neq 0$, and:

- if $x_i h_i = 0$, then $h_i = 0$ and $\|(\bar{x} + h)_T\|_2 + \|(x^* + h)_T\|_2 \geq |x_i + h_i| + |x_i^* + h_i| = 2|x_i|$;
- if $x_i h_i < 0$, then $1 + \text{sign}(x_i h_i) = 0$;
- if $x_i h_i > 0$, then $\|(\bar{x} + h)_T\|_2 + \|(x^* + h)_T\|_2 \geq |x_i + h_i| + |x_i^* + h_i| = 2|x_i|$.

Hence, for all $i \in S_1 \cap T$,

$$(1 + \text{sign}(x_i h_i)) \frac{2|x_i|}{\|(\bar{x} + h)_T\|_2 + \|(x^* + h)_T\|_2} \leq 1 + \text{sign}(x_i h_i). \quad (25)$$

Combining (24) and (25),

$$\|(\bar{x} + h)_T\|_2 - \|(x^* + h)_T\|_2 \leq \sum_{i \in S_1 \cap T} (1 + \text{sign}(x_i h_i)) |h_i| = \|h_{S_1 \cap T}\|_1 + \sum_{i \in S_1 \cap T} \text{sign}(x_i) h_i,$$

that is,

$$\sum_{i \in S_1 \cap T} \text{sign}(x_i) h_i - \|(\bar{x} + h)_T\|_2 \geq -\|h_{S_1 \cap T}\|_1 - \|(x^* + h)_T\|_2. \quad (26)$$

Moreover, by (16) and (17),

$$\|(\bar{x} + h)_T\|_1 - \|\bar{x}_T\|_1 = \sum_{i \in S_1 \cap T} \text{sign}(x_i) h_i + \|h_{S^c \cap T}\|_1 - \|h_{S_2 \cap T}\|_1, \quad \|\bar{x}_T\|_2 = \|x_T^*\|_2. \quad (27)$$

Therefore,

$$\begin{aligned} f_T(\bar{x}, h) &= \|(\bar{x} + h)_T\|_1 - \|(\bar{x} + h)_T\|_2 - (\|\bar{x}_T\|_1 - \|\bar{x}_T\|_2) \\ &= \sum_{i \in S_1 \cap T} \text{sign}(x_i) h_i - \|(\bar{x} + h)_T\|_2 + \|h_{S^c \cap T}\|_1 - \|h_{S_2 \cap T}\|_1 + \|\bar{x}_T\|_2 \\ &\geq -\|h_{S_1 \cap T}\|_1 - \|(x^* + h)_T\|_2 + \|h_{S^c \cap T}\|_1 - \|h_{S_2 \cap T}\|_1 + \|x_T^*\|_2 \\ &= \|h_{S^c \cap T}\|_1 - \|h_{S \cap T}\|_1 - \|(x^* + h)_T\|_2 + \|x_T^*\|_2 \\ &= f_T(x^*, h), \end{aligned} \quad (28)$$

where we used (27) in the second line, (26) and $\|\bar{x}_T\|_2 = \|x_T^*\|_2$ in the third line, and $S = S_1 \cup S_2$ in the fourth. In particular,

$$f_T(x^*, h) = \|h_{S^c \cap T}\|_1 - \|h_{S \cap T}\|_1 - \|(x^* + h)_T\|_2 + \|x_T^*\|_2. \quad (29)$$

Now enumerate $S_1 \cap T = \{i_1, \dots, i_l\}$ and define, for $1 \leq v \leq l$,

$$B_v = \sum_{j=1}^v (|x_{i_j}|^2 - |h_{i_j}|^2) + \|h_{S \cap T}\|_2^2, \quad A_v = \sum_{j=1}^v (|x_{i_j}| - |h_{i_j}|)^2 + \|h_{S^c \cap T}\|_2^2.$$

Lemma 2.14 implies that, for every $1 \leq v \leq l$,

$$\sqrt{B_v} - \sqrt{A_v} \geq \sqrt{B_{v-1}} - \sqrt{A_{v-1}} \quad \text{whenever } \sqrt{B_{v-1}} - \sqrt{A_{v-1}} \leq |h_{i_v}|, \quad (30)$$

$$\sqrt{B_v} - \sqrt{A_v} > |h_{i_v}| \quad \text{whenever } \sqrt{B_{v-1}} - \sqrt{A_{v-1}} > |h_{i_v}|, \quad (31)$$

where $A_0 = \|h_{S^C \cap T}\|_2^2$ and $B_0 = \|h_{S \cap T}\|_2^2$. Consequently,

$$\begin{aligned} \|x_T^*\|_2 - \|(x^* + h)_T\|_2 &= \sqrt{B_l} - \sqrt{A_l} \\ &\geq \min\{\sqrt{B_{l-1}} - \sqrt{A_{l-1}}, |h_{i_l}|\} \\ &\vdots \\ &\geq \min\{\sqrt{B_0} - \sqrt{A_0}, |h_{i_j}| : 1 \leq j \leq l\} \\ &= \min\left\{\|h_{S \cap T}\|_2 - \|h_{S^C \cap T}\|_2, \min_{i \in S_1 \cap T} |h_i|\right\}. \end{aligned} \quad (32)$$

Combining (23), (28), (29) and (32) proves (18).

It remains to prove the strict inequality (20) under (19).

Case 1: $\min_{i \in S \cap T} |h_i| = \min_{i \in S_1 \cap T} |h_i|$. In this case the inequalities $\sqrt{B_{v-1}} - \sqrt{A_{v-1}} \leq |h_{i_v}|$ cannot hold for all $v = 1, \dots, l$; otherwise (30) would give $\|h_{S \cap T}\|_2 - \|h_{S^C \cap T}\|_2 \leq \min_{i \in S \cap T} |h_i|$, contradicting (19). Hence there exists $1 \leq j \leq l$ such that $\sqrt{B_{j-1}} - \sqrt{A_{j-1}} > |h_{i_j}|$, and (31) yields

$$\|x_T^*\|_2 - \|(x^* + h)_T\|_2 = \sqrt{B_l} - \sqrt{A_l} \geq \dots \geq \sqrt{B_j} - \sqrt{A_j} > |h_{i_j}| \geq \min_{i \in S \cap T} |h_i|.$$

Case 2: $\min_{i \in S \cap T} |h_i| < \min_{i \in S_1 \cap T} |h_i|$. If $\sqrt{B_{v-1}} - \sqrt{A_{v-1}} \leq |h_{i_v}|$ holds for all v , then by (30)

$$\|x_T^*\|_2 - \|(x^* + h)_T\|_2 \geq \sqrt{B_0} - \sqrt{A_0} = \|h_{T \cap S}\|_2 - \|h_{T \cap S^C}\|_2 > \min_{i \in S \cap T} |h_i|,$$

using (19). Otherwise we are back in the situation of Case 1. In either case,

$$\|x_T^*\|_2 - \|(x^* + h)_T\|_2 > \min_{i \in S \cap T} |h_i|. \quad (33)$$

Combining (23), (28), (29), (32) and (33) yields (20). $\square$

Lemma 2.17. For a vector $h \in \mathbb{R}^n$ and an index set $\Omega \subset [n]$, define

$$\begin{aligned} \Phi_T(\Omega, h) &= \|h_{T \cap \Omega^C}\|_1 - \|h_{T \cap \Omega}\|_1 + \|h_{T \cap \Omega}\|_2 - \|h_{T \cap \Omega^C}\|_2, \\ \Psi_T(\Omega, h) &= \|h_{T \cap \Omega^C}\|_1 - \|h_{T \cap \Omega}\|_1 + \min_{i \in \Omega \cap T} |h_i|. \end{aligned}$$

Then:

- (i) $\Phi_T(D_1, h) \geq \Phi_T(D_2, h)$ whenever $D_1 \cap T \subset D_2 \cap T$;
- (ii) $\Psi_T(D_1, h) \geq \Psi_T(D_2, h)$ whenever $D_1 \cap T \subset D_2 \cap T$;
- (iii) $\Phi_T(S, h) \geq \Phi_T(I, h)$ and $\Psi_T(S, h) \geq \Psi_T(I, h)$ whenever $|S \cap T| \leq |I|$, where $I \subseteq T$ satisfies $\text{supp}((h_T)_{\max(k-|T^C|)}) \subseteq I$ and $|I| = k - |T^C|$.

Proof. (i) Set $\tilde{D}_1 = D_1 \setminus T^C$ and $\tilde{D}_2 = D_2 \setminus T^C$. Then $\tilde{D}_1 \cap T = D_1 \cap T$, $\tilde{D}_1^C \cap T = D_1^C \cap T$ (and similarly for D_2), with $\tilde{D}_1 \subseteq \tilde{D}_2 \subseteq T$, and

$$\Phi_T(D_1, h) = \Phi_T(\tilde{D}_1, h), \quad \Phi_T(D_2, h) = \Phi_T(\tilde{D}_2, h). \quad (34)$$

If $\tilde{D}_1 = \tilde{D}_2$ we are done. Otherwise pick $u \in \tilde{D}_2 \setminus \tilde{D}_1$ and set $D_3 = \tilde{D}_1 \cup \{u\}$, so that $\tilde{D}_1 \subset D_3 \subseteq \tilde{D}_2$. A direct computation gives

$$\begin{aligned} \Phi_T(\tilde{D}_1, h) - \Phi_T(D_3, h) &= 2|h_u| + (\|h_{T \cap \tilde{D}_1}\|_2 - \|h_{T \cap D_3}\|_2) + (\|h_{T \cap D_3^C}\|_2 - \|h_{T \cap \tilde{D}_1^C}\|_2) \\ &= 2|h_u| - \frac{|h_u|^2}{\|h_{T \cap \tilde{D}_1}\|_2 + \|h_{T \cap D_3}\|_2} - \frac{|h_u|^2}{\|h_{T \cap D_3^C}\|_2 + \|h_{T \cap \tilde{D}_1^C}\|_2} \\ &= |h_u| \left(2 - \frac{|h_u|}{\|h_{T \cap \tilde{D}_1}\|_2 + \|h_{T \cap D_3}\|_2} - \frac{|h_u|}{\|h_{T \cap D_3^C}\|_2 + \|h_{T \cap \tilde{D}_1^C}\|_2} \right) \geq |h_u|(2 - 1 - 1) = 0 \end{aligned}$$

Iterating this argument until $D_3 = \tilde{D}_2$ and using (34) proves (i).

(ii) The same reduction gives $\Psi_T(D_1, h) = \Psi_T(\tilde{D}_1, h)$ and $\Psi_T(D_2, h) = \Psi_T(\tilde{D}_2, h)$. With $u \in \tilde{D}_2 \setminus \tilde{D}_1$ and $D_3 = \tilde{D}_1 \cup \{u\}$ we have $\min_{i \in \tilde{D}_1 \cap T} |h_i| \geq \min_{i \in D_3 \cap T} |h_i|$, whence

$$\Psi_T(\tilde{D}_1, h) - \Psi_T(D_3, h) = 2|h_u| + \min_{i \in \tilde{D}_1 \cap T} |h_i| - \min_{i \in D_3 \cap T} |h_i| \geq 0.$$

Iteration yields (ii).

(iii) We prove $\Phi_T(S, h) \geq \Phi_T(I, h)$; the argument for Ψ_T is analogous. By (i) it suffices to treat the case $|S \cap T| = |I|$. Put $\tilde{S} = S \setminus T^C$, so that $\tilde{S} \cap T = S \cap T$, $\tilde{S}^C \cap T = S^C \cap T$ and $\Phi_T(S, h) = \Phi_T(\tilde{S}, h)$, with $\tilde{S} \subseteq T$, $I \subseteq T$, $|\tilde{S}| = |I|$. If $\tilde{S} = I$ we are done. Otherwise pick $u \in I \setminus \tilde{S}$ and $v \in \tilde{S} \setminus I$, and set $S' = (\tilde{S} \setminus \{v\}) \cup \{u\}$. By the definition of I we have $|h_u| \geq |h_v|$, and therefore

$$\begin{aligned} \Phi_T(\tilde{S}, h) - \Phi_T(S', h) &= 2(|h_u| - |h_v|) + (\|h_{T \cap \tilde{S}}\|_2 - \|h_{T \cap S'}\|_2) + (\|h_{T \cap S'^C}\|_2 - \|h_{T \cap \tilde{S}^C}\|_2) \\ &= 2(|h_u| - |h_v|) + \frac{|h_v|^2 - |h_u|^2}{\|h_{T \cap \tilde{S}}\|_2 + \|h_{T \cap S'}\|_2} + \frac{|h_v|^2 - |h_u|^2}{\|h_{T \cap S'^C}\|_2 + \|h_{T \cap \tilde{S}^C}\|_2} \\ &= (|h_v| - |h_u|) \left(-2 + \frac{|h_u| + |h_v|}{\|h_{T \cap \tilde{S}}\|_2 + \|h_{T \cap S'}\|_2} + \frac{|h_u| + |h_v|}{\|h_{T \cap S'^C}\|_2 + \|h_{T \cap \tilde{S}^C}\|_2} \right) \geq 0. \end{aligned}$$

Iterating gives $\Phi_T(\tilde{S}, h) \geq \Phi_T(I, h)$ and hence $\Phi_T(S, h) \geq \Phi_T(I, h)$. For Ψ_T , using (ii) we again reduce to $|\tilde{S}| = |I|$; then, by the definition of I ,

$$\|h_{T \cap \tilde{S}}\|_1 - \min_{i \in T \cap \tilde{S}} |h_i| \leq \|h_{T \cap I}\|_1 - \min_{i \in T \cap I} |h_i|, \quad \|h_{T \cap \tilde{S}^C}\|_1 \geq \|h_{T \cap I^C}\|_1,$$

which directly gives $\Psi_T(\tilde{S}, h) - \Psi_T(I, h) \geq 0$. $\square$

We are now in a position to characterize global recovery. As we shall see, its second condition is strong enough to absorb the rank condition (ii) of Theorem 2.10; this is recorded in Remark 2.19 below.

Theorem 2.18. *Let $A \in \mathbb{R}^{m \times n}$. Every nonzero k -sparse vector $x \in \mathbb{R}^n$ with $T^C \subset \text{supp}(x)$ can be recovered from $Ax = y$ via truncated ℓ_{1-2} global minimization if and only if the following two conditions hold:*

- (i) $\|h_{(I \cup T^C)^C}\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| \geq 0$ for all $h \in \ker A$ with $\|h\|_2 = 1$, where $I \subseteq T$ satisfies $\text{supp}((h_T)_{\max(k-|T^C|)}) \subseteq I$ and $|I| = k - |T^C|$;
- (ii) $\|h_{T \cap I^C}\|_1 - \|h_{T \cap I^C}\|_2 - (\|h_{T \cap I}\|_1 - \|h_{T \cap I}\|_2) > 0$ for all $h \in \ker A \setminus \{0\}$, where $I \subseteq T$ satisfies $\text{supp}((h_T)_{\max(k-|T^C|)}) \subseteq I$ and $|I| = k - |T^C|$.

Proof. Necessity. Let $h \in \ker A \setminus \{0\}$ and set $x = -h_I + \mathbf{1}_{T^C}$, where $\mathbf{1}_{T^C}$ is the indicator vector of T^C . Then $\text{supp}(x) \supseteq T^C$ and $\|x\|_0 \leq |I| + |T^C| = k$. Since $(x + h)_T = h_{T \cap I^C}$ and $x_T = -h_I$,

$$\|h_{T \cap I^C}\|_1 - \|h_{T \cap I^C}\|_2 - (\|h_{T \cap I}\|_1 - \|h_{T \cap I}\|_2) = f_T(x, h) > 0$$

by Definition 2.2, which is (ii). Finally, global recovery implies local recovery (Remark 2.3), so condition (i) follows from Theorem 2.10.

Sufficiency. Let $x \neq 0$ with $T^C \subseteq S := \text{supp}(x)$ and $|S| \leq k$, and let $h \in \ker A \setminus \{0\}$. If $x_T = 0$ then $S = T^C$ and Remark 2.5(b) applies: x is recoverable as soon as $|\text{supp}(h_T)| \geq 2$ for every nonzero $h \in \ker A$. This is implied by (ii), for if $|\text{supp}(h_T)| \leq 1$ then $h_{T \cap I^C}$ and $h_{T \cap I}$ each have at most one nonzero entry, whence $\|h_{T \cap I^C}\|_1 = \|h_{T \cap I^C}\|_2$ and $\|h_{T \cap I}\|_1 = \|h_{T \cap I}\|_2$, so the left-hand side of (ii) vanishes—a contradiction. Assume therefore $x_T \neq 0$, i.e. $S \cap T \neq \emptyset$. If $\|h_{T \cap S}\|_2 - \|h_{T \cap S^C}\|_2 \leq \min_{i \in S \cap T} |h_i|$, then Lemma 2.15 and Lemma 2.17(iii) give

$$\begin{aligned} f_T(x, h) &\geq \|h_{T \cap S^C}\|_1 - \|h_{T \cap S}\|_1 + (\|h_{T \cap S}\|_2 - \|h_{T \cap S^C}\|_2) \\ &= \Phi_T(S, h) \geq \Phi_T(I, h) \\ &= \|h_{T \cap I^C}\|_1 - \|h_{T \cap I^C}\|_2 - (\|h_{T \cap I}\|_1 - \|h_{T \cap I}\|_2) > 0. \end{aligned}$$

Otherwise, $\|h_{T \cap S}\|_2 - \|h_{T \cap S^C}\|_2 > \min_{i \in S \cap T} |h_i|$. Then the minimum in (18) is attained by the second entry, and Lemma 2.15 together with Lemma 2.17(iii) and condition (i) gives

$$\begin{aligned} f_T(x, h) &> \|h_{T \cap S^C}\|_1 - \|h_{T \cap S}\|_1 + \min_{i \in S \cap T} |h_i| = \Psi_T(S, h) \geq \Psi_T(I, h) \\ &= \left\| h_{(I \cup T^C)^C} \right\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| \geq 0, \end{aligned} \tag{35}$$

where the strict inequality is (20) of Lemma 2.15 (whose hypothesis is exactly the current case) and the last two steps use Lemma 2.17(iii) and condition (i). Hence $f_T(x, h) > 0$ in both cases, for every $h \in \ker A \setminus \{0\}$. By Definition 2.2, x can be recovered via truncated ℓ_{1-2} global minimization, and the proof is complete. $\square$

Remark 2.19. Condition (ii) of Theorem 2.18 automatically implies condition (ii) of Theorem 2.10, which is why the latter is not listed separately in Theorem 2.18. Indeed, if some $j \in [n] \setminus T^C$ were such that the columns of $A_{T^C \cup \{j\}}$ are linearly dependent, then there would exist $h \in \ker A \setminus \{0\}$ with $\text{supp}(h) \subseteq T^C \cup \{j\}$, hence $|\text{supp}(h_T)| \leq 1$ and therefore $\|h_{T \cap I^C}\|_1 = \|h_{T \cap I^C}\|_2$ and $\|h_{T \cap I}\|_1 = \|h_{T \cap I}\|_2$, contradicting the strict inequality in (ii).

2.3 Consequences: verifiable conditions and strict improvement

The conditions of Theorems 2.10 and 2.18 are *null-space* conditions: they are sharp, but they cannot be verified for a given matrix in polynomial time. In this subsection we derive consequences of Theorem 2.10 that are of a different nature:

- a *restricted isometry* (RIP) condition, which is verifiable for the random matrix ensembles used in practice (Theorem 2.23);
- the boundary case $|T^C| = k - 1$, in which the recovery conditions collapse to a simple rank condition (Corollary 2.25);
- an explicit example showing that the guarantee of Theorem 2.10 is *strictly* stronger than that of the conventional ℓ_{1-2} model (Theorem 2.26).

We recall the restricted isometry constant.

Definition 2.20 (RIC). *The s -th restricted isometry constant $\delta_s = \delta_s(A)$ of $A \in \mathbb{R}^{m \times n}$ is the smallest $\delta \geq 0$ such that*

$$(1 - \delta) \|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta) \|z\|_2^2$$

for all s -sparse vectors $z \in \mathbb{R}^n$.

We need two elementary estimates. The first one bounds the energy of a kernel vector on a small support by the ℓ_1 -norm of its tail.

Lemma 2.21 (RIP tail estimate). *Let $h \in \ker A \setminus \{0\}$, let $\delta_{2k} < 1$, and let $\Omega \subseteq [n]$ with $|\Omega| = k$. Then*

$$\|h_\Omega\|_2 \leq \frac{\sqrt{1 + \delta_k}}{\sqrt{1 - \delta_{2k}}} \frac{\|h_{\Omega^c}\|_1}{\sqrt{k}}.$$

Proof. Partition Ω^c into consecutive blocks $\Omega_1, \Omega_2, \dots$ by repeatedly removing the k entries of largest magnitude: thus $|\Omega_j| = k$ for every j except possibly the last one, and $\min_{i \in \Omega_j} |h_i| \geq \max_{i \in \Omega_{j+1}} |h_i|$ for every j . Note that whenever Ω_{j+1} is defined and nonempty the preceding block satisfies $|\Omega_j| = k$, so that $\|h_{\Omega_j}\|_1 \geq k \min_{i \in \Omega_j} |h_i|$. Since $|\Omega_{j+1}| \leq k$,

$$\|h_{\Omega_{j+1}}\|_2 \leq \sqrt{k} \max_{i \in \Omega_{j+1}} |h_i| \leq \sqrt{k} \min_{i \in \Omega_j} |h_i| \leq \frac{\|h_{\Omega_j}\|_1}{\sqrt{k}}.$$

Summing over $j \geq 2$ and telescoping (see (36)),

$$\sum_{j \geq 2} \|h_{\Omega_j}\|_2 \leq \frac{1}{\sqrt{k}} \sum_{j \geq 1} \|h_{\Omega_j}\|_1 = \frac{\|h_{\Omega^c}\|_1}{\sqrt{k}}. \quad (36)$$

Since $Ah = 0$, we have $A(h_\Omega + h_{\Omega_1}) = -\sum_{j \geq 2} Ah_{\Omega_j}$, and hence

$$\|A(h_\Omega + h_{\Omega_1})\|_2 \leq \sum_{j \geq 2} \|Ah_{\Omega_j}\|_2 \leq \sqrt{1 + \delta_k} \sum_{j \geq 2} \|h_{\Omega_j}\|_2 \leq \sqrt{1 + \delta_k} \frac{\|h_{\Omega^c}\|_1}{\sqrt{k}}.$$

On the other hand $|\Omega \cup \Omega_1| \leq 2k$, so the lower RIP bound gives

$$\|A(h_\Omega + h_{\Omega_1})\|_2 \geq \sqrt{1 - \delta_{2k}} \|h_\Omega + h_{\Omega_1}\|_2 = \sqrt{1 - \delta_{2k}} \sqrt{\|h_\Omega\|_2^2 + \|h_{\Omega_1}\|_2^2} \geq \sqrt{1 - \delta_{2k}} \|h_\Omega\|_2,$$

where we used that Ω and Ω_1 are disjoint. Combining the two displays yields the assertion. $\square$

Lemma 2.22 (Dropping the smallest entry). *Let $h \in \mathbb{R}^n$ and let $I \subseteq [n]$ with $|I| = p \geq 1$. Then*

$$\|h_I\|_1 - \min_{i \in I} |h_i| \leq \sqrt{p-1} \|h_I\|_2.$$

Proof. Choose $i_0 \in \arg \min_{i \in I} |h_i|$. Then

$$\|h_I\|_1 - \min_{i \in I} |h_i| = \|h_{I \setminus \{i_0\}}\|_1 \leq \sqrt{p-1} \|h_{I \setminus \{i_0\}}\|_2 \leq \sqrt{p-1} \|h_I\|_2,$$

by the Cauchy–Schwarz inequality. $\square$

Theorem 2.23 (RIP condition for local recovery). *Let $k \geq 2$ and let $t = |T^C|$ satisfy $0 \leq t \leq k - 2$. If*

$$\delta_{2k} < \frac{t+1}{2k-t-1}, \quad (37)$$

where δ_{2k} is the restricted isometry constant of Definition 2.20, then every nonzero k -sparse vector $x \in \mathbb{R}^n$ with $T^C \subset \text{supp}(x)$ can be recovered from $Ax = y$ via truncated ℓ_{1-2} local minimization.

Proof. Write $\delta = \delta_{2k}$. Since $t \leq k - 2$ we have $t + 1 \leq k - 1 < 2k - t - 1$, so the right-hand side of (37) is strictly less than 1; thus $\delta < 1$.

Condition (ii) of Theorem 2.10. For $j_0 \in [n] \setminus T^C$ the submatrix $A_{T^C \cup \{j_0\}}$ has $t+1 \leq k \leq 2k$ columns. If its columns were linearly dependent, there would be a nonzero z supported on $T^C \cup \{j_0\}$ with $Az = 0$, contradicting the lower RIP bound $\|Az\|_2^2 \geq (1 - \delta) \|z\|_2^2 > 0$. Hence the columns are linearly independent.

Condition (i) of Theorem 2.10. Let $h \in \ker A$ with $\|h\|_2 = 1$, let I be as in Theorem 2.10, and put $S = I \cup T^C$, so that $|S| = k$, $|I| = k - t \geq 2$ and $S^C = (I \cup T^C)^C$. First, we claim that $h_{S^C} \neq 0$: otherwise h would be k -sparse, and the lower RIP bound would give $0 = \|Ah\|_2^2 \geq (1 - \delta) \|h\|_2^2 > 0$, a contradiction. By Lemma 2.22 with $p = k - t$ and by $I \subseteq S$,

$$\|h_I\|_1 - \min_{i \in I} |h_i| \leq \sqrt{k - t - 1} \|h_I\|_2 \leq \sqrt{k - t - 1} \|h_S\|_2,$$

while Lemma 2.21 and the monotonicity $\delta_k \leq \delta_{2k}$ give

$$\|h_S\|_2 \leq \frac{\sqrt{1 + \delta}}{\sqrt{1 - \delta}} \frac{\|h_{S^C}\|_1}{\sqrt{k}}.$$

Therefore

$$\|h_I\|_1 - \min_{i \in I} |h_i| \leq \underbrace{\sqrt{\frac{k - t - 1}{k}} \frac{\sqrt{1 + \delta}}{\sqrt{1 - \delta}}}_{=: \gamma} \|h_{S^C}\|_1.$$

Now $\gamma \leq 1$ is equivalent to $\frac{1 + \delta}{1 - \delta} \leq \frac{k}{k - t - 1}$, i.e. to $(1 + \delta)(k - t - 1) \leq k(1 - \delta)$, i.e. to $\delta(2k - t - 1) \leq t + 1$, which holds by (37). Since $\|h_{S^C}\|_1 > 0$, we conclude that

$$\left\| h_{(I \cup T^C)^C} \right\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| = \|h_{S^C}\|_1 - (\|h_I\|_1 - \min_{i \in I} |h_i|) \geq 0,$$

which is condition (i). The conclusion now follows from Theorem 2.10. $\square$

Remark 2.24 (Interpretation of (37)). The bound $\frac{t+1}{2k-t-1}$ is increasing in $t = |T^C|$: it equals $\frac{1}{2k-1}$ when no prior information is used ($t = 0$) and tends to 1 as $t \rightarrow k$. This quantifies, in a verifiable currency, the intuition already expressed in Remark 2.13: the more active indices are known in advance, the weaker the requirement on the measurement matrix. For example, for $k = 22$ the bound reads $\delta_{44} < \frac{1}{43} \approx 0.023$ when $t = 0$, but improves to $\delta_{44} < \frac{5}{39} \approx 0.128$ when $t = 4$ and to $\delta_{44} < \frac{9}{35} \approx 0.257$ when $t = 8$. We do not claim that (37) is sharp; obtaining optimal RIP constants (in the spirit of the sharp ℓ_1 bounds) requires the refined multi-block technique of Cai and Zhang and is left for future work.

Corollary 2.25 (The case of maximal prior information). Let $|T^C| = k - 1$. Then:

- (a) condition (i) of Theorem 2.10 holds automatically, hence every nonzero k -sparse x with $T^C \subset \text{supp}(x)$ is recoverable via truncated ℓ_{1-2} local minimization if and only if the columns of $A_{T^C \cup \{j_0\}}$ are linearly independent for every $j_0 \in [n] \setminus T^C$;
- (b) for global minimization, condition (i) of Theorem 2.18 holds automatically and condition (ii) reduces to

$$|\text{supp}(h_{T \cap I^C})| \geq 2 \quad \text{for all } h \in \ker A \setminus \{0\}.$$

Proof. Here $|I| = k - |T^C| = 1$, say $I = \{i_1\}$. Then $\min_{i \in I} |h_i| = |h_{i_1}| = \|h_I\|_1$, so condition (i) of Theorem 2.10 becomes $\left\| h_{(I \cup T^C)^C} \right\|_1 \geq 0$, which is always true; this gives (a). For (b), the same computation shows that condition (i) of Theorem 2.18 is automatic, while in condition (ii) we have $\|h_{T \cap I}\|_1 = \|h_{T \cap I}\|_2 = |h_{i_1}|$, so (ii) becomes $\|h_{T \cap I^C}\|_1 - \|h_{T \cap I^C}\|_2 > 0$. Since $\|v\|_1 = \|v\|_2$ holds if and only if $|\text{supp}(v)| \leq 1$, this is equivalent to $|\text{supp}(h_{T \cap I^C})| \geq 2$. $\square$

Corollary 2.25 is conceptually satisfying: if $k - 1$ of the k active indices are already known, then recovery amounts to locating a single remaining index, and the only requirement is the mild rank condition (ii).

Theorem 2.26 (Strict improvement over the conventional ℓ_{1-2} model). *There exist a measurement matrix A , a sparsity level k and an index set T^C such that*

- (i) *every nonzero k -sparse vector satisfying $T^C \subset \text{supp}(x)$ is recoverable via truncated ℓ_{1-2} local minimization, while*
- (ii) *the conventional ℓ_{1-2} minimization model (3) fails to recover all k -sparse vectors.*

Consequently, the recovery guarantee of Theorem 2.10 is strictly stronger than the corresponding guarantee for the conventional model [15].

Proof. Let $n = 3$, $m = 2$, $k = 2$, $T^C = \{3\}$ (so $T = \{1, 2\}$) and

$$A = \begin{pmatrix} 1 & 1 & -5 \\ 1 & -1 & -1 \end{pmatrix}, \quad v = (3, 2, 1)^\top.$$

One checks that $Av = 0$ and that $\text{rank}(A) = 2$ (its first two columns are linearly independent), hence $\ker A = \text{span}\{v\}$.

The truncated model. Here $t = |T^C| = 1$ and $|I| = k - t = 1$. For $h = cv \neq 0$ we have $h_T = (3c, 2c)$, so $I = \{1\}$ and $(I \cup T^C)^C = \{2\}$. Condition (i) of Theorem 2.10 reads

$$\|h_{(I \cup T^C)^C}\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| = |h_2| - |h_1| + |h_1| = |h_2| = 2|c| \geq 0,$$

and it holds. Condition (ii) requires the linear independence of the columns of $A_{\{3,1\}}$ and $A_{\{3,2\}}$, which follows from

$$\det \begin{pmatrix} -5 & 1 \\ -1 & 1 \end{pmatrix} = -4 \neq 0, \quad \det \begin{pmatrix} -5 & 1 \\ -1 & -1 \end{pmatrix} = 6 \neq 0.$$

By Theorem 2.10, every 2-sparse x with $x_3 \neq 0$ is recoverable, which is claim (i).

The conventional model. Take $T^C = \emptyset$, i.e. $T = [3]$. Then $|I| = k = 2$ and $I = \text{supp}(v_{\max(2)}) = \{1, 2\}$, so that condition (i) of Theorem 2.10 reads

$$\|h_{I^C}\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| = |h_3| - (|h_1| + |h_2|) + \min\{|h_1|, |h_2|\} = |c| - 5|c| + 2|c| = -2|c| < 0$$

for $h = cv \neq 0$. Hence the condition fails and, by Theorem 2.10, the conventional ℓ_{1-2} model does not recover all 2-sparse signals. Explicitly, for $x = (-3, -2, 0)^\top$ and $h = v$ the quantity in (9) (which is positively homogeneous in h) equals

$$\sum_{i \in T \cap S} \text{sign}(x_i) h_i + \|h_{S^C}\|_1 - \left\langle h_T, \frac{x_T}{\|x_T\|_2} \right\rangle = (-3 - 2) + 1 - \frac{-13}{\sqrt{13}} = -4 + \sqrt{13} < 0,$$

so this x is not recoverable by ℓ_{1-2} local minimization. This proves (ii). Note that $x_3 = 0$, i.e. x violates the prior information $T^C \subset \text{supp}(x)$ and therefore is not required to be recovered by the truncated model, so the two statements are consistent. $\square$

3 Numerical Algorithm

In this section we design an algorithm for solving the nonconvex problem (6). Let

$$C = \{x \in \mathbb{R}^n : Ax = y\}, \quad \iota_C(x) = \begin{cases} 0, & x \in C, \\ +\infty, & x \notin C. \end{cases}$$

Then (6) is equivalent to the unconstrained problem

$$\min_{x \in \mathbb{R}^n} \|x_T\|_1 - \|x_T\|_2 + \iota_C(x). \quad (38)$$

Difference-of-convex reformulation. The objective in (38) is a difference of two convex functions, $g(x) - h(x)$ with $g(x) = \|x_T\|_1 + \iota_C(x)$ and $h(x) = \|x_T\|_2$. Since $\frac{x_T}{\|x_T\|_2} \in \partial \|x_T\|_2$ for $x_T \neq 0$, the classical DCA replaces h at the current iterate x^k by its linearization, leading to the convex subproblem

$$x^{k+1} = \arg \min_{x \in \mathbb{R}^n} \|x_T\|_1 - \langle x_T, v^k \rangle + \iota_C(x), \quad v^k = \begin{cases} 0, & x_T^k = 0, \\ \frac{x_T^k}{\|x_T^k\|_2}, & x_T^k \neq 0. \end{cases} \quad (39)$$

The outer iteration is summarized in Algorithm 1.

Algorithm 1: Truncated ℓ_{1-2} minimization via DCA

Input: $A \in \mathbb{R}^{m \times n}$, $y \in \mathbb{R}^m$, index set T , tolerance $\epsilon > 0$, maximum number of outer iterations **Outmax**
Output: approximate solution x^*

```

2 Initialize  $x^0 = A^+y$ ,  $k = 0$ 
3 while  $k < \text{Outmax}$  do
5   if  $x_T^k = 0$  then
6      $v^k = 0$ 
7   else
8      $v^k = x_T^k / \|x_T^k\|_2$ 
9   end
11  Compute  $x^{k+1}$  by solving (39) with Algorithm 2
13  if  $\frac{\|x^{k+1} - x^k\|_2}{\max\{1, \|x^k\|_2\}} > \epsilon$  then
14     $k \leftarrow k + 1$ 
15  else
16    break
17  end
18 end
20 Set  $x^* = x^{k+1}$ 

```

By the standard convergence theory of DCA (see, e.g., [12]), the sequence of objective values generated by Algorithm 1 is non-increasing and convergent, and every cluster point of $\{x^k\}$ is a critical point of (38).

Solving the subproblem. Problem (39) has no closed-form solution, so we solve it by a primal-dual splitting method. Recall the proximity operator of a convex function φ :

$$\text{prox}_\varphi(x) = \arg \min_{u \in \mathbb{R}^n} \left\{ \frac{1}{2} \|u - x\|_2^2 + \varphi(u) \right\}.$$

Set

$$\rho(x) = \|x_T\|_1 - \langle x_T, v^k \rangle, \quad \delta(x) = \iota_C(x).$$

Given parameters $\alpha > \beta > 0$ and $t = \beta/\alpha$, the inner iteration reads as follows.

Both proximity operators can be evaluated in closed form.

Step (i). Writing $w^k = x^k + t y^k - t x^k - t z^k$, the first step is separable:

$$x^{k+1} = \arg \min_{x \in \mathbb{R}^n} \left\{ \frac{\alpha}{2} \|x - w^k\|_2^2 + \|x_T\|_1 - \langle x_T, v^k \rangle \right\},$$

Algorithm 2: Primal–dual solver for the subproblem (39)

Input: v^k , parameters $\alpha > \beta > 0$, tolerance $\epsilon > 0$, maximum number of inner iterations **Inmax**

Output: approximate solution of (39)

```

2 Initialize  $x^0 = 0, y^0 = 0, z^0 = 0, k = 0, t = \beta/\alpha$ 
3 for  $k = 0, 1, 2, \dots$  do
5    $x^{k+1} = \text{prox}_{\rho/\alpha}(x^k + t y^k - t x^k - t z^k)$ 
7    $y^{k+1} = \text{prox}_{\delta/\beta}(z^k + x^{k+1})$ 
9    $z^{k+1} = z^k + x^{k+1} - y^{k+1}$ 
11  if  $\max\left\{\frac{\|x^{k+1} - x^k\|_2}{\max\{1, \|x^k\|_2\}}, \frac{\|y^{k+1} - y^k\|_2}{\max\{1, \|y^k\|_2\}}, \frac{\|z^{k+1} - z^k\|_2}{\max\{1, \|z^k\|_2\}}\right\} > \epsilon$  and  $k < \text{Inmax}$  then
12     $k \leftarrow k + 1$ 
13  else
14     $x^* = x^{k+1}$ ; break
15  end
16 end

```

so that, componentwise,

$$x_i^{k+1} = \begin{cases} w_i^k, & i \notin T, \\ \text{soft}(w_i^k + \frac{v_i^k}{\alpha}, \frac{1}{\alpha}), & i \in T, \end{cases} \quad (40)$$

where $\text{soft}(a, \tau) = \text{sign}(a) \max\{|a| - \tau, 0\}$ is the soft-thresholding operator.

Step (ii). Let $\eta_1, \dots, \eta_{n-m}$ be a basis of $\ker A$ and let ξ be any particular solution of $Ax = y$. Every solution of $Ax = y$ can be written as $B\lambda + \xi$, where $B = (\eta_1, \dots, \eta_{n-m}) \in \mathbb{R}^{n \times (n-m)}$ and $\lambda \in \mathbb{R}^{n-m}$. Hence

$$y^{k+1} = \text{prox}_{\delta/\beta}(z^k + x^{k+1}) = \arg \min_{y \in C} \|y - z^k - x^{k+1}\|_2^2 = \arg \min_{\lambda} \|B\lambda + \xi - z^k - x^{k+1}\|_2^2,$$

whose solution is

$$y^{k+1} = B(B^\top B)^{-1}B^\top(z^k + x^{k+1} - \xi) + \xi, \quad (41)$$

i.e., the orthogonal projection of $z^k + x^{k+1}$ onto the affine subspace C . Note that $B^\top B$ is invertible whenever A has full row rank, and the matrix $B(B^\top B)^{-1}B^\top$ can be precomputed once.

Remark 3.1 (Complexity). *Each outer iteration requires one application of Algorithm 2. Every inner iteration costs $O(n)$ for the soft-thresholding step (40) and $O(n(n-m))$ for the projection (41) if the projector is applied densely (this can be reduced by precomputing $B(B^\top B)^{-1}B^\top$). The overall cost is therefore comparable to that of DCA for the conventional ℓ_{1-2} model, the difference being that the shrinkage in (40) is applied only on T .*

4 Numerical Experiments

This section reports numerical experiments designed to (1) verify the effectiveness of the truncated ℓ_{1-2} minimization model (6); (2) compare it with the conventional ℓ_{1-2} minimization model (3); and (3) quantify how the gain depends on the amount and on the *quality* of the prior information.

Experimental setup. For both models we use the same DCA framework (Algorithms 1 and 2) and identical parameter settings, so that the comparison isolates the effect of truncation. In fact, both models are solved by the very same routine: the conventional model (3) is obtained from (6) by setting $T^C = \emptyset$, i.e. $T = [n]$. The measurement matrix $A \in \mathbb{R}^{64 \times 256}$ is a Gaussian random matrix with i.i.d. entries normalized as $A_{ij} \sim \mathcal{N}(0, 1/m)$; the unknown signal x is a random k -sparse signal with standard Gaussian entries on its support. In Algorithm 1 we set `Outmax` = 15 and $\epsilon = 10^{-7}$, and we use the least-squares solution $x^0 = A^+y$ as the initial point. In Algorithm 2 we set $\alpha = 2$, $\beta = 1$, $x^0 = y^0 = z^0 = 0$, $\epsilon = 10^{-7}$ and `Inmax` = 1000. Unless stated otherwise, the prior information T^C consists of indices drawn at random from the true support of x . A recovery is declared successful if

$$\frac{\|x - x^*\|_2}{\|x\|_2} \leq 10^{-3},$$

where x is the ground-truth signal and x^* is the recovered signal. Each success rate is averaged over 200 independent trials (using 4 independent draws of A), and all random seeds are fixed, so the experiments are exactly reproducible. MATLAB (and a reference Python/NumPy) implementation of Algorithms 1–2, of all experiments below, and of the fixed instance used in Figure 1 accompanies the paper.

Single-trial recovery. We first illustrate a single experiment with $|T^C| = 8$ and sparsity $k = 22$, i.e. in the hardest regime of Table 1 (Figure 1). The conventional ℓ_{1-2} model produces visible spurious components outside the true support and distorts several amplitudes (relative error 6.26×10^{-1} , a failure), whereas the signal reconstructed by the truncated model is visually indistinguishable from the ground truth (relative error 8.57×10^{-7}). This confirms that exploiting prior support information substantially improves the quality of the reconstruction.

Recovery rates versus sparsity. Figure 2 and Table 1 display the empirical success rates as the sparsity level k varies, for $|T^C| \in \{0, 4, 8\}$ ($|T^C| = 0$ being the conventional model). Three observations can be made:

- (1) As k grows, the success rates of both models decrease, which is expected since the problem becomes increasingly underdetermined.
- (2) The curve of the conventional ℓ_{1-2} model decays much faster than those of the truncated model.
- (3) The truncated model stays above the conventional model for every sparsity level tested. The gain is negligible in the easy regime ($k \leq 14$, where both models are close to 100%) and becomes substantial once the conventional model approaches its breakdown point: the gap grows from about 1 percentage point at $k = 12$ to 25.5 points ($|T^C| = 4$) and 51 points ($|T^C| = 8$) at $k = 22$. Equivalently, at $k = 22$ the conventional model fails in 62% of the trials, whereas the truncated model with $|T^C| = 8$ still succeeds in 89% of them.

This widening gap is precisely what Remark 2.13 predicts: enlarging T^C weakens the condition $\|h_{(I \cup T^C)^c}\|_1 - \|h_I\|_1 + \min_{i \in I} |h_i| \geq 0$, and the effect is largest in the regime where the conventional model is close to its breakdown point.

Recovery rates versus the number of measurements. Varying the number of measurements m (with $k = 18$ fixed) shows the same effect and makes it quantitative (Table 2 and Figure 3). The conventional model needs about $m = 76$ measurements to reach a 100%

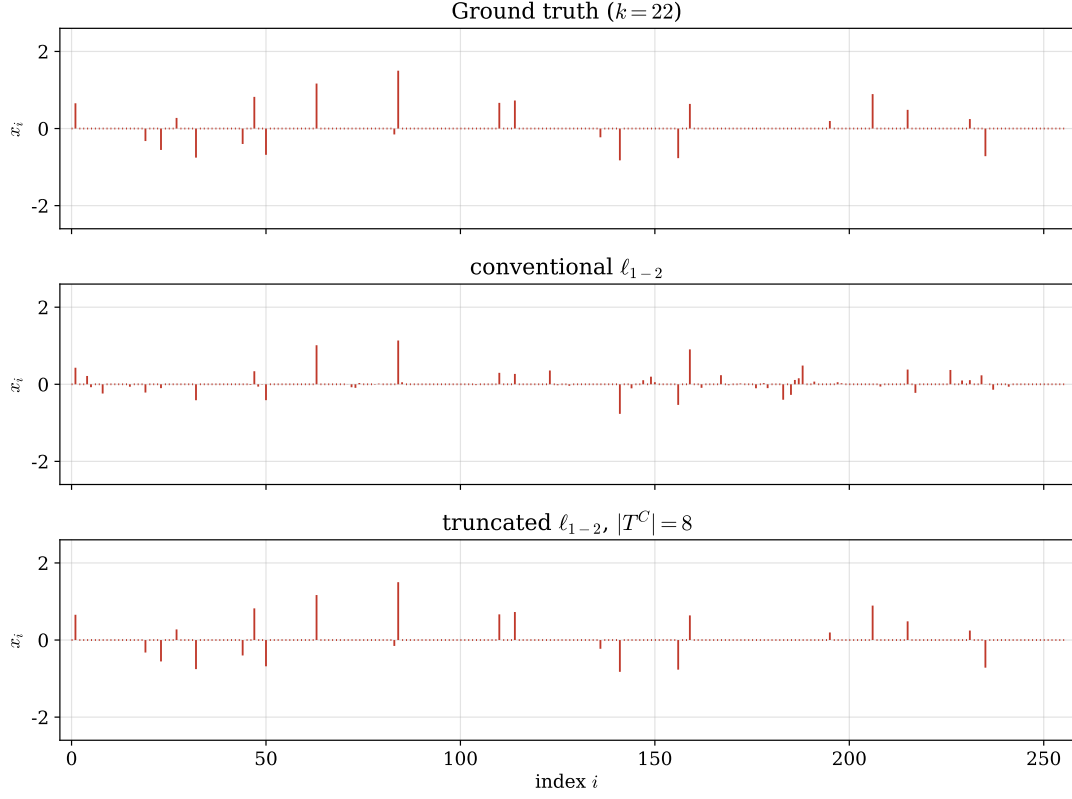


Figure 1: Single recovery instance with $n = 256$, $m = 64$, $k = 22$, $|T^C| = 8$: (a) ground-truth signal, (b) recovered by the conventional ℓ_{1-2} model, (c) recovered by the truncated ℓ_{1-2} model.

Table 1: Success rate (%) versus sparsity k ($n = 256$, $m = 64$, 200 trials per entry).

Model	$k = 12$	14	16	18	20	22
ℓ_{1-2} (conventional)	98.5	98.5	93.0	82.0	60.0	38.0
T- ℓ_{1-2} , $ T^C = 4$	99.5	98.0	100.0	96.0	86.0	63.5
T- ℓ_{1-2} , $ T^C = 8$	99.0	99.0	99.5	98.0	96.5	89.0

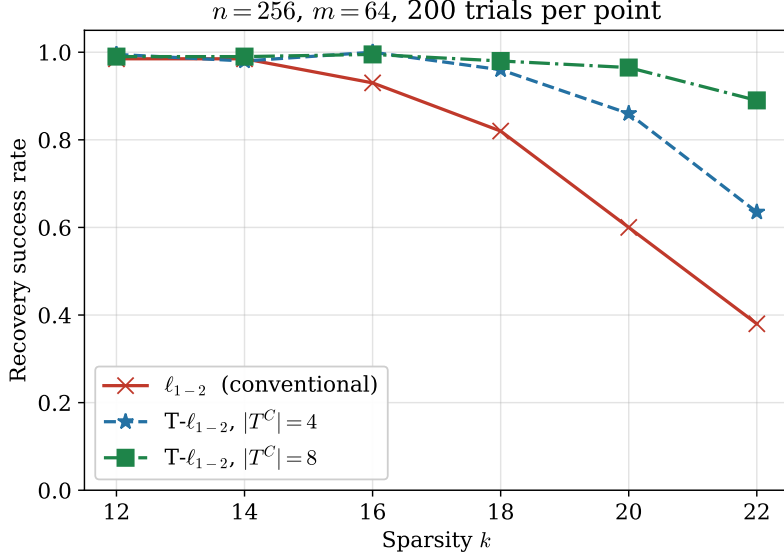


Figure 2: Recovery success rates versus sparsity k ($n = 256$, $m = 64$). Red crosses: conventional ℓ_{1-2} ; blue stars: truncated ℓ_{1-2} with $|T^C| = 4$; green squares: truncated ℓ_{1-2} with $|T^C| = 8$.

success rate, whereas the truncated model with $|T^C| = 8$ already reaches 97.5% at $m = 60$ and 88.8% at $m = 52$. Interpolating the two curves, a given success level is attained by the truncated model with about 11–14 fewer measurements than by the conventional one (e.g. 90% success at $m \approx 53$ instead of $m \approx 67$), i.e. a saving of roughly 20% of the sampling budget. In other words, exploiting $|T^C| = 8$ known active indices shifts the phase transition to the left by about 12 measurements. This is consistent with the RIP bound of Theorem 2.23, since a larger $t = |T^C|$ permits a larger δ_{2k} and hence a smaller m .

Table 2: Success rate (%) versus number of measurements m ($n = 256$, $k = 18$, 80 trials per entry).

Model	$m = 44$	52	60	68	76	84
ℓ_{1-2} (conventional)	0.0	12.5	61.3	95.0	100.0	100.0
T- ℓ_{1-2} , $ T^C = 4$	12.5	43.8	85.0	96.2	98.8	98.8
T- ℓ_{1-2} , $ T^C = 8$	52.5	88.8	97.5	100.0	98.8	100.0

Robustness to erroneous prior information. An important practical question is what happens when the prior information is not fully reliable: in applications the set T^C is typically produced by a support-detection heuristic and may contain indices that are in fact zero. We therefore repeat the experiment with $k = 18$ while injecting a controlled number of *wrong* indices into T^C , i.e. indices drawn from $[n] \setminus \text{supp}(x)$ (Table 3 and Figure 4). The truncated model degrades gracefully: even when *half* of the indices in T^C are wrong (2 out of 4, resp. 4 out of 8), its success rate (88.8%) still exceeds that of the conventional model (82.0%), which uses no prior information at all. Hence the model is useful even with imperfect side information, although—as one would expect—the gain over the conventional model shrinks as the error rate increases.

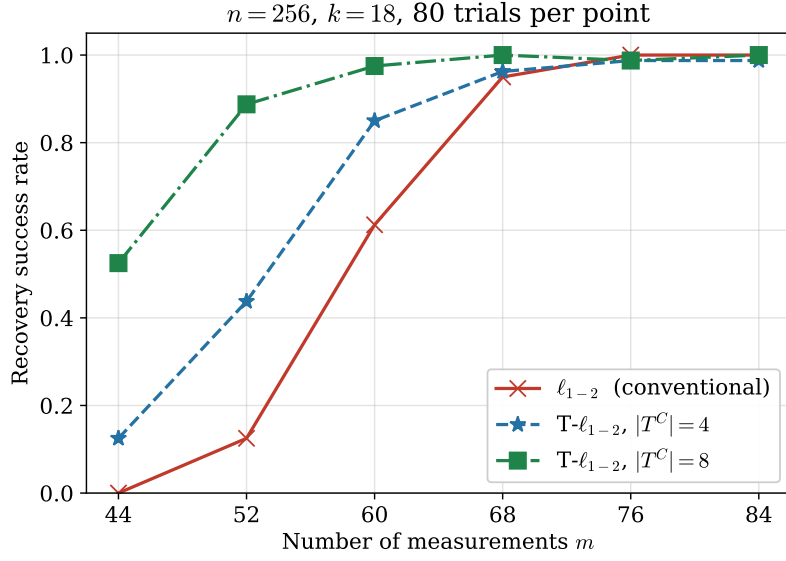


Figure 3: Recovery success rates versus the number of measurements m ($n = 256, k = 18$). Same colour and marker convention as in Figure 2.

Table 3: Success rate (%) versus the number of erroneous indices in T^C ($n = 256, m = 64, k = 18, 80 \text{ trials per entry}$). The conventional ℓ_{1-2} model achieves 82.0% and is independent of T^C .

Model	wrong indices in T^C		
	0	25%	50%
$T\text{-}\ell_{1-2}, T^C = 4$	90.0	85.0 (1/4)	88.8 (2/4)
$T\text{-}\ell_{1-2}, T^C = 8$	100.0	97.5 (2/8)	88.8 (4/8)

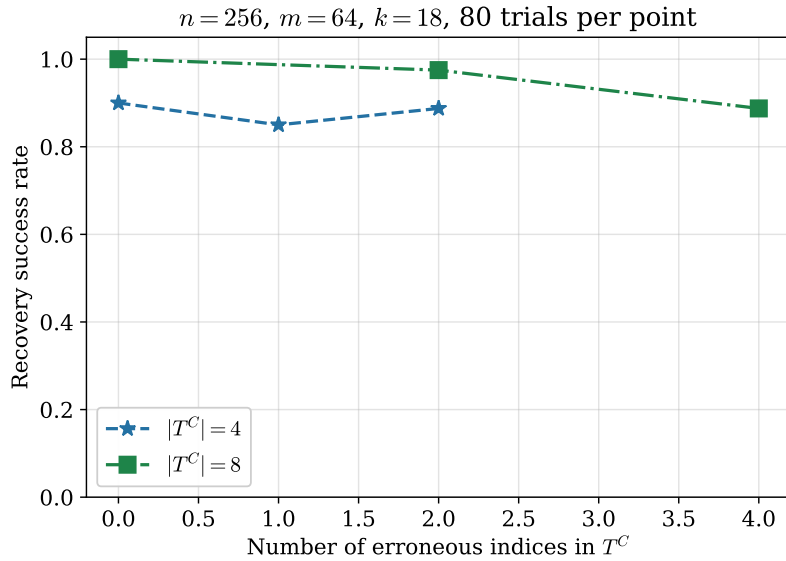


Figure 4: Robustness to erroneous prior information ($n = 256, m = 64, k = 18$). The dashed red line is the conventional ℓ_{1-2} baseline.

Behaviour under noisy measurements. The theory of Section 2 treats the noiseless case, but it is natural to ask whether the advantage persists when the data are contaminated. We replace $y = Ax$ by $y = Ax + e$ with Gaussian noise e at a prescribed signal-to-noise ratio and keep the equality-constrained formulation (6) unchanged. Table 4 and Figure 5 report the median relative reconstruction error. The truncated model is consistently more accurate than the conventional one (by 23% at 10 dB and by 42% at 40 dB for $|T^C| = 8$), and the error of both models decreases steadily with the SNR—by a factor of roughly 2.5–2.9 per 10 dB, i.e. slightly slower than the $\text{SNR}^{-1/2}$ rate one would expect from an ideally stable method. This suggests that the benefit of truncation is not an artefact of the noiseless setting, and motivates the stable-recovery analysis mentioned in Section 5.

Table 4: Median relative reconstruction error under additive noise ($n = 256$, $m = 64$, $k = 18$, 80 trials per entry).

Model	10 dB	20 dB	30 dB	40 dB
ℓ_{1-2} (conventional)	5.62e-1	2.20e-1	7.71e-2	2.70e-2
T- ℓ_{1-2} , $ T^C = 4$	4.62e-1	1.80e-1	5.79e-2	1.98e-2
T- ℓ_{1-2} , $ T^C = 8$	4.32e-1	1.53e-1	4.73e-2	1.58e-2

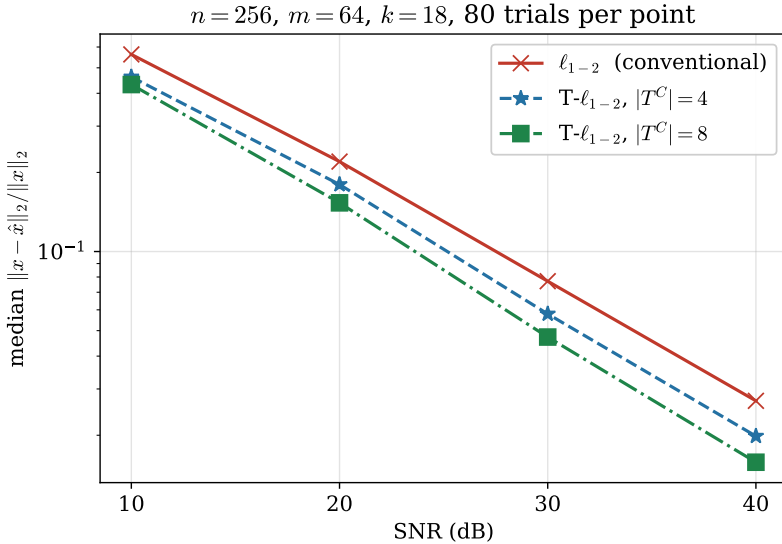


Figure 5: Median relative reconstruction error versus SNR ($n = 256$, $m = 64$, $k = 18$).

Behaviour of the algorithm. Finally we illustrate the monotone decrease of the objective predicted by the DCA theory (Figure 6). On the instance shown, the objective is already stationary after 2 outer iterations for the conventional model and after 1 iteration for the truncated model, and the outer loop terminates after 4 (resp. 3) iterations. In practice `Outmax` = 10 is therefore amply sufficient; we nevertheless keep `Outmax` = 15 throughout. Note that the objective values of the two models are not directly comparable, since they minimize different functions ($\|x_T\|_1 - \|x_T\|_2$ versus $\|x\|_1 - \|x\|_2$) over the same feasible set; the meaningful comparison is the recovery accuracy reported above.

In summary, the numerical results confirm that the truncated ℓ_{1-2} minimization model recovers sparse signals effectively and outperforms the conventional ℓ_{1-2} model; that its advantage increases with the amount of available prior support information and with the difficulty of

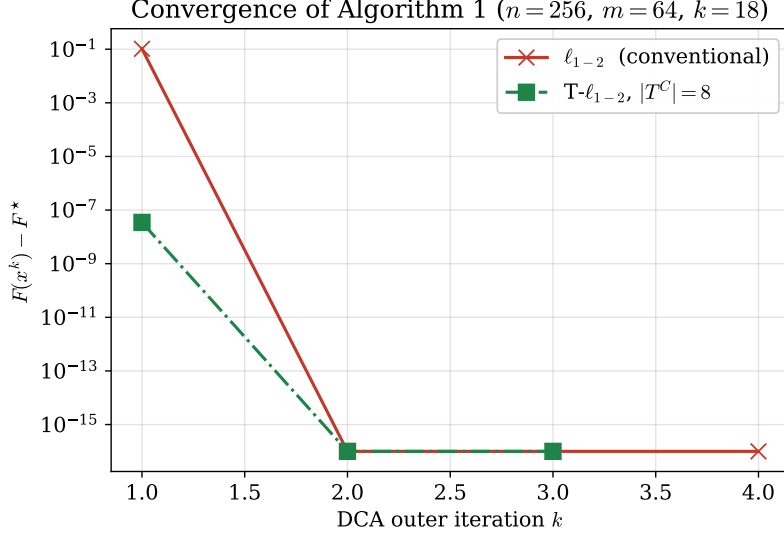


Figure 6: Objective value versus the DCA outer iteration for a single instance ($n = 256$, $m = 64$, $k = 18$).

the problem (large k , small m); that this advantage is robust to errors in the prior information and to measurement noise; and that the proposed algorithm converges in a handful of outer iterations.

5 Conclusion

In this paper we extended the conventional ℓ_{1-2} minimization model to the truncated ℓ_{1-2} minimization model by incorporating prior information about a subset T^C of the support of the target signal. We established necessary and sufficient conditions for the recovery of a fixed sparse signal, of all k -sparse signals by the local minimization model, and of all k -sparse signals by the global minimization model. We showed that these conditions are weaker than the existing ones for the conventional ℓ_{1-2} model [15] and for a related truncated model [17], and that they degenerate to the former when $T^C = \emptyset$. Going beyond null-space conditions, we derived a verifiable RIP sufficient condition whose tolerance $\delta_{2k} < (t+1)/(2k-t-1)$ grows monotonically with the size $t = |T^C|$ of the known active set (Theorem 2.23 and Remark 2.24), we completely settled the extreme case $t = k - 1$, and we proved by an explicit example that the resulting gain is strict: there are matrices for which the truncated model recovers every admissible signal while the conventional model fails. We further designed a DCA-based algorithm with a primal–dual subroutine and derived closed-form expressions for all involved proximity operators. Numerical experiments demonstrate that the proposed model recovers sparse signals reliably, outperforms the conventional ℓ_{1-2} model, and benefits from additional prior support information.

Several directions remain open for future research. First, the RIP bound of Theorem 2.23 is presumably not optimal; sharpening it via the multi-block technique of Cai and Zhang, and deriving coherence-type conditions, would make the theory more directly applicable. Second, the present analysis treats the noiseless and exactly sparse setting; the experiments of Section 4 indicate numerically that the advantage of truncation persists under additive noise and under errors in the prior information, but a rigorous extension to noisy measurements and to approximately sparse (compressible) signals—in the form of stable and robust recovery guarantees, and of a noise-aware variant of (6) that relaxes the equality constraint—remains open and is of significant practical interest. Third, a theoretical study of the global minimization

model beyond the characterization of Theorem 2.18, together with convergence guarantees for the proposed algorithm, deserves further investigation. Finally, designing adaptive strategies for selecting the set T^C from data would make the model applicable when prior information is only partially available.

Declarations

Funding. This research received no external funding.

Author contributions. All authors contributed to the conception of the study, to the mathematical proofs, and to the writing of the manuscript. All authors read and approved the final version.

Data availability. All data generated or analysed during this study are included in this published article.

Conflict of interest. The authors declare that they have no financial or non-financial interests that are relevant to the content of this article.

References

- [1] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, New York, 2013.
- [2] E. Candès, The restricted isometry property and its implications for compressed sensing, *C. R. Math. Acad. Sci. Paris*, 346 (2008), 589–592.
- [3] Q. Mo and S. Li, New bounds on the restricted isometry constant δ_{2k} , *Appl. Comput. Harmon. Anal.*, 31 (2011), 460–468.
- [4] T. Cai and A. Zhang, Sparse representation of a polytope and recovery of sparse signals and low-rank matrices, *IEEE Trans. Inform. Theory*, 60 (2014), 122–132.
- [5] P. Li and W. Chen, Signal recovery under cumulative coherence, *J. Comput. Appl. Math.*, 343 (2018), 266–275.
- [6] Y. Zhao, RSP-based analysis for sparsest and least ℓ_1 -norm solutions to underdetermined linear systems, *IEEE Trans. Signal Process.*, 61 (2013), 5777–5788.
- [7] Y. Zhao, Equivalence and strong equivalence between the sparsest and least ℓ_1 -norm nonnegative solutions of linear systems and their applications, *J. Oper. Res. Soc. China*, 2 (2014), 171–193.
- [8] R. Gribonval and M. Nielsen, Highly sparse representations from dictionaries are unique and independent of the sparseness measure, *Appl. Comput. Harmon. Anal.*, 22 (2007), 335–355.
- [9] R. Zhang and S. Li, Optimal RIP bounds for sparse signals recovery via ℓ_p minimization, *Appl. Comput. Harmon. Anal.*, 47 (2019), 566–584.
- [10] H. Rahimi, C. Wang, H. Dong and Y. Lou, A scale invariant approach for sparse signal recovery, *SIAM J. Sci. Comput.*, 41 (2019), A3649–A3672.

- [11] C. Wang, M. Yan, H. Rahimi and Y. Lou, Accelerated schemes for the L_1/L_2 minimization, *IEEE Trans. Signal Process.*, 68 (2020), 2660–2674.
- [12] P. Yin, Y. Lou, H. Qi and J. Xin, Minimization of ℓ_{1-2} for compressed sensing, *SIAM J. Sci. Comput.*, 37 (2015), A536–A563.
- [13] J. Wen, J. Weng, C. Tong, C. Ren and Z. Zhou, Sparse signal recovery with minimization of 1-norm minus 2-norm, *IEEE Trans. Veh. Technol.*, 68 (2019), 6847–6854.
- [14] N. Bi, J. Tan and W. Tang, A new sufficient condition for sparse vector recovery via L_1 - L_2 local minimization, *Anal. Appl.*, 19 (2021), 1019–1031.
- [15] N. Bi and W. Tang, A necessary and sufficient condition for sparse vector recovery via L_1 - L_2 minimization, *Appl. Comput. Harmon. Anal.*, 56 (2022), 337–350.
- [16] Y. Wang and W. Yin, Sparse signal reconstruction via iterative support detection, *SIAM J. Imaging Sci.*, 3 (2010), 462–491.
- [17] S. Xie and J. Li, k -sparse vector recovery via truncated L_1 - L_2 local minimization, *Optim. Lett.*, 18 (2024), 291–305.
- [18] S. Xie, J. Li and K. Liang, k -sparse vector recovery via $\ell_1 - a\ell_2$ local minimization, *J. Optim. Theory Appl.*, 201 (2024), 75–102.
- [19] F. Gu and A. Wan, Nonconvex L_p - aL_q minimization method and p -RIP condition for stable recovery of approximately k -sparse signals, *Comput. Appl. Math.*, 43 (2024), 33.
- [20] X. Xie, N. Bi and W. Chen, Sparse recovery with coherent frames via ℓ_{1-2} -analysis, *Int. J. Wavelets Multiresolut. Inf. Process.*, 22 (2024), 2450010.
- [21] H. Ge, Y. Xie and W. Chen, Uniform RIP analysis for the $\ell_p - w\ell_q$ -minimization, *J. Comput. Appl. Math.*, 451 (2024), 116100.
- [22] Y. Lu, N. Bi and A. Wan, A reliable iteration algorithm for one-bit compressive sensing on the unit sphere, *Acta Math. Appl. Sin.*, 40 (2024), 2649–2662.
- [23] Y. Lu, N. Bi and A. Wan, $\ell_1 - \ell_p$ ($1 < p \leq 2$)-minimization for one-bit compressed sensing, *IEEE Trans. Signal Process.*, 72 (2024), 3835–3848.