Performance bounds for sparse signal reconstruction with multiple side information [arXiv]

Luong, Huynh Van; Seiler, Jurgen; Kaup, Andre; Forchhammer, Søren; Deligiannis, Nikos

Published in:
ArXiv

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Performance Bounds for Sparse Signal Reconstruction with Multiple Side Information

Huynh Van Luong*, Jürgen Seiler, André Kaup, Søren Forchhammer, and Nikos Deligiannis

***Abstract*—In the context of compressive sensing (CS), this paper considers the problem of reconstructing sparse signals with the aid of other given correlated sources as multiple side information (SI). To address this problem, we propose a reconstruction algorithm with multiple SI (RAMSI) that solves a general weighted n-ℓ_1 norm minimization. The proposed RAMSI algorithm takes advantage of both CS and the n-ℓ_1 minimization by adaptively computing optimal weights among SI signals at every reconstructed iteration. In addition, we establish theoretical performance bounds on the number of measurements that are required to successfully reconstruct the original sparse source using RAMSI under arbitrary support SI conditions. The analyses of the established bounds reveal that RAMSI can achieve sharper bounds and significant performance improvements compared to classical CS. We evaluate experimentally the proposed algorithm and the established bounds using synthetic sparse signals as well as correlated feature histograms, extracted from a multiview image database for object recognition. The obtained results show clearly that the proposed RAMSI algorithm outperforms classical CS and CS with single SI in terms of both the theoretical bounds and the practical performance.**

***Index Terms*—Sparse signal recovery, compressive sensing, multiple side information, n-ℓ_1 minimization, measurement bound.**

I. Introduction

COMPRESSIVE sensing (CS) is a recent theory to perform sparse signal reconstruction, which has attracted significant attention [1]–[17] in the past decade. CS enables sparse signals to be recovered in a computationally tractable manner from a relative limited number of random measurements. This can be done by solving a basic pursuit problem, which involves the l_1-norm minimization of the sparse signal subject to the measurements. It has been shown that sparse signal reconstruction can be even further improved by replacing the l_1-norm by a weighted l_1-norm [5], [6], [12], [13]. The works in [14], [15] give general conditions for exact and robust recovery to provide accurate CS bounds required for successful reconstruction based on convex optimization problems. Furthermore, distributed compressive sensing [16], [17] allows a correlated ensemble of sparse signals to be jointly recovered by exploiting intra- and inter-signal correlations.

In addition, CS with prior information [13], [18]–[25] has drawn considerable attention in investigating theoretical bounds as well as related applications. It was shown that if the prior information is good enough, a l_1-l_1 minimization, which integrates SI into the classical CS, improves the reconstruction dramatically. The bounds on the number of measurements that are required to guarantee perfect reconstruction for CS with prior information are established in [18], [19]. These bounds demonstrate that CS with prior information requires a smaller number of measurements than those of classical CS. Moreover, the works in [13], [22], [23] show efficient reconstructions with single SI for sparse signals from a limited number of measurements in compressive foreground extraction and MRI imaging applications.

Recent emerging applications [26]–[28] such as visual sensor surveillance and mobile augmented reality are following a distributed sensing scenario where a plethora of tiny heterogeneous devices collect information from the environment. In certain scenarios, we may need to deal with very high dimensional data where sensing and processing are reasonably expensive under time-resource constraints. These distributed sensing challenges can be addressed by casting the problem into a distributed sparse representation of multiple sources [28]. The problem in this setup is to represent and reconstruct the sparse sources along with exploiting the correlation among them. One of the key questions is how to robustly reconstruct a compressed source from a small number of measurements given supported observations, i.e., other available sources as SI. Recent attempts to address these challenges in [13], [18], [19], [22], [23] are restricted to considering only one SI and the SI is typically of good quality. However, we are aiming at robustly reconstructing the compressed source given multiple SI signals in the scenarios of the multiple heterogeneous sources changing in time, i.e., arbitrary SI qualities and unknown correlations among them. The challenge raises key interesting questions for solving the distributed sensing problem:

- How can we take advantage of the multiple heterogeneous SI signals? This implies an optimal strategy able to effectively exploit the useful information from the multiple SI signals as well as adaptively eliminate negative effects of poor SI samples.
- How many measurements are required to successfully reconstruct the sparse source given multiple SI signals? This calls for performance bounds on the number of mea-

A part of this work, the RAMSI algorithm, has been presented at the Data Compression Conference 2016 [29].

H. V. Luong, J. Seiler, and A. Kaup are with the Chair of Multimedia Communications and Signal Processing, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany (e-mail: huynh.luong@fau.de, juergen.seiler@fau.de, and andre.kaup@fau.de).

S. Forchhammer is with the Department of Photonics Engineering, Technical University of Denmark, 2800 Lyngby, Denmark (e-mail: sofo@fotonik.dtu.dk).

N. Deligiannis is with the Department of Electronics and Informatics, Vrije Universiteit Brussel, 1050 Brussels, and also with iMinds, 9050 Ghent, Belgium (e-mail: ndeligia@etro.vub.ac.be).

*Corresponding author. Tel: +49 9131 85 27664. Fax: +49 9131 85 28849. E-mail address: huynh.luong@fau.de (H. V. Luong).

surements required to guarantee successful data recovery.

To address the aforementioned challenges, we propose a multi-hypothesis reconstruction using multiple SI signals leading to higher signal recovery performance. This paper contributes in a twofold way. First, an efficient reconstruction algorithm with multiple SI (RAMSI) is proposed. The second contribution is to theoretically establish the measurement bounds of RAMSI under different SI conditions.

The proposed RAMSI algorithm can reconstruct a sparse source with the aid of multiple SI signals. This is done by solving a general weighted n-ℓ_1 minimization, where correlations among sparse sources are taken into account efficiently. The algorithm solves the n-ℓ_1 problem by adaptively selecting optimal weights on multiple SI signals per iteration during the reconstruction process. Contrary to the existing works [13], [18], [19], [22], [23], which exploit only one SI signal, the proposed algorithm can leverage the correlations among multiple heterogenous sources to adapt to on-the-fly changes of the correlations.

We also establish the performance bounds of the proposed RAMSI algorithm to guarantee successful reconstruction using convex optimization tools [14], [15]. The theoretical bounds depend on the support of the source signal to be recovered and the correlations between the source signal and the multiple SI signals. The correlations are expressed via the supports of differences between the source and SI signals, namely, the supports of the subtractions of the source and SI signals. We will show that the RAMSI bounds are sharper compared to those of classic CS [1]–[3] and recent ℓ_1-ℓ_1 reconstruction [18], [19] methods. Furthermore, we show that RAMSI with higher number of SI signals always provides better measurement bounds compared to the one with smaller number of SI signals no matter how the SI qualities vary. These theoretical bounds evidently depict the advantage of RAMSI to deal with heterogeneous SI signals including possible poor SI signals.

The rest of this paper is organized as follows. Section II reviews related works on CS and CS with prior information as well as their corresponding performance bounds. We propose the RAMSI algorithm in Section III. Our established bounds and the corresponding bound analysis are presented in Section IV. We assess the derived bounds and the performance of RAMSI on different sparse sources in Section V.

II. Related Work

In this section, we review the fundamental problem of signal recovery from low-dimensional measurements [2]–[6] including CS (Sec. II-A1) and CS with prior information [13], [18]–[20], [22], [24] (Sec. II-A2) as well as their corresponding known bounds. In addition, we consider the background for measurement bounds (Sec. II-B).

A. Sparse Signal Recovery

1) Compressive Sensing: Low-dimensional signal recovery arises in a wide range of applications such as statistical inference and signal processing. Most signals in such applications have sparse representations in some domain. Let $\boldsymbol{x} \in \mathbb{R}^n$ denote a high-dimensional sparse vector, which is compressible. The source $\boldsymbol{x}$ can be reduced by sampling via a linear projection [3] at the encoder. We denote a random measurement matrix for $\boldsymbol{x}$ by $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}(m < n)$, whose elements are sampled from an i.i.d. Gaussian distribution. Thus, we get a measurement vector $\boldsymbol{y} = \boldsymbol{\Phi x}$, consisting of m elements. At the decoder, the source $\boldsymbol{x}$ can be recovered [2], [3] by solving the Basic Pursuit problem:

$$\min_{\boldsymbol{x}} ||\boldsymbol{x}||_1 \text{ subject to } \boldsymbol{y} = \boldsymbol{\Phi x}, \tag{1}$$

where $||\boldsymbol{x}||_p := (\sum_{i=1}^{n} |x_i|^p)^{1/p}$ is ℓ_p norm of $\boldsymbol{x}$ wherein x_i is an element of $\boldsymbol{x}$.

Problem (1) becomes an instance of finding a general solution:

$$\min_{\boldsymbol{x}}\{H(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})\}, \tag{2}$$

where $f := \mathbb{R}^n \to \mathbb{R}$ is a smooth convex function and $g := \mathbb{R}^n \to \mathbb{R}$ is a continuous convex function possibly non-smooth. Problem (1) is a special case of (2) with $g(\boldsymbol{x}) = \lambda||\boldsymbol{x}||_1$ and $f(\boldsymbol{x}) = \frac{1}{2}||\boldsymbol{\Phi x} - \boldsymbol{y}||_2^2$ with Lipschitz $L_{\nabla f}$ [4]. The results of using proximal gradient methods [4] give that $\boldsymbol{x}^{(k)}$ at iteration k can be iteratively computed by:

$$\boldsymbol{x}^{(k)} = \Gamma_{\frac{1}{L}g}\Big(\boldsymbol{x}^{(k-1)} - \frac{1}{L}\nabla f(\boldsymbol{x}^{(k-1)})\Big), \tag{3}$$

where $L \geq L_{\nabla f}$ and $\Gamma_{\frac{1}{L}g}(\boldsymbol{x})$ is a proximal operator that is defined by:

$$\Gamma_{\frac{1}{L}g}(\boldsymbol{x}) = \underset{\boldsymbol{v}\in\mathbb{R}^n}{\arg\min}\Big\{\frac{1}{L}g(\boldsymbol{v}) + \frac{1}{2}||\boldsymbol{v} - \boldsymbol{x}||_2^2\Big\}. \tag{4}$$

The classical ℓ_1 minimization of CS [1]–[3] requires m_{ℓ_1} measurements [14], [18], [19] for successful reconstruction bounded as:

$$m_{\ell_1} \geq 2s_0 \log\frac{n}{s_0} + \frac{7}{5}s_0 + 1, \tag{5}$$

where $s_0 := \text{nnz}(\boldsymbol{x}) = |i : x_i \neq 0|$ and the $|.|$ denotes the cardinality of a set and $\text{nnz}(.)$ denotes the number of non-zero elements.

2) CS with Prior Information: CS with prior information or SI via ℓ_1-ℓ_1 minimization improves the reconstruction dramatically if the SI has good enough quality [18], [19], [22]. The ℓ_1-ℓ_1 minimization considers reconstructing $\boldsymbol{x}$ given a SI, $\boldsymbol{z} \in \mathbb{R}^n$ by solving the problem (2) with $g(\boldsymbol{x}) = \lambda(||\boldsymbol{x}||_1 + ||\boldsymbol{x} - \boldsymbol{z}||_1)$, i.e., solving:

$$\min_{\boldsymbol{x}}\Big\{\frac{1}{2}||\boldsymbol{\Phi x} - \boldsymbol{y}||_2^2 + \lambda(||\boldsymbol{x}||_1 + ||\boldsymbol{x} - \boldsymbol{z}||_1)\Big\}. \tag{6}$$

The ℓ_1-ℓ_1 minimization problem in (6) has an expression for the bound on the number of measurements required to successfully reconstruct $\boldsymbol{x}$ which is a function of the quality of SI $\boldsymbol{z}$ as given by [18], [19], [22]:

$$m_{\ell_1\text{-}\ell_1} \geq 2\overline{h}\log\Big(\frac{n}{s_0 + \xi/2}\Big) + \frac{7}{5}\Big(s_0 + \frac{\xi}{2}\Big) + 1, \tag{7}$$

where

$$\xi := |\{i : z_i \neq x_i = 0\}| - |\{i : z_i = x_i \neq 0\}| \tag{8a}$$

$$\bar{h} := |\{i : x_i > 0, x_i > z_i\}| \cup |\{i : x_i < 0, x_i < z_i\}|, \tag{8b}$$

wherein x_i, z_i are corresponding elements of $\boldsymbol{x}$, $\boldsymbol{z}$. The authors of [18], [19] have shown that Problem (6) improves over

Problem (1) provided that the SI has good enough quality. The good quality is expressed by a high number of good elements z_i, which are equal to x_i, to lead ξ in (8a) being small. The method presented in this work, however, leads to a high recovery performance independent of the quality of the SI signals.

B. Background for Measurement Bounds

We introduce some key definitions and conditions in convex optimization as well as linear inverse problems, based on the concepts in [14], [15], to be used to derive the measurement bounds for the proposed RAMSI algorithm.

1) Convex Cone: A *convex cone* $C \subset \mathbb{R}^n$ is a convex set that satisfies $C = \tau C, \forall \tau \geq 0$ [15]. For the cone $C \subset \mathbb{R}^n$, a *polar cone* C° is the set of outward normals of C, defined by:

$$C^\circ := \{\boldsymbol{u} \in \mathbb{R}^n : \boldsymbol{u}^{\mathrm{T}}\boldsymbol{x} \leq 0,\ \forall \boldsymbol{x} \in C\}. \tag{9}$$

A *descent cone* (Definition 2.7 [15]) $\mathcal{D}(g, \boldsymbol{x})$ or alias *tangent cone* [14] of a convex function $g := \mathbb{R}^n \to \mathbb{R}$ at a point $\boldsymbol{x} \in \mathbb{R}^n$ at which g is not increasing is defined:

$$\mathcal{D}(g, \boldsymbol{x}) := \bigcup_{\tau \geq 0} \{\boldsymbol{y} \in \mathbb{R}^n : g(\boldsymbol{x} + \tau \boldsymbol{y}) \leq g(\boldsymbol{x})\}, \tag{10}$$

where $\bigcup$ denotes the union operator.

2) Gaussian Width: Before moving to the recovery conditions for linear inverse problems, let us define the Gaussian width [14] on which the recovery conditions are based. This width can be considered as a summary parameter for convex cones. For a convex cone $C \subset \mathbb{R}^n$, considering a subset $C \cap \mathbb{S}^{n-1}$ where $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ is a unit sphere, the *Gaussian width* (Definition 3.1 [14]) is defined as

$$\omega(C) := \mathbb{E}[\sup_{\boldsymbol{u} \in C \cap \mathbb{S}^{n-1}} \boldsymbol{g}^{\mathrm{T}}\boldsymbol{u}]. \tag{11}$$

The Gaussian width is widely used as a summary parameter to measure the aperture of a convex cone to determine the recovery conditions [14]. The Gaussian width (Proposition 3.6 [14]) can further be bounded as follows

$$\omega(C) \leq \mathbb{E}\big[\mathrm{dist}(\boldsymbol{g}, C^\circ)\big], \tag{12}$$

where the $\mathbb{E}[.]$ is an expectation operator and $\mathrm{dist}(\boldsymbol{g}, C^\circ)$ denotes the Euclidean distance of $\boldsymbol{g}$ to the set C° that is defined as

$$\mathrm{dist}(\boldsymbol{g}, C^\circ) := \inf\{||\boldsymbol{g} - \boldsymbol{u}||_2 : \boldsymbol{u} \in C^\circ\}. \tag{13}$$

More recently, a new summary parameter called the statistical dimension $\delta(C)$ of cone C [15], is introduced to estimate the convex cone (Theorem 4.3 [15]). The statistical dimension $\delta(C)$ has an interesting relationship with the Gaussian width $\omega(C)$ which is given in Proposition 10.2 in [15]:

$$\omega^2(C) \leq \delta(C) \leq \omega^2(C) + 1. \tag{14}$$

This relationship gives a convenient bound for the Gaussian width that is to be used in our following computations. The statistical dimension (Proposition 3.1(4) [15]) can be expressed in terms of the polar cone C° (9) by:

$$\delta(C) := \mathbb{E}\big[\mathrm{dist}^2(\boldsymbol{g}, C^\circ)\big]. \tag{15}$$

3) Measurement condition: An optimality condition (Proposition 2.1 [14] and Fact 2.8 [15]) for linear inverse problems states that $\boldsymbol{x}_0$ is the unique solution of (2) if and only if

$$\mathcal{D}(g, \boldsymbol{x}_0) \cap \mathrm{null}(\boldsymbol{\Phi}) = \{\boldsymbol{0}\}, \tag{16}$$

where $\mathrm{null}(\boldsymbol{\Phi}) := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{\Phi}\boldsymbol{x} = \boldsymbol{0}\}$ is the null space of $\boldsymbol{\Phi}$.

We consider the number of measurements m required to successfully reconstruct a given signal $\boldsymbol{x}_0 \in \mathbb{R}^n$. Corollary 3.3(1) [14] shows that if we have measurement $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}_0$, $\boldsymbol{x}_0$ is the unique solution of (2) with probability at least $1 - \exp(-\frac{1}{2}(\sqrt{m} - \omega(\mathcal{D}(g, \boldsymbol{x}_0)))^2)$ provided that $m \geq \omega^2(\mathcal{D}(g, \boldsymbol{x}_0)) + 1$. Furthermore, combined with the relationship in (14), we can interpret the successful recovery of $\boldsymbol{x}_0$ in an equivalent condition:

$$m \geq \delta(\mathcal{D}(g, \boldsymbol{x}_0)) + 1. \tag{17}$$

4) Bound for the measurement condition: The key remaining question is how to calculate the statistical dimension $\delta(\mathcal{D}(g, \boldsymbol{x}))$ of a descent cone $\mathcal{D}(g, \boldsymbol{x})$.

From (15), we can calculate $\delta(\mathcal{D}(g, \boldsymbol{x}))$ as

$$\delta(\mathcal{D}(g, \boldsymbol{x})) = \mathbb{E}\big[\mathrm{dist}^2(\boldsymbol{g}, \mathcal{D}(g, \boldsymbol{x})^\circ)\big], \tag{18}$$

where $\mathcal{D}(g, \boldsymbol{x})^\circ$ is the polar cone of $\mathcal{D}(g, \boldsymbol{x})$ as defined in (9). Let us consider the subdifferential ∂g [30] of a convex function g at a point $\boldsymbol{x} \in \mathbb{R}^n$ is given by $\partial g := \{\boldsymbol{u} \in \mathbb{R}^n : g(\boldsymbol{y}) \geq g(\boldsymbol{x}) + \boldsymbol{u}^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x})$ for all $\boldsymbol{y} \in \mathbb{R}^n\}$. From (18) and Proposition 4.1 in [15], we obtain the upper bound of $\delta(\mathcal{D}(g, \boldsymbol{x}))$ by:

$$\begin{aligned} \delta(\mathcal{D}(g, \boldsymbol{x})) &= \mathbb{E}\big[\inf_{\tau \geq 0} \mathrm{dist}^2(\boldsymbol{g}, \tau \cdot \partial g(\boldsymbol{x}))\big] \\ &\leq \inf_{\tau \geq 0} \mathbb{E}\big[\mathrm{dist}^2(\boldsymbol{g}, \tau \cdot \partial g(\boldsymbol{x}))\big]. \end{aligned} \tag{19}$$

In short, we conclude the following proposition.

Proposition II.1 (The measurement bound of the convex norm function)**.** *In order to obtain the measurement bound for the recovery condition, $m \geq U_g + 1$, we need to calculate the quantity U_g given a convex norm function $g := \mathbb{R}^n \to \mathbb{R}$ by:*

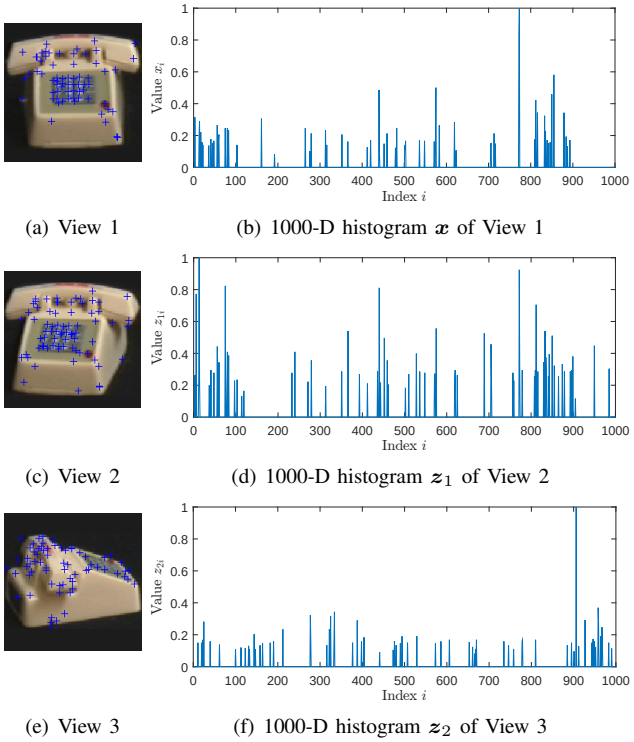
$$U_g = \inf_{\tau \geq 0} \mathbb{E}\big[\mathrm{dist}^2(\boldsymbol{g}, \tau \cdot \partial g(\boldsymbol{x}))\big] \tag{20}$$

where $\boldsymbol{g}$ is a standard normal vector in $\mathbb{R}^n$.

III. Reconstruction with Multiple SI Signals

A. Problem Statement

We consider the problem of efficiently reconstructing a sparse source from low-dimensional random measurements given multiple correlated SI signals. When trying to reconstruct a signal, it is possible to have access to correlated signals, called SI signals, which have spatial or temporal similarities with the target signal. Recently, CS with prior information [13], [18], [19], [22], [23] emerged as an elegant technique to exploit these similarities. For instance, the work in [13] exploits the temporal similarity in MRI longitudinal scans to accelerate MRI acquisition. In the application scenario of video background subtraction [20], [22], [23], prior information, which is generated from a past frame, is used to reduce the number of measurements for a subsequent frame.



(a) View 1 (b) 1000-D histogram $\boldsymbol{x}$ of View 1

(c) View 2 (d) 1000-D histogram $\boldsymbol{z}_1$ of View 2

(e) View 3 (f) 1000-D histogram $\boldsymbol{z}_2$ of View 3

Fig. 1. SIFT-feature histograms of object 60 in the COIL-100 [31].

Existing attempts to incorporate SI in compressed sensing are typically considering one SI signal. We are aiming at considering multiple SI signals which may have both spatial and temporal correlations with the target signal. In the context of the distributed sensing of sparse sources, this problem appears in applications, e.g., distributed visual tracking and recognitions [26]–[28], which involve multiple heterogeneous sources that operate under resource and time constraints. These applications are strained by expensive sensing and processing of very high-dimensional data. The known SI signals can be some other existing reconstructed sources, where we can exploit inter-source redundancies for a robust reconstruction from the reduced sparse source. Therefore, we need an advanced reconstruction method, able to flexibly adapt to on-the-fly changes according to the heterogeneous sparse sources. In this paper, we introduce "the multi-hypothesis reconstruction algorithm using multiple SI signals" (RAMSI) to address the challenge.

Let us consider the scenario of tiny cameras for multiview object recognition in Fig. 1, where a corresponding feature histogram acquired from a given camera is considered as a sparse source $\boldsymbol{x}$. Figure 1 shows three-view images, View 1, View 2, View 3, of Object 60 in the COIL-100 database [31] and the corresponding SIFT-feature [32] points that create sparse feature histogram vectors. The feature histograms are created by extracting all SIFT [32] features from the image then propagating down a hierarchical vocabulary tree based on a hierarchical k-means algorithm [33]. In reality, as we may need very high-dimensional histograms, it is essential to reduce the source dimension by CS before further processing. The idea of CS is reducing the source without prior knowledge of the source distribution. Thus a reduced $\boldsymbol{y}$, which is obtained by compressing 1000 dimensions (D) $\boldsymbol{x}$, is to be conveyed to the joint decoder. We use the available compressed $\boldsymbol{y}$ to reconstruct $\boldsymbol{x}$ with given known SI signals, 1000-D $\boldsymbol{z}_1$ (Fig. 1(d)) and 1000-D $\boldsymbol{z}_2$ (Fig. 1(f)), which are feature histograms of neighbor views. We observe that $\boldsymbol{x}, \boldsymbol{z}_1, \boldsymbol{z}_2$ are naturally correlated in some degree.

In order to reconstruct $\boldsymbol{x}$ from $\boldsymbol{y}$ given $\boldsymbol{z}_1, \boldsymbol{z}_2$, we may go straight to the ℓ_1-ℓ_1 minimization solution (6) with only one input SI either $\boldsymbol{z}_1$ or $\boldsymbol{z}_2$. The number of measurements of $\boldsymbol{y}$ required to successfully reconstruct $\boldsymbol{x}$ is a function of the quality of $\boldsymbol{z}_1$ or $\boldsymbol{z}_2$. There may be a chance that the ℓ_1-ℓ_1 reconstruction performs worse than the ℓ_1 reconstruction, incurred by a not good enough SI, e.g., $\boldsymbol{z}_2$, thereby exposing the drawback of the ℓ_1-ℓ_1 reconstruction. Thus, we propose a new Reconstruction Algorithm with Multiple Side Information (RAMSI), which aims at automatically and optimally utilizing information from the multiple SI signals gleaned from heterogeneous sources. RAMSI is built up not only using the advance of the ℓ_1-ℓ_1 minimization (6) [18], [19], [22], [23] but robustly working on multiple SI signals. RAMSI uses the measurement $\boldsymbol{y}$, which is compressed by $\boldsymbol{y}=\boldsymbol{\Phi x}$, and J known SI signals, $\boldsymbol{z}_1, \boldsymbol{z}_2, ..., \boldsymbol{z}_J \in \mathbb{R}^n$ as inputs. The objective function of RAMSI shall be created as an n-ℓ_1 minimization based on Problem (2) with

$$g(\boldsymbol{x}) = \lambda \sum_{j=0}^{J} ||\mathbf{W}_j(\boldsymbol{x}-\boldsymbol{z}_j)||_1, \tag{21}$$

where $\boldsymbol{z}_0 = \mathbf{0}$ and $\mathbf{W}_j$ are diagonal weight matrices, $\mathbf{W}_j = \mathrm{diag}(w_{j1}, w_{j2}, ..., w_{jn})$, wherein $w_{ji} > 0$ is the weight in $\mathbf{W}_j$ at index i. We formulate the objective function of the n-ℓ_1 minimization problem by:

$$\min_{\boldsymbol{x}} \left\{ H(\boldsymbol{x}) = \frac{1}{2}||\boldsymbol{\Phi x}-\boldsymbol{y}||_2^2 + \lambda \sum_{j=0}^{J} ||\mathbf{W}_j(\boldsymbol{x}-\boldsymbol{z}_j)||_1 \right\}. \tag{22}$$

B. The Proposed RAMSI Algorithm

An important question arises when trying to solve the n-ℓ_1 problem in (22): how to determine the weight values to improve the reconstruction as well as to effectively leverage the multiple SI signals? This also calls for a method to avoid recovery performance degradation when the quality of the SI signals is poor, namely, their correlation with the source signal of interest decreases. We should distribute relevant weights not only to one SI but among SI signals. To optimize the objective function (22) among SI signals, we impose a constraint on all $\mathbf{W}_j$, by doing so, we will be able to assign weights on multiple SI signals according to their qualities during the iterative process. We propose to solve the n-ℓ_1 problem (22) based on the proximal gradient method [4], i.e., at every iteration k we need to update, on the one hand, the weights $\mathbf{W}_j$ and on the other hand compute $\boldsymbol{x}$.

Firstly, for intermediately determining the considered weights $\{w_{ji}\}$ at indices $j \in [0, J]; i \in [1, n]$, we minimize the objective function $H(\boldsymbol{x})$ in (22) by considering $\boldsymbol{x}$ fixed. We may have different strategies to update the weights $\{w_{ji}\}$ depending on our constraint which is defined on multiple SI signals. In this work, we use the constraint $\sum_{j=0}^{J} \mathbf{W}_j = \mathbf{I}$, where $\mathbf{I}$ is a unit diagonal matrix (size of $n \times n$), $\mathbf{I} = \mathrm{diag}(1, 1, ..., 1)$.

Consequently, we compute $\{w_{ji}\}$ by separately optimizing Problem (22) at a given index i of $\mathbf{W}_j$ by:

$$\arg\min_{\{w_{ji}\}}\{H(\boldsymbol{x})\} = \arg\min_{\{w_{ji}\}}\Big\{\lambda\sum_{j=0}^{J} w_{ji}|x_i - z_{ji}|\Big\}, \quad (23)$$

where z_{ji} is an element of $\boldsymbol{z}_j$ at index i. We achieve the minimization of (23) (according to Cauchy's inequality) when all $w_{ji}|x_i{-}z_{ji}|$ are equal to a positive parameter η_i, i.e., $w_{ji}{=}\eta_i/|x_i{-}z_{ji}|$. To ensure that zero-valued components of $|x_i{-}z_{ji}|$ do not prohibit the iterative computation of w_{ji}, we introduce a small parameter $\epsilon{>}0$ that is added to $|x_i{-}z_{ji}|$. As a result, we obtain

$$w_{ji} = \frac{\eta_i}{|x_i - z_{ji}| + \epsilon}, \quad (24)$$

Additionally, by considering the constraint $\sum_{j=0}^{J}\mathbf{W}_j{=}\mathbf{I}$, i.e., $\sum_{j=0}^{J} w_{ji}{=}1$ at an index i of $\mathbf{W}_j$, we can rewrite each weight component w_{ji} as

$$w_{ji} = \frac{1}{1 + (|x_i - z_{ji}| + \epsilon)(\sum\limits_{l=0,\neq j}^{J}(|x_i - z_{li}| + \epsilon)^{-1})}. \quad (25)$$

Given $\mathbf{W}_j$, RAMSI subsequently computes $\boldsymbol{x}^{(k)}$ at iteration k to minimize Problem (22) from (3), where the proximal operator $\Gamma_{\frac{1}{L}g}(x_i)$, which is introduced in (4), is computed by the following proposition (the proof is given in Appendix A).

Proposition III.1 (The proximal operator for multiple SI)**.** *The proximal operator* $\Gamma_{\frac{1}{L}g}(\boldsymbol{x})$ *in* (4) *with* $g(\boldsymbol{x}){=}\lambda\sum_{j=0}^{J}||\mathbf{W}_j(\boldsymbol{x}{-}\boldsymbol{z}_j)||_1$ *is given by:*

$$\Gamma_{\frac{1}{L}g}(x_i){=}\begin{cases} x_i{-}\frac{\lambda}{L}\sum\limits_{j=0}^{J} w_{ji}(-1)^{\mathfrak{b}(l<j)} & \text{if (27a);}\\ z_{li} & \text{if (27b);}\end{cases} \quad (26)$$

$$z_{li}{+}\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}(-1)^{\mathfrak{b}(l<j)}{<}x_i{<}z_{l+1i}{+}\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}(-1)^{\mathfrak{b}(l<j)}; \quad (27a)$$

$$z_{li}{+}\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}(-1)^{\mathfrak{b}(l-1<j)}{\leq}x_i{\leq}z_{li}{+}\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}(-1)^{\mathfrak{b}(l<j)}, \quad (27b)$$

in which we have assumed that, without loss of generality, $-\infty = z_{-1i} \leq z_{0i} \leq z_{1i} \leq ... \leq z_{Ji} \leq z_{J+1i} = \infty$*, and we have defined a boolean function* $\mathfrak{b}(l_i < j) = 1$ *if* $l_i < j$*, otherwise* $\mathfrak{b}(l_i{<}j){=}0$*, here* $-1{\leq}l_i{\leq}J$*.*

Finally, we sum up the proposed RAMSI in Algorithm 1, which is based on a fast iterative FISTA algorithm [4]. It can be noted that the *Stopping criterion* in Algorithm 1 can be either a maximum iteration number $k_{\max}$, a relative variation of the objective function $H(\boldsymbol{x})$ in (22), or a change of the number of nonzero components of the estimate $\boldsymbol{x}^{(k)}$. In this work, the relative variation of $H(\boldsymbol{x})$ is chosen.

IV. Measurement Bounds for RAMSI

We will theoretically establish the bounds of the RAMSI algorithm (Sec. III-B) and then analyze the bounds in relation to the bounds of classical CS in (5) and the ℓ_1–ℓ_1 minimization method in (7).

Algorithm 1: The proposed RAMSI algorithm.

Input: $\boldsymbol{y}, \boldsymbol{\Phi}, \boldsymbol{z}_1, \boldsymbol{z}_2, ..., \boldsymbol{z}_J$;
Output: $\widehat{\boldsymbol{x}}$;
// Initialization of variables and constants.
$\mathbf{W}_0^{(1)}{=}\mathbf{I}$; $\mathbf{W}_j^{(1)}{=}\mathbf{0}$ $(1{\leq}j{\leq}J)$; $\boldsymbol{u}^{(1)}{=}\boldsymbol{x}^{(0)}{=}\mathbf{0}$; $L{=}L_{\nabla f}$; $\lambda, \epsilon{>}0$; $t_1{=}1$; $k{=}0$;
while *Stopping criterion is false* **do**
  $k = k + 1$;
  // Solving the solution given the weights.
  $\nabla f(\boldsymbol{u}^{(k)}) = \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\Phi}\boldsymbol{u}^{(k)} - \boldsymbol{y})$;
  $\boldsymbol{x}^{(k)} = \Gamma_{\frac{1}{L}g}\Big(\boldsymbol{u}^{(k)} - \frac{1}{L}\nabla f(\boldsymbol{u}^{(k)})\Big)$; $\Gamma_{\frac{1}{L}g}(.)$ is given by (26);
  // Computing the updated weights.
  $w_{ji}^{(k+1)} = \frac{1}{1+(|x_i^{(k)}-z_{ji}|+\epsilon)(\sum\limits_{l=0,\neq j}^{J}(|x_i^{(k)}-z_{li}|+\epsilon)^{-1})}$;
  // Updating new values for the next iteration.
  $t_{k+1} = (1 + \sqrt{1 + 4t_k^2})/2$;
  $\boldsymbol{u}^{(k+1)} = \boldsymbol{x}^{(k)} + \frac{t_k - 1}{t_{k+1}}(\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)})$;
end
return $\boldsymbol{x}^{(k)}$;

A. Bound for the RAMSI

In this section, we theoretically compute the measurement bound $U_{g_{n\text{-}\ell_1}}$ (20), which corresponds to the number of measurements required for successful signal recovery, for the RAMSI algorithm. Let us recall from the RAMSI setup (Sec. III) that the regularization function in Problem (22) is $g(\boldsymbol{x}) = \sum_{j=0}^{J}||\mathbf{W}_j(\boldsymbol{x} - \boldsymbol{z}_j)||_1$. First, we take the characteristics of the considered signals into account.

1) Signal Setups: We begin with some definitions regarding the source $\boldsymbol{x}$ and the SI signals $\boldsymbol{z}_j$.

Definition IV.1. *The original sparse signal* $\boldsymbol{x}$ *has* s_0 *nonzero elements and each difference vector* $\boldsymbol{x}{-}\boldsymbol{z}_j$ *has* s_j *nonzero elements for a given SI* $\boldsymbol{z}_j$*, in other words,* $\mathrm{nnz}(\boldsymbol{x}{-}\boldsymbol{z}_j){=}s_j$*, where* $j{\in}[0, J]$*.*

In addition, without loss of generality, we assume that

Definition IV.2. *There are* p *indices* $i{\in}[1, p]$ *where all* $\{x_i{-}z_{ji}\}_{j=0}^{J}$ *values are not equal to zero, meanwhile there are* n–q *indices* $i \in [q{+}1, n]$ *for which all* $\{x_i{-}z_{ji}\}_{j=0}^{J}$ *are equal to zero. As a consequence,* $p{\leq}\inf\{s_j\}$ *and* $q{\geq}\sup\{s_j\}$*.*

By Definition IV.2, we represent the difference vectors in terms of forms as

$$\begin{aligned}
&\boldsymbol{x}{-}\boldsymbol{z}_0{=}(x_1 \quad , ..., x_p \quad , x_{p+1} \quad , ..., x_q \quad , 0, ..., 0)\\
&\boldsymbol{x}{-}\boldsymbol{z}_1{=}(x_1{-}z_{11}, ..., x_p{-}z_{1p}, x_{p+1}{-}z_{1p+1}, ..., x_q{-}z_{1q}, 0, ..., 0)\\
&\cdots\\
&\boldsymbol{x}{-}\boldsymbol{z}_J{=}(x_1{-}z_{J1}, ..., x_p{-}z_{Jp}, x_{p+1}{-}z_{Jp+1}, ..., x_q{-}z_{Jq}, 0, ..., 0).
\end{aligned} \quad (28)$$

As another consequence from Definition IV.2, without loss of generality, we assume that

Definition IV.3. *At a given index $i \in [p+1, q]$, J values of $\{x_i - z_{ji}\}_{j=0}^{J}$ have $d_i \in [1, J]$ values that are zeros, i.e., $\{x_i - z_{ji}\}_{j=l_i+1}^{l_i+d_i} = 0$, where the subtractions of x_i and a SI from z_{l_i+1i} to $z_{l_i+d_i i}$ are zeros.*

By Definitions IV.1, IV.2, IV.3, we can consequently express the total number of zero elements in (28) with the following equality

$$\sum_{i=p+1}^{q} d_i + (J+1)(n-q) = (J+1)n - \sum_{j=0}^{J} s_j. \tag{29}$$

From Definitions IV.2, IV.3, and w_{ji} in (24), we derive

$$w_{ji} = \begin{cases} \frac{\eta_i}{\epsilon}, & \Big\{ i \in [p+1, q], j \in [l_i+1, l_i+d_i] \Big\}, \Big\{ i \in [q+1, n], \forall j \Big\} \\ \frac{\eta_i}{|x_i - z_{ji}| + \epsilon}, & \text{otherwise.} \end{cases} \tag{30}$$

From the constraint $\sum_{j=1}^{J} w_{ji} = 1$, we can derive:

$$\eta_i = \Big(\sum_{j=0}^{J} \frac{1}{|x_i - z_{ji}| + \epsilon} \Big)^{-1}. \tag{31}$$

Definition IV.4. *Focusing on a given index i of the source $\boldsymbol{x}$ and J SI signals $\boldsymbol{z}_j$, without loss of generality regarding the values of $\{z_{ji}\}_{j=0}^{J}$, we assume that $-\infty \le z_{0i} \le z_{1i} \le ... \le z_{Ji} \le \infty$. For convenience, let z_{-1i} and z_{J+1i} denote $-\infty$ and ∞, respectively. We assume $x_i \in (z_{l_i i}, z_{l_i+1 i}]$ with $-1 \le l_i \le J$. Let $\mathfrak{b}(.)$ denote a boolean function, i.e., $\mathfrak{b}(l_i < j) = 1$ if $l_i < j$, otherwise $\mathfrak{b}(l_i < j) = 0$, and let $\operatorname{sign}(.)$ denote a sign function. Consequently, $\operatorname{sign}(x_i - z_{ji}) = (-1)^{\mathfrak{b}(l_i < j)}$.*

2) The Measurement Bound: We derive a useful bound to determine the number of measurements for the successful reconstruction of RAMSI. A general bound is obtained by Theorem IV.5. For convenience, the bound could also be approximately computed via a simpler bound given in (58).

Theorem IV.5 (The measurement bound for multiple SI)**.** *RAMSI requires $m_{n\text{-}\ell_1}$ measurements to successfully reconstruct the source $\boldsymbol{x}$ given J SI signals $\boldsymbol{z}_j$:*

$$m_{n\text{-}\ell_1} \ge 2\bar{a}_{n\text{-}\ell_1} \log \frac{n}{\bar{s}_{n\text{-}\ell_1}} + \frac{7}{5}\bar{s}_{n\text{-}\ell_1} + 1 + \delta_{n\text{-}\ell_1}, \tag{32}$$

where $\bar{a}_{n\text{-}\ell_1}$, $\bar{s}_{n\text{-}\ell_1}$, and $\delta_{n\text{-}\ell_1}$ are defined in (43a)*,* (43c)*, and* (47)*, respectively.*

Proof: We drive the bound based on Proposition II.1 by firstly computing the subdifferential $\partial g(\boldsymbol{x})$ and then the distance between the standard normal vector $\boldsymbol{g}$ to the $\partial g(\boldsymbol{x})$. With the conditions in Definitions IV.2, IV.3, IV.4 and the weights in (30), the $\boldsymbol{u} \in \partial g(\boldsymbol{x})$ of $g(\boldsymbol{x})$ is derived by

$$\boldsymbol{u} = \begin{cases} u_i = a_i, & i = 1, ..., p \\ u_i \in [b_i - c_i, b_i + c_i], & i = p+1, ..., q \\ u_i \in \big[-\sum_{j=0}^{J} w_{ji}, \sum_{j=0}^{J} w_{ji} \big] = [-1, 1], & i = q+1, ..., n, \end{cases} \tag{33}$$

where

$$a_i = \sum_{j=0}^{J} w_{ji} (-1)^{\mathfrak{b}(l_i < j)}, \qquad i \in [1, p] \tag{34a}$$

$$b_i = \sum_{j \notin [l_i+1, l_i+d_i]} w_{ji} (-1)^{\mathfrak{b}(l_i < j)}, \qquad i \in [p+1, q] \tag{34b}$$

$$c_i = \sum_{j=l_i+1}^{l_i+d_i} w_{ji} = d_i \frac{\eta_i}{\epsilon}, \qquad i \in [p+1, q]. \tag{34c}$$

We can compute the distance from the standard normal vector $\boldsymbol{g}$ to the subdifferential $\partial g(\boldsymbol{x})$ based on (13) and (33) as

$$\begin{aligned} \operatorname{dist}^2(\boldsymbol{g}, \tau \cdot \partial g(\boldsymbol{x})) = & \sum_{i=1}^{p} (g_i - \tau a_i)^2 \\ & + \sum_{i=p+1}^{q} \Big(\mathcal{P}^2(g_i - \tau(b_i + c_i)) + \mathcal{P}^2(-g_i + \tau(b_i - c_i)) \Big) \\ & + \sum_{i=q+1}^{n} \mathcal{P}^2(|g_i| - \tau), \end{aligned} \tag{35}$$

where $\mathcal{P}(a) := \max\{a, 0\}$ returns the maximum value between $a \in \mathbb{R}$ and 0. Taking the expectation of (35) delivers

$$\begin{aligned} \mathbb{E}[\operatorname{dist}^2(\boldsymbol{g}, \tau \cdot \partial g(\boldsymbol{x}))] = & \, p + \tau^2 \sum_{i=1}^{p} a_i^2 \\ & + \frac{1}{\sqrt{2\pi}} \sum_{i=p+1}^{q} \int_{\tau(b_i + c_i)}^{\infty} (v - \tau(b_i + c_i))^2 \mathrm{e}^{-v^2/2} \mathrm{d}v \\ & + \frac{1}{\sqrt{2\pi}} \sum_{i=p+1}^{q} \int_{-\infty}^{\tau(b_i - c_i)} (v - \tau(b_i - c_i))^2 \mathrm{e}^{-v^2/2} \mathrm{d}v \\ & + \sqrt{\frac{2}{\pi}} \sum_{i=q+1}^{n} \int_{\tau}^{\infty} (v - \tau)^2 \mathrm{e}^{-v^2/2} \mathrm{d}v. \end{aligned} \tag{36}$$

Replacing the formulations from (76a), (76b) (Appendix B) in (36) gives

$$\begin{aligned} \mathbb{E}[\operatorname{dist}^2(\boldsymbol{g}, \tau \cdot \partial g(\boldsymbol{x}))] = & \, p + \tau^2 \sum_{i=1}^{p} a_i^2 + \sum_{i=p+1}^{q} \mathcal{A}(\tau(b_i + c_i)) \\ & + \sum_{i=p+1}^{q} \mathcal{B}(\tau(b_i - c_i)) + 2 \sum_{i=q+1}^{n} \mathcal{A}(\tau), \end{aligned} \tag{37}$$

where

$$\mathcal{A}(x) := \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} (v - x)^2 e^{-v^2/2} \mathrm{d}v \tag{38a}$$

$$\mathcal{B}(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} (v - x)^2 e^{-v^2/2} \mathrm{d}v. \tag{38b}$$

We may emphasize the advantage of the adaptive weights in (30), which take higher value when the SI signals' values are closer to the target values. Let us focus on a given index $i \in [p+1, q]$, that is, we can observe the weight distributions in the expressions of b_i (34b), c_i (34c). Obviously, the weights w_{ji} taking part in c_i are considerably higher than those of in b_i since their SI signals z_{ji} are equal to source x_i. Moreover, we recall that the small positive parameter ϵ is introduced aiming at the zero values of $x_i - z_{ji}$ which do not prohibit the iterative

computations. This reasonably small ϵ can ensure that c_i (34c) is always greater than $|b_i|$ (34b). Consequently, we have $b_i{+}c_i{>}0$ and $b_i{-}c_i{<}0$. Using the inequality (78a) (Appendix B) with $x>0$ for the formulations $\mathcal{A}$ and the inequality (78b) (Appendix B) with $x<0$ for the formulation $\mathcal{B}$ in (37) is to reach the bound in (20) as follows

$$U_{g_{n\text{-}\ell_1}} \le \inf_{\tau\ge 0}\Big\{p{+}\tau^2\sum_{i=1}^{p}a_i^2{+}\sum_{i=p+1}^{q}\Big(\frac{\psi(\tau(b_i{+}c_i))}{\tau(b_i{+}c_i)}{+}\frac{\psi(\tau(c_i{-}b_i))}{\tau(c_i{-}b_i)}\Big) + 2\sum_{i=q+1}^{n}\frac{\psi(\tau)}{\tau}\Big\}. \tag{39}$$

Applying the inequality (79) (Appendix B) on the second sum in (39) delivers

$$U_{g_{n\text{-}\ell_1}} \le \inf_{\tau\ge 0}\Big\{p{+}\tau^2\sum_{i=1}^{p}a_i^2{+}\sum_{i=p+1}^{q}\Big(\frac{1}{\sqrt{2\pi}}\frac{1-(b_i{+}c_i)^2}{\tau(b_i{+}c_i)} + \frac{1}{\sqrt{2\pi}}\frac{1{-}(c_i{-}b_i)^2}{\tau(c_i{-}b_i)}{+}c_i\frac{2\psi(\tau)}{\tau}\Big){+}2(n-q)\frac{\psi(\tau)}{\tau}\Big\}.$$

$$\Leftrightarrow U_{g_{n\text{-}\ell_1}} \le \inf_{\tau\ge 0}\Big\{p{+}\tau^2\sum_{i=1}^{p}a_i^2{+}\sum_{i=p+1}^{q}\frac{1}{\sqrt{2\pi}}\frac{2c_i}{\tau}\Big(\frac{1}{c_i^2-b_i^2}-1\Big) + \Big(\sum_{i=p+1}^{q}c_i+(n-q)\Big)2\frac{\psi(\tau)}{\tau}\Big\}. \tag{40}$$

From the definitions of b_i (34b) and c_i (34c), we have

$$b_i \le \sum_{j\notin[l_i+1,l_i+d_i]} w_{ji} = 1-c_i, \tag{41}$$

due to the constraint $\sum_{j=0}^{J} w_{ji}=1$. Thus the second sum in (40) gives

$$\sum_{i=p+1}^{q}\frac{1}{\sqrt{2\pi}}\frac{2c_i}{\tau}\Big(\frac{1}{c_i^2-b_i^2}-1\Big) \le \sum_{i=p+1}^{q}\frac{4}{\sqrt{2\pi}\tau}\frac{c_i}{2c_i-1}(1-c_i) \le \frac{4\inf\{c_i\}}{\sqrt{2\pi}\tau(2\inf\{c_i\}-1)}\sum_{i=p+1}^{q}(1-c_i). \tag{42}$$

For simplicity, let us denote

$$\bar{a}_{n\text{-}\ell_1} = \sum_{i=1}^{p}a_i^2 \tag{43a}$$

$$\kappa_{n\text{-}\ell_1} = \frac{4\inf\{c_i\}}{\sqrt{2\pi}\tau(2\inf\{c_i\}-1)} \tag{43b}$$

$$\bar{s}_{n\text{-}\ell_1} = q-\sum_{i=p+1}^{q}c_i = p+\sum_{i=p+1}^{q}(1-c_i). \tag{43c}$$

Substituting the quantities of (43a), (43b), (43c), and inequality (42) in formula (40) delivers

$$U_{g_{n\text{-}\ell_1}} \le \inf_{\tau\ge 0}\Big\{\bar{a}_{n\text{-}\ell_1}\tau^2{+}(n{-}\bar{s}_{n\text{-}\ell_1})\frac{2\psi(\tau)}{\tau}{+}p{+}\kappa_{n\text{-}\ell_1}(\bar{s}_{n\text{-}\ell_1}{-}p)\Big\}. \tag{44}$$

Furthermore, we have

$$U_{g_{n\text{-}\ell_1}} \le \inf_{\tau\ge 0}\Big\{\bar{a}_{n\text{-}\ell_1}\tau^2{+}(n{-}\bar{s}_{n\text{-}\ell_1})\frac{2}{\sqrt{2\pi}}\frac{e^{-\frac{\tau^2}{2}}}{\tau} + \bar{s}_{n\text{-}\ell_1}{+}(\kappa_{n\text{-}\ell_1}{-}1)(\bar{s}_{n\text{-}\ell_1}{-}p)\Big\}. \tag{45}$$

To give a bound as a function of the given $\boldsymbol{x}$ and other related parameters, we can select a parameter $\tau>0$ to obtain an useful bound in (45). Setting $\tau=\sqrt{2\log(n/\bar{s}_{n\text{-}\ell_1})}$ gives

$$U_{g_{n\text{-}\ell_1}} \le 2\bar{a}_{n\text{-}\ell_1}\log\frac{n}{\bar{s}_{n\text{-}\ell_1}} + \frac{\bar{s}_{n\text{-}\ell_1}(1-\bar{s}_{n\text{-}\ell_1}/n)}{\sqrt{\pi\log(n/\bar{s}_{n\text{-}\ell_1})}} + \bar{s}_{n\text{-}\ell_1} + \delta_{n\text{-}\ell_1}, \tag{46}$$

where

$$\delta_{n\text{-}\ell_1} = (\kappa_{n\text{-}\ell_1}-1)(\bar{s}_{n\text{-}\ell_1}-p). \tag{47}$$

Eventually, applying inequality (75) (Appendix B) to the second term on the right side of inequality (46) gives

$$U_{g_{n\text{-}\ell_1}} \le 2\bar{a}_{n\text{-}\ell_1}\log\frac{n}{\bar{s}_{n\text{-}\ell_1}} + \frac{7}{5}\bar{s}_{n\text{-}\ell_1} + \delta_{n\text{-}\ell_1}. \tag{48}$$

As a result, the RAMSI bound requiring $m_{n\text{-}\ell_1}$ measurements in (32) is obtained. ■

B. The Bound Analysis

We analyze the relations as well as compare the characteristics of the bounds $m_{n\text{-}\ell_1}$ (32) with the classical CS bound (5) and the ℓ_1-ℓ_1 bound (7).

1) Bound Relations: We can establish the following consequent relations to the known bounds in Sec. II-A.

Corollary IV.5.1 (Bound relations)**.** *There are two consequent relations, the relation to the ℓ_1 bound and the relation to the ℓ_1-ℓ_1 bound.*

(a) *The RAMSI bound $m_{n\text{-}\ell_1}$ in* (32) *becomes the ℓ_1 bound m_{ℓ_1} in* (5) *when $\mathbf{W}_0{=}\mathbf{I}$ and $\mathbf{W}_j{=}\mathbf{0}$ for $j{\ge}1$, i.e,*

$$m_{n\text{-}\ell_1} \equiv m_{\ell_1} \ge 2s_0\log\frac{n}{s_0}+\frac{7}{5}s_0+1, \tag{49}$$

where $s_0{=}\mathrm{nnz}(\boldsymbol{x})$.

(b) *The RAMSI bound $m_{n\text{-}\ell_1}$ in* (32) *is computed to become the equal-weight bound $m_{\ell_1\text{-}\ell_1}$ in* (7) *when $\mathbf{W}_0{=}\mathbf{W}_1{=}\mathbf{I}/2$ and $\mathbf{W}_j{=}\mathbf{0}$ for $j\ge 2$, here we use $\mathbf{I}/2$ to ensure that $\mathbf{W}_0{+}\mathbf{W}_1{=}\mathbf{I}$, i.e.,*

$$m_{n\text{-}\ell_1} \equiv m_{\ell_1\text{-}\ell_1} \ge 2\bar{h}\log\frac{n}{\bar{s}_{\ell_1\text{-}\ell_1}}+\frac{7}{5}\bar{s}_{\ell_1\text{-}\ell_1}+1, \tag{50}$$

where $\bar{h}$ is given by (8b) *and $\bar{s}_{\ell_1\text{-}\ell_1}{=}(s_0+s_1)/2$.*

Proof: To obtain Relation (a) under the condition that $\mathbf{W}_0=\mathbf{I}$ and $\mathbf{W}_j=\mathbf{0}$ for $j\ge 1$, we have $\bar{s}_{n\text{-}\ell_1}=p=s_0$ and $\bar{a}_{n\text{-}\ell_1}=p=s_0$ based on the definitions in (43a), (43c). Consequently, from (47), $\delta_{n\text{-}\ell_1}=0$ due to $\bar{s}_{n\text{-}\ell_1}=p$. Using these $\bar{a}_{n\text{-}\ell_1}$, $\bar{s}_{n\text{-}\ell_1}$, $\delta_{n\text{-}\ell_1}$ for inequality (32) in Theorem IV.5 gives $m_{n\text{-}\ell_1}{\ge}2s_0\log(n/s_0){+}(7/5)s_0{+}1$, which is the ℓ_1 bound m_{ℓ_1} in (5).

To reach Relation (b) given that $\mathbf{W}_0{=}\mathbf{W}_1{=}\mathbf{I}/2$ and $\mathbf{W}_j{=}\mathbf{0}$ for $j\ge 2$, let us first denote two subsets, $\mathbb{I}_1$ and $\mathbb{I}_2$, as

$$\mathbb{I}_1 := \{i\in[p{+}1,q] : b_i{+}c_i{=}1, b_i{-}c_i{=}0\} \tag{51a}$$

$$\mathbb{I}_2 := \{i\in[p{+}1,q] : b_i{+}c_i{=}0, b_i{-}c_i{=}{-}1\}. \tag{51b}$$

From the definitions of b_i in (34b) and c_i in (34c) with $i \in [p{+}1, q]$, we observe that i is in either $\mathbb{I}_1$ or $\mathbb{I}_2$.

Replacing the above conditions in (37) leads to

$$\begin{aligned}\mathbb{E}[\mathrm{dist}^2(\boldsymbol{g}, \tau{\cdot}\partial g(\boldsymbol{x}))] &= p + \tau^2 \sum_{i=1}^{p} a_i^2 + \sum_{i\in\mathbb{I}_1} \mathcal{A}(\tau) \\ &+ \sum_{i\in\mathbb{I}_1} \mathcal{B}(0) + \sum_{i\in\mathbb{I}_2} \mathcal{A}(0) + \sum_{i\in\mathbb{I}_2} \mathcal{B}(-\tau) + 2\sum_{i=q+1}^{n} \mathcal{A}(\tau).\end{aligned} \tag{52}$$

Applying the equalities (77) and the inequalities (78a),(78b) from Appendix B on the inequality (52) gives

$$U_{\ell_1\text{-}\ell_1} \leq \inf_{\tau\geq 0}\Big\{ p + \tau^2 \sum_{i=1}^{p} a_i^2 + \sum_{i=p+1}^{q} \frac{1}{2} + \sum_{i=p+1}^{q} \frac{\psi(\tau)}{\tau} + 2\sum_{i=q+1}^{n} \frac{\psi(\tau)}{\tau} \Big\}, \tag{53}$$

which can be further elaborated to

$$U_{\ell_1\text{-}\ell_1} \leq \inf_{\tau\geq 0}\Big\{ \tau^2 \sum_{i=1}^{p} a_i^2 + \frac{1}{2}(p{+}q) + (2n{-}(p{+}q)) \frac{e^{\frac{-\tau^2}{2}}}{\sqrt{2\pi}\tau} \Big\}. \tag{54}$$

We are considering this case with one SI and equal weights, i.e., $d_i = 1, J = 1$, in (29), we thus derive $p + q = s_0 + s_1$ and combining with the definition (8b) gives $\sum_{i=1}^{p} a_i^2 = \bar{h}$. Let us denote $\bar{s}_{\ell_1\text{-}\ell_1} = (s_0{+}s_1)/2$ and set $\tau = \sqrt{2log(n/\bar{s}_{\ell_1\text{-}\ell_1})}$, we have

$$U_{\ell_1\text{-}\ell_1} \leq 2\bar{h} \log \frac{n}{\bar{s}_{\ell_1\text{-}\ell_1}} + \bar{s}_{\ell_1\text{-}\ell_1} + \frac{\bar{s}_{\ell_1\text{-}\ell_1}(1 - \bar{s}_{\ell_1\text{-}\ell_1}/n)}{\sqrt{2\pi \log(n/\bar{s}_{\ell_1\text{-}\ell_1})}}. \tag{55}$$

Applying (75) to the third term on the right side of (55) gives

$$U_{\ell_1\text{-}\ell_1} \leq 2\bar{h} \log \frac{n}{\bar{s}_{\ell_1\text{-}\ell_1}} + \frac{7}{5}\bar{s}_{\ell_1\text{-}\ell_1}. \tag{56}$$

Finally, we obtain the required measurement $m_{\ell_1\text{-}\ell_1}$

$$m_{\ell_1\text{-}\ell_1} \geq 2\bar{h} \log \frac{n}{\bar{s}_{\ell_1\text{-}\ell_1}} + \frac{7}{5}\bar{s}_{\ell_1\text{-}\ell_1} + 1. \tag{57}$$

It can be noted that $\bar{s}_{\ell_1\text{-}\ell_1} = s_0{+}\xi/2$ due to the definition of ξ (8a), i.e., the bounds $m_{\ell_1\text{-}\ell_1}$ in both (57) and (7) are identical as in (50).

■

2) Bound Comparisons: In this section, we theoretically evaluate and compare the RAMSI bound $m_{n\text{-}\ell_1}$ in (32) with the bound $m_{\ell_1\text{-}\ell_1}$ in (57). We first consider the last quantity $\delta_{n\text{-}\ell_1}$ in bound $m_{n\text{-}\ell_1}$ in (32). By the result of Lemma B.2 (Appendix B), we observe that $\delta_{n\text{-}\ell_1}$ is negative. In addition, we derive a simpler bound $\widetilde{m}_{n\text{-}\ell_1}$ in (58) for conveniently computing an approximation for bound $m_{n\text{-}\ell_1}$ (32) (Theory IV.5). For evaluating the bounds more easily, we introduce two looser bounds that are independent from the values of $\boldsymbol{x}, \boldsymbol{z}_j$. Using these bounds, we can theoretically figure out the characteristics of bound $m_{\ell_1\text{-}\ell_1}$ (57) and bound $\widetilde{m}_{n\text{-}\ell_1}$ in (58).

(a) **The simpler bound**. Bound $m_{n\text{-}\ell_1}$ in (32) approximately becomes

$$\widetilde{m}_{n\text{-}\ell_1} \geq 2\bar{a}_{n\text{-}\ell_1} \log \frac{n}{p} + \frac{7}{5}p + 1, \tag{58}$$

where $\bar{a}_{n\text{-}\ell_1}$ is defined by (43a).

(b) **The looser bounds**. Bound $\widetilde{m}_{n\text{-}\ell_1}$ in (58) and bound $m_{\ell_1\text{-}\ell_1}$ in (57) have corresponding looser bounds that are independent from the values of $\boldsymbol{x}, \boldsymbol{z}_j$, denoted by $\widehat{m}_{n\text{-}\ell_1}$ and $\widehat{m}_{\ell_1\text{-}\ell_1}$, given by

$$\widehat{m}_{n\text{-}\ell_1} \geq 2p \log \frac{n}{p} + \frac{7}{5}p + 1, \tag{59}$$

$$\widehat{m}_{\ell_1\text{-}\ell_1} \geq 2\rho \log \frac{n}{\bar{s}_{\ell_1\text{-}\ell_1}} + \frac{7}{5}\bar{s}_{\ell_1\text{-}\ell_1} + 1, \tag{60}$$

where p ($\leq \inf\{s_j\}$) is defined in Definition IV.2, $\rho = \min\{s_0, s_1\}$, and $\bar{s}_{\ell_1\text{-}\ell_1} = (s_0{+}s_1)/2$.

To reach the simpler bound in (58), we consider the quantity $\bar{s}_{n\text{-}\ell_1}$ (43c) in the bound $m_{n\text{-}\ell_1}$ (32). Let us recall that we have assumed that parameter ϵ is small and thus $c_i \approx 1$ (see Eq. (82) in Lemma B.2). Consequently, from (43c), $\bar{s}_{n\text{-}\ell_1} \approx p$ as well as $\delta_{n\text{-}\ell_1} \approx 0$ from (47). Thus bound $m_{n\text{-}\ell_1}$ (32) approximately becomes (58). The simpler bound $\widetilde{m}_{n\text{-}\ell_1}$ (58) is conveniently computed rather than the more complex bound $m_{n\text{-}\ell_1}$ (32) due to the fact that we only need to compute $\bar{a}_{n\text{-}\ell_1}$ and p. Furthermore, $\bar{a}_{n\text{-}\ell_1} \leq p \leq \inf\{s_j\}$ based on Eq. (43a) and Definition IV.2, that is, p in (58) is less than s_0 in (5) and $\bar{s}_{\ell_1\text{-}\ell_1}$ in (57). Consequently, bound $\widetilde{m}_{n\text{-}\ell_1}$ (58) of RAMSI are sharper than bound m_{ℓ_1} (5) and bound $m_{\ell_1\text{-}\ell_1}$ (57).

To obtain the two bounds in (59) and (60), we also consider the worst-case scenarios of both bound $\widetilde{m}_{n\text{-}\ell_1}$ (58) and bound $m_{\ell_1\text{-}\ell_1}$ (57), i.e., these looser bounds are independent from the values of $\boldsymbol{x}$, $\boldsymbol{z}_j$. We observe that the bounds $\widetilde{m}_{n\text{-}\ell_1}$ (58) and $m_{\ell_1\text{-}\ell_1}$ (57) depend on the values of $\boldsymbol{x}, \boldsymbol{z}_j$ via quantities $\bar{a}_{n\text{-}\ell_1}$ and $\bar{h}$, respectively. Bound $\widetilde{m}_{n\text{-}\ell_1}$ (58) and bound $m_{\ell_1\text{-}\ell_1}$ (57) are looser when the quantities, $\bar{a}_{n\text{-}\ell_1}$ and $\bar{h}$, take their maximum value. From the definitions of $\bar{a}_{n\text{-}\ell_1}$ (43a) and $\bar{h}$ (8b), we observe that $\bar{a}_{n\text{-}\ell_1} \leq p$ and $\bar{h} \leq \min\{s_0, s_1\}$. As a result, we derive the looser bounds $\widehat{m}_{n\text{-}\ell_1}$ in (59) and $\widehat{m}_{\ell_1\text{-}\ell_1}$ in (60).

Considering bounds (59) and (60), they reveal that additional SI signals taken in RAMSI will reduce $\widehat{m}_{n\text{-}\ell_1}$ (59) due to the fact that $p \leq \inf\{s_j\}$. Clearly, $p \leq \rho$, therefore, $\widehat{m}_{n\text{-}\ell_1} \leq \widehat{m}_{\ell_1\text{-}\ell_1}$. In particular, for bound $\widehat{m}_{\ell_1\text{-}\ell_1}$ (60), if the SI $\boldsymbol{z}_1$ is not good enough, i.e., $s_1 \gg s_0$, bound $\widehat{m}_{\ell_1\text{-}\ell_1}$ (60) would be higher than bound m_{ℓ_1} (5) due to $\bar{s}_{\ell_1\text{-}\ell_1} \gg s_0$. This analysis theoretically explains the drawback of the ℓ_1-ℓ_1 minimization as well as the advantages of RAMSI.

V. Experimental Results

We present numerical experiments that demonstrate the established bounds and performance of the RAMSI algorithm on sparse signals with different characteristics. We also analyze how the SI qualities have effects on the measurements of RAMSI. In addition, we test RAMSI on correlated feature histogram vectors as sparse sources, which are extracted from a multiview image database [31].

A. Experimental Setup

In this experiment, we consider the reconstruction of a generated sparse source $\boldsymbol{x}$ given some known SI signals, $\boldsymbol{z}_1$, $\boldsymbol{z}_2$, $\boldsymbol{z}_3$. We generate $\boldsymbol{x}$ with $n = 1000$, and support $s_0 = 128$, that is, 128 of 1000 elements are nonzeros, which are

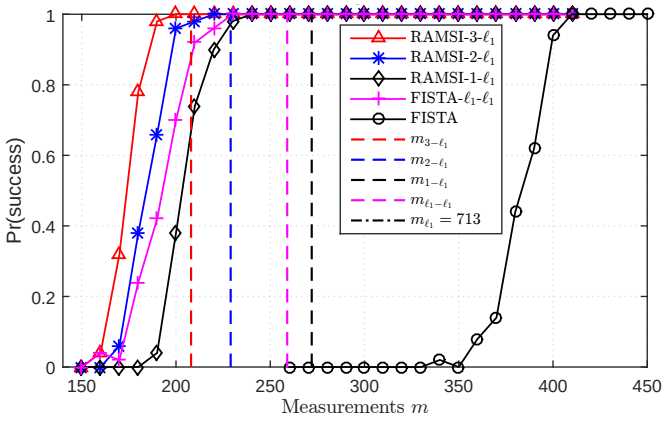
generated from the standard i.i.d. Gaussian distribution. Firstly, we would like to consider a scenario where the SI signals $\boldsymbol{z}_j$ are well-correlated to the source $\boldsymbol{x}$ leading to a small number of nonzeros in $\boldsymbol{x}-\boldsymbol{z}_j$. In this scenario, the SI signals, $\boldsymbol{z}_1$, $\boldsymbol{z}_2$, $\boldsymbol{z}_3$, are generated satisfying $s_j = \text{nnz}(\boldsymbol{x}-\boldsymbol{z}_j) = 64$. Moreover, similar to the setup in [18], [19], a parameter is controlling the number of positions of nonzeros for which both $\boldsymbol{x}$ and $\boldsymbol{x}-\boldsymbol{z}_j$ coincide. Let us denote this number by r_j. For instance, if $r_j = 51$, $\boldsymbol{x}$ has 51 nonzero positions that coincide with 51 nonzero positions of $\boldsymbol{x}-\boldsymbol{z}_j$. This incurs a significant error between the source and the SI, which is given by $||\boldsymbol{z}_j-\boldsymbol{x}||_2/||\boldsymbol{x}||_2 \approx 0.56$.

To assess the performance of the algorithm and the bounds when the quality of SI is poor, we generate SI signals that are not well-correlated to the source $\boldsymbol{x}$. Their poor qualities are expressed via higher values of s_j, for example, $s_j = 256$ and $s_j = 352$ are tested in this experiment. Furthermore, we set $r_j = 128$, namely, 128 nonzero positions of $\boldsymbol{x}$ coincide with 128 nonzero positions out of the total 256 or 352 nonzero positions of the SI signals, $\boldsymbol{z}_1$, $\boldsymbol{z}_2$, $\boldsymbol{z}_3$. This leads to very high errors, e.g., $||\boldsymbol{z}_j-\boldsymbol{x}||_2/||\boldsymbol{x}||_2 \approx 1.12$ for $s_j = 256$, and the supports s_j of $\boldsymbol{x}-\boldsymbol{z}_j$ are much higher than that of $\boldsymbol{x}$. It may be noted that we set all s_j equal to avoid numerous scenarios so that comparisons are carried out under similar SI qualities.

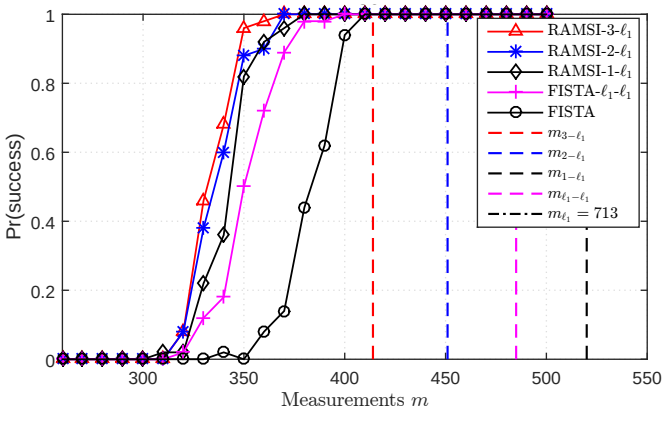
Furthermore, we experimentally conduct the reconstruction of multiview feature histograms as sparse sources used in multiview object recognition. Given an image, its feature histogram is formed as in Sec. III-A. The size of the vocabulary tree [33] depends on the value of k and the number of hierarchies, for example, if $k = 10$ and 3 hierarchies, $n = 1000$ vocabularies as 1000-D. Because of the small number of features in a single image, the histogram vector $\boldsymbol{x}$ is indeed sparse and compressible. Therefore, $\boldsymbol{x}$ is first projected into the compressed vector $\boldsymbol{y}$ which is to be sent to the decoder. At the joint decoder, we take $\boldsymbol{x}$ to be reconstructed given some already decoded histograms of neighbor views, e.g., $\boldsymbol{z}_1$, $\boldsymbol{z}_2$, $\boldsymbol{z}_3$. In this work, we use the COIL-100 data set [31] containing multiview images of 100 small objects with different angle degrees. In order to ensure that our experimental setup reflects a realistic scenario, we randomly select the 4 neighbor views (3 neighbors as SI signals) of objects 16 (Fig. 4(a)) and 60 (Fig. 4(b)) over 72 views captured through 360 degrees in the COIL-100 [31] multiview database. Specifically, the four neighbor views are assigned to $\boldsymbol{z}_1$, $\boldsymbol{x}$, $\boldsymbol{z}_2$, $\boldsymbol{z}_3$, respectively, of which the third SI $\boldsymbol{z}_3$ is set to the furthest neighbor of the source $\boldsymbol{x}$.

B. Performance Evaluation

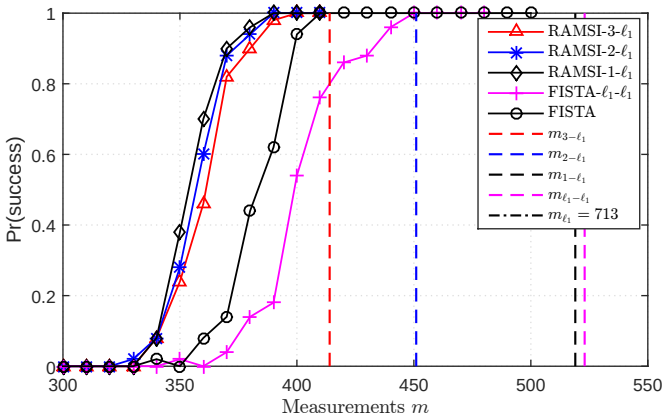
1) Synthetic Signal Reconstruction: We evaluate and compare the obtained bounds $m_{n\text{-}\ell_1}$ (32), $m_{\ell_1\text{-}\ell_1}$ (7), and m_{ℓ_1} (5) along with the reconstruction accuracy of RAMSI. For evaluating the reconstruction of RAMSI, we introduce a probability of successful recovery, denoted as $\Pr(\text{success})$. For a fixed dimension or number of measurements m, $\Pr(\text{success})$ is the number of times, in which the source $\boldsymbol{x}$ is recovered as $\widehat{\boldsymbol{x}}$ with an error $||\widehat{\boldsymbol{x}}-\boldsymbol{x}||_2/||\boldsymbol{x}||_2 \leq 10^{-2}$, divided by the total number of 100 trials (each trial considered different generated $\boldsymbol{x}, \boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3, \boldsymbol{\Phi}$).



(a) $s_0 = 128$, $s_1 = s_2 = s_3 = 64$



(b) $s_0 = 128$, $s_1 = s_2 = s_3 = 256$



(c) $s_0 = 128$, $s_1 = s_2 = s_3 = 352$

Fig. 2. Successful probabilities and measurement bounds and of the original 1000-D $\boldsymbol{x}$ vs. number of measurements m for the RAMSI using 1,2,3 SI signals.

Let RAMSI-J-ℓ_1 denote the RAMSI reconstruction, where J indicates the number of SI signals, e.g., RAMSI-1-ℓ_1, RAMSI-2-ℓ_1, RAMSI-3-ℓ_1 are RAMSI using 1 ($\boldsymbol{z}_1$), 2 ($\boldsymbol{z}_1$, $\boldsymbol{z}_2$), and 3 ($\boldsymbol{z}_1$, $\boldsymbol{z}_2$, $\boldsymbol{z}_3$) SI signals, respectively. Let FISTA-ℓ_1-ℓ_1 denote the ℓ_1-ℓ_1 CS reconstruction with one SI ($\boldsymbol{z}_1$) [18]. The existing FISTA [4] and FISTA-ℓ_1-ℓ_1 reconstructions [18] are used for comparison. Let $m_{3\text{-}\ell_1}$, $m_{2\text{-}\ell_1}$, $m_{1\text{-}\ell_1}$ (32), $m_{\ell_1\text{-}\ell_1}$ (7), and m_{ℓ_1} (5) denote the corresponding bounds of RAMSI-3-ℓ_1, RAMSI-2-ℓ_1, RAMSI-1-ℓ_1, FISTA-ℓ_1-ℓ_1, and FISTA. The original source 1000-D $\boldsymbol{x}$ is compressed into different lower-dimensions $\boldsymbol{y}$. We assess the bounds and the accuracy of a reconstructed $\widehat{\boldsymbol{x}}$ versus the $\boldsymbol{x}$ via the successful rate versus

the number of measurements m, which is the dimensionality of $\boldsymbol{y}$. RAMSI (in Algorithm 1) thus has $n = 1000$, $m < n$, $J = 1, 2, 3$ and we set $\epsilon = 10^{-5}$, $\lambda = 10^{-5}$.

Figure 2(a) depicts the bounds as well as successful probabilities of recovering $\boldsymbol{x}$ versus the number of measurements m, given the SI signals with $s_j = 64$. The figure shows clearly that RAMSI-3-ℓ_1 gives the sharpest bound $m_{3\text{-}\ell_1}$ and the best successful probability. Furthermore, the performance of RAMSI-2-ℓ_1 is higher than those of RAMSI-1-ℓ_1 and FISTA-ℓ_1-ℓ_1. In particular, FISTA-ℓ_1-ℓ_1 outperforms RAMSI-1-ℓ_1 in this scenario of the quite good SI quality. We can infer that in these kinds of sparse sources, the equal weights of FISTA-ℓ_1-ℓ_1 gain more than the adaptive weights of RAMSI-1-ℓ_1. This can be explained by looking at bound $m_{1\text{-}\ell_1}$ (32) of RAMSI-1-ℓ_1 and bound $m_{\ell_1\text{-}\ell_1}$ (7) of FISTA-ℓ_1-ℓ_1. It is clear that $\bar{a}_{1\text{-}\ell_1}$ (43a) is greater than $\bar{h}$ (8b). Thus, combining with the small number s_1 in case of the good SI $\boldsymbol{z}_1$ resulting in small $\bar{s}_{\ell_1\text{-}\ell_1}$, the bound as well as the successful probability of FISTA-ℓ_1-ℓ_1 are better than those of RAMSI-1-ℓ_1 shown by magenta and black lines in Fig. 2(a). We can conclude that exploiting multiple SI signals gives the best performance and when dealing with only one SI for this kind of scenarios, we may choose equal weights.

Moreover, Figs. 2(b) and 2(c) present the bounds and the reconstruction performance versus the number of measurements when the SI signals are less correlated with the signal of interest, e.g., $s_j = 256$ and $s_j = 352$. All RAMSI configurations outperform FISTA and FISTA-ℓ_1-ℓ_1. The performance of RAMSI-1-ℓ_1 is better than that of FISTA-ℓ_1-ℓ_1, where the same SI is exploited. More interestingly, in Fig. 2(c), we observe that the accuracy of FISTA-ℓ_1-ℓ_1 is worse than that of FISTA, i.e., the SI $\boldsymbol{z}_1$ does not help, however, RAMSI-1-ℓ_1 still outperforms FISTA. These results may highlight the drawback of ℓ_1-ℓ_1 minimization when the SI quality is not good enough. Though encountering poor SI signals, all the RAMSI versions achieve better results due to performing the proposed adaptive-weighted n-ℓ_1 minimization. Specially, we observe that the performances of RAMSI-2-ℓ_1 and RAMSI-3-ℓ_1 are slightly worse than that of RAMSI-1-ℓ_1. These small penalties can be explained by the not so good SI signals which interfere in the iterative update process. We will see more on these occurrences in the following analysis on the SI quality.

2) SI Quality-Dependence Analysis: This subsection considers how the SI qualities impact the number of measurements required to successfully reconstruct the original source. To this end, we perform the proposed RAMSI algorithm on variations of SI qualities, $\boldsymbol{z}_1$, $\boldsymbol{z}_2$, $\boldsymbol{z}_3$, through different $s_1 = s_2 = s_3$ for reconstructing the source $\boldsymbol{x}$. For a fixed value of $s_1 = s_2 = s_3$, we measure the number of measurements of RAMSI-3-ℓ_1, RAMSI-2-ℓ_1, RAMSI-1-ℓ_1, FISTA-ℓ_1-ℓ_1, FISTA as well as the corresponding bounds, $m_{3\text{-}\ell_1}$, $m_{2\text{-}\ell_1}$, $m_{1\text{-}\ell_1}$ (32), $m_{\ell_1\text{-}\ell_1}$ (7), m_{ℓ_1} (5). In this experiment, we get the number of measurements once the algorithms achieve $\Pr(\text{success}) \geq 0.98$. We conduct the experiment for the range of $s_1 = s_2 = s_3$ from 20 to 400, where SI qualities with supports greater than 400 are too poor to be considered. We evaluate the performance on the source $\boldsymbol{x}$ with $s_0 = 128$, which is generated as in Sec. V-A.

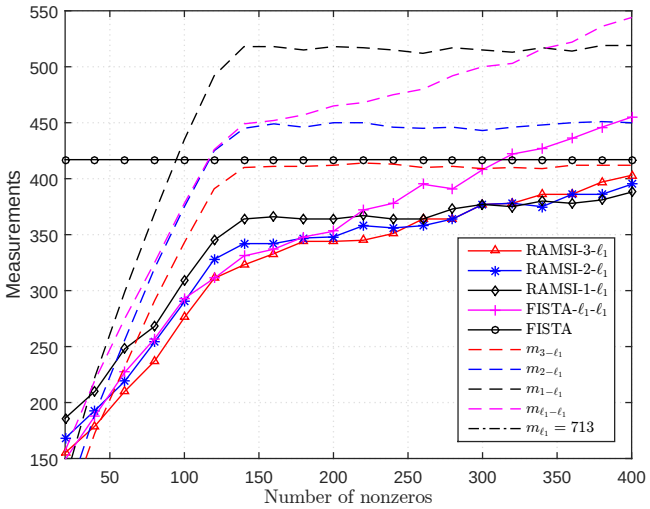


Fig. 3. Number of measurements vs. number of nonzeros $s_1 = s_2 = s_3$ of SI signals given the source $s_0 = 128$ for the RAMSI using 1,2,3 SI signals.

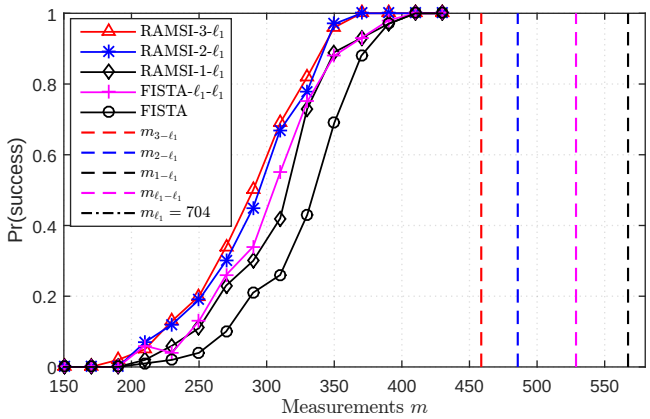
Figure 3 illustrates the number of measurements versus the number of nonzeros as values of $s_1 = s_2 = s_3$ of SI signals representing the quality variations. Evidently, we can observe that the reconstructions with SI signals significantly reduce the number of measurements until the quality of SI signals meets the number of nonzeros s_0 of the source $\boldsymbol{x}$. When $s_j > s_0$, the performance and the measurement bound of FISTA-ℓ_1-ℓ_1 are linearly increased according to increasing number of s_1. Specifically, at $s_1 = 315$ in Fig. 3, the FISTA-ℓ_1-ℓ_1 performance is worse than that of FISTA. This illustration along with the performance in Fig. 2(c) once again show the drawback of FISTA-ℓ_1-ℓ_1.

As shown in Fig. 3, the performance of RAMSI, both in terms of the theoretical bounds and the practical results, is robust against poor-quality SI signals. The theoretical bounds are sharper when the number of SI signals increases and remains constant when the number of nonzero values s_j increase (indicating increasing-poor SI quality). Considering the performance of RAMSI, with $s_j > 300$, the number of measurements of RAMSI-3-ℓ_1 and RAMSI-2-ℓ_1 are slightly worse than those of RAMSI-1-ℓ_1 and approaching to the measurements of FISTA as shown in Fig. 3. We can also see these occurrences in Fig. 2(c), where $s_j = 352$ indicating poor SI signals. To avoid these circumstances, we may find a solution to adaptively select the best performance among the RAMSI configurations. For instance, when we have only one SI, we may use equal weights over the adaptive weights for the good SI qualities such as the number of nonzeros less than 210 shown in Fig. 3. It can be re-emphasized that the advantage of RAMSI is that we can control it via the weights and parameters to enable RAMSI to deal with such poor SI signals. Therefore, we ensure that the RAMSI performance is not worse than FISTA by weighing dominantly on the source rather than the poor SI signals during the reconstruction.

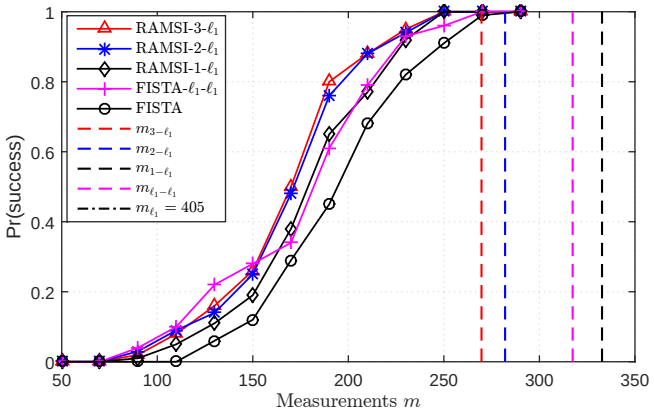
3) Feature Histogram Reconstruction: In this subsection, we evaluate the RAMSI configurations and the bounds on sparse sources extracted from the multiview image database [31]. Figure 4 presents the performance of RAMSI with 1, 2, 3 SI signals, FISTA [4], FISTA-ℓ_1-ℓ_1 [18] in terms of

(a) Object 16

(b) Object 60



(c) Reconstruction performance of Object 16



(d) Reconstruction performance of Object 60

Fig. 4. Successful probabilities and measurement bounds vs. number of measurements m for objects 16 and 60 in the COIL-100 [31].

the successful probabilities and theoretical bounds versus the number of measurements. Similar to the generated signals (Sec. V-B1), RAMSI (in Algorithm 1) has $n=1000$, $m<n$, $J=1,2,3$ and we set $\epsilon=10^{-5}$, $\lambda=10^{-5}$. Pr(success) is also the number of times divided by the total number of 100 trials (each trial considered randomly selecting 4 neighbor views for objects 16 (Fig. 4(a)) and 60 (Fig. 4(b)). Figure 4(c) and Figure 4(d) show clearly that RAMSI significantly improves the reconstruction accuracy for the representative objects 16 (Fig. 4(a)) and 60 (Fig. 4(b)), respectively. Furthermore, the results using 3 SI signals, RAMSI-3-ℓ_1, provide for highest accuracy and the performance of RAMSI-2-ℓ_1 is higher than that of RAMSI-1-ℓ_1. It can be noted that RAMSI-3-ℓ_1 gains not much over RAMSI-2-ℓ_1 since we set the third SI as the furthest neighbor as in Sec. V-A. These results are consistent with what we have shown in the performances of the generated sparse signals (Secs. V-B1+V-B2).

VI. Conclusion

This paper presented the RAMSI algorithm and theoretically established its performance bounds. The proposed algorithm incorporates multiple SI signals in the problem of sparse signal recovery and iteratively weights the various SI signals so as to optimize the reconstruction performance. The obtained bounds confirm the advantage of RAMSI in utilizing multiple SI signals to significantly reduce the number of measurements and to deal with variations in the quality of the SI signals. We experimentally assessed the established bounds and the performance of RAMSI against state-of-the-art methods using both synthetic and real-life sparse signals. The results showed that the measurement bounds of RAMSI are sharper and that RAMSI outperformed the conventional ℓ_1 CS and the recent ℓ_1-ℓ_1 minimization reconstruction methods. Moreover, RAMSI can efficiently incorporate multi-hypothesis SI signals, where the higher number of SI signals, the higher performance of the proposed RAMSI algorithm.

Appendix A
Proof of Proposition III.1

We shall compute the proximal operator $\Gamma_{\frac{1}{L}g}(\boldsymbol{x})$ (4) with $g(\boldsymbol{x})=\lambda\sum_{j=0}^{J}||\mathbf{W}_j(\boldsymbol{x}-\boldsymbol{z}_j)||_1$. From (4), $\Gamma_{\frac{1}{L}g}(\boldsymbol{x})$ is expressed by:

$$\Gamma_{\frac{1}{L}g}(\boldsymbol{x})=\arg\min_{\boldsymbol{v}\in\mathbb{R}^n}\left\{\frac{\lambda}{L}\sum_{j=0}^{J}||\mathbf{W}_j(\boldsymbol{v}-\boldsymbol{z}_j)||_1+\frac{1}{2}||\boldsymbol{v}-\boldsymbol{x}||_2^2\right\}. \quad (61)$$

We note that both terms in (61) are separable in $\boldsymbol{v}$ and thus we can minimize each element v_i of $\boldsymbol{v}$ individually as

$$\Gamma_{\frac{1}{L}g}(x_i)=\arg\min_{v_i\in\mathbb{R}}\left\{h(v_i)=\frac{\lambda}{L}\sum_{j=0}^{J}w_{ji}|v_i-z_{ji}|+\frac{1}{2}(v_i-x_i)^2\right\}. \quad (62)$$

We consider $\partial h(v_i)/\partial v_i$. Without loss of generality, we assume $-\infty\le z_{0i}\le z_{1i}\le\ldots\le z_{Ji}\le\infty$. For convenience, let us denote $z_{-1i}=-\infty$ and $z_{J+1i}=\infty$. When v_i is located in one of the intervals, we suppose $v_i\in(z_{li},z_{l+1i})$ with $-1\le l\le J$, where $\partial h(v_i)$ exists. Taking the derivative of $h(v_i)$ in (z_{li},z_{l+1i}) delivers

$$\frac{\partial h(v_i)}{\partial v_i}=\frac{\lambda}{L}\sum_{j=0}^{J}w_{ji}\text{sign}(v_i-z_{ji})+(v_i-x_i), \quad (63)$$

where sign(.) is a sign function. In addition, let $\mathfrak{b}(.)$ denote a boolean function, i.e., $\mathfrak{b}(l<j)=1$ if $l<j$, otherwise $\mathfrak{b}(l<j)=0$. Consequently, $\text{sign}(v_i-z_{ji})=(-1)^{\mathfrak{b}(l<j)}$ and from (63), we rewrite:

$$\frac{\partial h(v_i)}{\partial v_i}=\frac{\lambda}{L}\sum_{j=0}^{J}w_{ji}(-1)^{\mathfrak{b}(l<j)}+(v_i-x_i). \quad (64)$$

When setting $\partial h(v_i)/\partial v_i=0$ to minimize the $h(v_i)$, we derive:

$$v_i=x_i-\frac{\lambda}{L}\sum_{j=0}^{J}w_{ji}(-1)^{\mathfrak{b}(l<j)}. \quad (65)$$

This v_i (65) is only valid in (z_{li},z_{l+1i}), i.e.,

$$z_{li}+\frac{\lambda}{L}\sum_{j=0}^{J}w_{ji}(-1)^{\mathfrak{b}(l<j)}<x_i<z_{l+1i}+\frac{\lambda}{L}\sum_{j=0}^{J}w_{ji}(-1)^{\mathfrak{b}(l<j)}. \quad (66)$$

In case of that x_i does not belong to alike intervals in (66), i.e.,

$$z_{li}+\frac{\lambda}{L}\sum_{j=0}^{J}w_{ji}(-1)^{\mathfrak{b}(l-1<j)}\le x_i\le z_{li}+\frac{\lambda}{L}\sum_{j=0}^{J}w_{ji}(-1)^{\mathfrak{b}(l<j)}. \quad (67)$$

We will prove that $h(v_i)$ (62) is minimum when $v_i = z_{li}$ in the following Lemma A.1.

Lemma A.1. *Given x_i belonging to the intervals represented in* (67)*, the $h(v_i)$ in* (62) *is minimum when $v_i = z_{li}$.*

Proof: To prove it, we re-express the $h(v_i)$ (62) by:

$$h(v_i)=\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}|(v_i-z_{li})-(z_{ji}-z_{li})|+\frac{1}{2}((v_i-z_{li})-(x_i-z_{li}))^2. \quad (68)$$

Applying a simple inequality $|a-b| \geq |a|-|b|$, where $a, b \in \mathbb{R}$, to the first term and expanding the second term in (68), we obtain:

$$\begin{aligned} h(v_i)\geq \frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}|v_i-z_{li}|-\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}|z_{ji}-z_{li}|+ \\ \frac{1}{2}(v_i-z_{li})^2-(v_i-z_{li})(x_i-z_{li})+\frac{1}{2}(x_i-z_{li})^2. \end{aligned} \quad (69)$$

It can be noted that $-(v_i-z_{li})(x_i-z_{li}) \geq -|v_i-z_{li}||x_i-z_{li}|$. Thus the (69) inequality is equivalent to:

$$\begin{aligned} h(v_i)\geq |v_i-z_{li}|\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}-|v_i-z_{li}||x_i-z_{li}|+\frac{1}{2}(v_i-z_{li})^2 \\ -\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}|z_{ji}-z_{li}|+\frac{1}{2}(x_i-z_{li})^2. \end{aligned} \quad (70)$$

Without difficulty, from the expression (67), we get:

$$-\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}\leq x_i-z_{li}\leq \frac{\lambda}{L}\sum_{j=0}^{J} w_{ji} \Leftrightarrow |x_i-z_{li}|\leq \frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}. \quad (71)$$

Eventually, we observe that the part including v_i in the right hand side of the $h(v_i)$ inequality in (70) is

$$|v_i - z_{li}|\Big(\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}-|x_i - z_{li}|\Big)+\frac{1}{2}(v_i-z_{li})^2. \quad (72)$$

With (71), the expression (72) is minimum when $v_i = z_{li}$. Therefore, we deduce that $h(v_i)$ (62) is minimum when $v_i = z_{li}$. ∎

In summary, from (65) with conditions in (66), (67) and the Lemma A.1, we obtain:

$$\Gamma_{\frac{1}{L}g}(x_i)=\begin{cases} x_i-\frac{\lambda}{L}\sum_{j=0}^{J} w_{ji}(-1)^{\mathfrak{b}(l<j)} & \text{if (66);} \\ z_{li} & \text{if (67);} \end{cases} \quad (73)$$

Appendix B
Useful Functions and Inequalities

Let us consider two functions relating to the normal standard distribution, the probability density of the normal distribution $\mathcal{N}(0,1)$ with zero-mean and unit variance $\psi(x)$ given by:

$$\psi(x) := \frac{1}{\sqrt{2\pi}}e^{-x^2/2}. \quad (74)$$

The meaningful inequality [19] of a specific log expression is also used in our bound computation, expressed as follows

$$\frac{(1-x^{-1})}{\sqrt{\pi\log(x)}} \leq \frac{1}{\sqrt{2\pi}} \leq \frac{2}{5}, \quad (75)$$

for all $x > 1$.

Moreover, we will use two formulations which are frequently used in our computations denoted as $\mathcal{A}(x)$ and $\mathcal{B}(x)$ same as in [19]:

$$\mathcal{A}(x) := \frac{1}{\sqrt{2\pi}}\int_x^{\infty}(v-x)^2e^{-v^2/2}\mathrm{d}v \quad (76a)$$

$$\mathcal{B}(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}(v-x)^2e^{-v^2/2}\mathrm{d}v. \quad (76b)$$

Without difficulty, if $x = 0$ we obtain

$$\mathcal{A}(0) = \mathcal{B}(0) = 1/2. \quad (77)$$

For $x \neq 0$, we have useful inequalities [19] expressed by:

$$\mathcal{A}(x) \leq \begin{cases} \psi(x)/x, & x > 0 \\ x^2+1, & x < 0 \end{cases} \quad (78a)$$

$$\mathcal{B}(x) \leq \begin{cases} -\psi(x)/x, & x < 0 \\ x^2+1, & x > 0. \end{cases} \quad (78b)$$

Lemma B.1. *Given $x \in (0,1]$ and $\tau > 0$, for $\psi(x)$ given in* (74)*, we have*

$$\frac{\psi(\tau x)}{\tau x} \leq \frac{1}{\sqrt{2\pi}}\frac{1-x^2}{\tau x}+x\frac{\psi(\tau)}{\tau}. \quad (79)$$

Proof: We have the left hand of the inequality (79) is expressed based on (74) as

$$\frac{\psi(\tau x)}{\tau x} = \frac{1}{\sqrt{2\pi}}\frac{e^{-\tau^2x^2/2}}{\tau x}. \quad (80)$$

Applying the Bernoulli's inequality on $e^{-\tau^2x^2/2}$ obtains

$$e^{-\tau^2x^2/2} = \Big(1+(e^{-\tau^2/2}-1)\Big)^{x^2} \leq 1+x^2(e^{-\tau^2/2}-1), \quad (81)$$

where $0 < x \leq 1$ and $(e^{-\tau^2/2}-1) > -1$ given $\tau > 0$.

Combining (81) and (80), we derive the inequality (79). ∎

Lemma B.2. *Given a sparse signal $\boldsymbol{x}\in\mathbb{R}^n$ with $\mathrm{nnz}(\boldsymbol{x})=s_0$ and J SI signals, $\boldsymbol{z}_j\in\mathbb{R}^n$ with $\mathrm{nnz}(\boldsymbol{x}-\boldsymbol{z}_j)=s_j$, considering bound $m_{n\text{-}\ell_1}$ in* (32)*, the last quantity $\delta_{n\text{-}\ell_1}$* (47) *is negative.*

Proof: Let us recall the last quantity in $m_{n\text{-}\ell_1}$ (32), $\delta_{n\text{-}\ell_1} = (\kappa_{n\text{-}\ell_1}-1)(\bar{s}_{n\text{-}\ell_1}-p)$. Based on the definition of $\bar{s}_{n\text{-}\ell_1}$ (43c), clearly $(\bar{s}_{n\text{-}\ell_1}-p) > 0$ due to $c_i < 1$ (34c).

We consider $\kappa_{n\text{-}\ell_1}$ (43b), wherein c_i is derived from (31) and (34c) as

$$c_i = d_i\Big(d_i + \sum_{j\notin[l_i+1,l_i+d_i]}\frac{\epsilon}{|x_i - z_{ji}| + \epsilon}\Big)^{-1}. \quad (82)$$

We have set the pretty small parameter ϵ, thus from (82) $c_i \approx 1$. As a result, $\kappa_{n\text{-}\ell_1}$ (43b) is approximately obtained by

$$\kappa_{n\text{-}\ell_1} \approx \frac{4}{\sqrt{2\pi}\tau} \approx \frac{2}{\sqrt{\pi\log(n/\bar{s}_{n\text{-}\ell_1})}}, \quad (83)$$

due to the setting $\tau = \sqrt{2\log(n/\bar{s}_{n\text{-}\ell_1})}$. We observe that $\kappa_{n\text{-}\ell_1} < 1$ if $\bar{s}_{n\text{-}\ell_1}/n < 0.28$. From (43c), $\bar{s}_{n\text{-}\ell_1} \approx p \leq \inf\{s_j\}$ (Definition IV.2). We have assumed that our input source signal is supposed to be sparse with certain degree. In case of the source signal is not sparse enough and also all SI signals are very bad qualities, in other words, $\bar{s}_{n\text{-}\ell_1} \approx p \leq \inf\{s_j\}$ is relative high, the recovery algorithm would not work well as

expected. To show this statement, look at the looser bound $\widehat{m}_{n\text{-}\ell_1}$ (59) for RAMSI or the similar form bound m_{ℓ_1} (5) without SI signals, we refer to Lemma B.3. We prove for $\widehat{m}_{n\text{-}\ell_1}$ (59) in Lemma B.3 that if the sparse degree p/n is less than 0.23, bound $\widehat{m}_{n\text{-}\ell_1}$ (59) is less than the source dimension n. Otherwise, bound $\widehat{m}_{n\text{-}\ell_1}$ (59) is higher than n, i.e., the recovery is useless in this context. Therefore, in our feasible context, i.e., $p/n < 0.23$ or $\bar{s}_{n\text{-}\ell_1}/n < 0.23$, we get $(\kappa_{n\text{-}\ell_1} - 1) < 0$ and combine with $(\bar{s}_{n\text{-}\ell_1} - p) > 0$ to conclude the quantity $\delta_{n\text{-}\ell_1} < 0$. ∎

Lemma B.3. *Given a sparse signal $\boldsymbol{x}\in\mathbb{R}^n$ with $\mathrm{nnz}(\boldsymbol{x})=s_0$ and J SI signals, $\boldsymbol{z}_j\in\mathbb{R}^n$ with $\mathrm{nnz}(\boldsymbol{x}-\boldsymbol{z}_j)=s_j$, considering bound $\widehat{m}_{n\text{-}\ell_1}$ (59), if the sparse degree $p/n < 0.23$, bound $\widehat{m}_{n\text{-}\ell_1}$ (59) is less than the source dimension n, otherwise $\widehat{m}_{n\text{-}\ell_1} > n$.*

Proof: Let us recall bound $\widehat{m}_{n\text{-}\ell_1}$ (59) and suppose the bound satisfies the condition $m<n$ as

$$2p\log\frac{n}{p}+\frac{7}{5}p<n. \tag{84}$$

Let $x\in\mathbb{R}$ denote $x=n/p$, substituting x for the inequality (84) gives

$$\begin{aligned} &2\log x+\frac{7}{5}<x \\ \Leftrightarrow\ &f(x)=x-2\log x+\frac{7}{5}>0, \end{aligned} \tag{85}$$

where $x>1$. From (85), we find an extremum at $x=2$ due to $f'(x)=(x-2)/x$. Without difficulty, we can find the condition of $x>4.33$ or $p/n<0.23$ to get the inequality (85) as the proof. ∎

REFERENCES

[1] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] ——, "For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution," *Communications on Pure and Applied Math*, vol. 59, no. 6, pp. 797–829, 2006.

[3] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Apr. 2006.

[4] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2(1), pp. 183–202, 2009.

[5] E. Candès, M. B. Wakin, , and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 877–905, 2008.

[6] M. S. Asif and J. Romberg, "Fast and accurate algorithms for re-weighted ℓ_1-norm minimization," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5905–5916, Dec. 2013.

[7] N. Vaswani, "Stability (over time) of modified-cs for recursive causal sparse reconstruction," in *Allerton Conf. Communications, Control, and Computing*, Monticello, Illinois, USA, Sep. 2010.

[8] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4595–4607, Sep. 2010.

[9] M. P. Friedlander, H. Mansour, R. Saab, and O. Yilmaz, "Recovering compressively sampled signals using partial support information," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, p. 11221134, Feb. 2012.

[10] J. Zhan and N. Vaswani, "Robust pca with partial subspace knowledge," in *IEEE Intern. Symposium on Information Theory (ISIT)*, Hawaii, USA, Jun. 2010.

[11] ——, "Time invariant error bounds for modified-cs-based sparse signal sequence recovery," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1389–1409, Mar. 2015.

[12] M. A. Khajehnejad, W. Xu, A. S. Avestimehr, and B. Hassibi, "Improving the thresholds of sparse recovery: An analysis of a two-step reweighted basis pursuit algorithm," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 5116–5128, Sep. 2015.

[13] L. Weizman, Y. C. Eldar, and D. B. Bashat, "Compressed sensing for longitunal MRI: An adaptive-weighted approach," *Medical Physics*, vol. 42, no. 9, pp. 5195–5207, 2015.

[14] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational Mathematics*, vol. 12, no. 6, pp. 805–849, 2012.

[15] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp, "Living on the edge: phase transitions in convex programs with random data," *Information and Inference*, vol. 3, no. 3, pp. 224–294, 2014.

[16] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing," ArXiv e-prints, Jan. 2009.

[17] M. Duarte, M. Wakin, D. Baron, S. Sarvotham, and R. Baraniuk, "Measurement bounds for sparse signal ensembles via graphical models," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4280–4289, Jul. 2013.

[18] J. F. Mota, N. Deligiannis, and M. R. Rodrigues, "Compressed sensing with side information: Geometrical interpretation and performance bounds," in *IEEE Global Conf. on Signal and Information Processing*, Austin, Texas, USA, Dec. 2014.

[19] ——, "Compressed sensing with prior information: Optimal strategies, geometry, and bounds," ArXiv e-prints, Aug. 2014.

[20] J. Scarlett, J. Evans, and S. Dey, "Compressed sensing with prior information: Information-theoretic limits and practical decoders," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 427–439, Jan. 2013.

[21] X. Wang and J. Liang, "Side information-aided compressed sensing reconstruction via approximate message passing," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Florence, Italy, May 2014.

[22] J. F. Mota, N. Deligiannis, A. Sankaranarayanan, V. Cevher, and M. R. Rodrigues, "Dynamic sparse state estimation using ℓ_1-ℓ_1 minimization: Adaptive-rate measurement bounds, algorithms and applications," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Brisbane, Australia, Apr. 2015.

[23] J. F. Mota, N. Deligiannis, A. C. Sankaranarayanan, V. Cevher, and M. R. Rodrigues, "Adaptive-rate sparse signal reconstruction with application in compressive background subtraction," *IEEE Trans. Signal Process.*, in press, 2016.

[24] G. Warnell, S. Bhattacharya, R. Chellappa, and T. Basar, "Adaptive-rate compressive sensing using side information," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3846–3857, 2015.

[25] E. Zimos, J. F. C. Mota, M. R. D. Rodrigues, and N. Deligiannis, "Bayesian compressed sensing with heterogeneous side information," in *Data Compression Conference*, Snowbird, Utah, USA, Apr. 2016.

[26] M. Taj and A. Cavallaro, "Distributed and decentralized multicamera tracking," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 46–58, 2011.

[27] A. Sankaranarayanan, A. Veeraraghavan, and R. Chellappa, "Object detection, tracking and recognition for multiple smart cameras," *Proc. of IEEE*, vol. 96, no. 10, pp. 1606–1624, 2008.

[28] A. Y. Yang, M. Gastpar, R. Bajcsy, and S. Sastry, "Distributed sensor perception via sparse representation," *Proc. of IEEE*, vol. 98, no. 6, pp. 1077–1088, 2010.

[29] H. V. Luong, J. Seiler, A. Kaup, and S. Forchhammer, "A reconstruction algorithm with multiple side information for distributed compression of sparse sources," in *Data Compression Conference*, Snowbird, Utah, USA, Apr. 2016.

[30] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of Convex Analysis*. Springer, 2004.

[31] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia object image library (coil-100)," Technical Report CUCS-006-96, Feb. 1996.

[32] D. G. Lowe, "Object recognition from local scale-invariant features," in *IEEE Int. Conf. on Computer Vision*, Kerkyra, Greece, Sep. 1999.

[33] D. Nistér and H. Stewénius, "Scalable recognition with a vocabulary tree," in *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, New York, USA, Jun. 2006.