Digital Signal Processing 133 (2023) 103833

Contents lists available at ScienceDirect

Digital Signal Processing

journal homepage: www.elsevier.com/locate/dsp

Analysis of general weights in weighted ℓ_{1-2} minimization through applications

K.Z. Najiya *, C.S. Sastry

Department of Mathematics, Indian Institute of Technology, Hyderabad, 502285, India

ARTICLE INFO

Article history:
Available online 23 November 2022

Keywords:
Compressed sensing
ℓ_{1-2} minimization
Weighted minimization
Reconstruction in MRI

ABSTRACT

The weighted ℓ_{1-2} minimization has recently attracted some attention due to its capability to deal with highly coherent matrices. Notwithstanding the availability of its stable recovery guarantees, there appear to be some issues not addressed in the literature, which are (i). convergence of the solver for the weighted ℓ_{1-2} minimization analytically, and (ii). detailed analysis of relevance of general weights to applications. While establishing the convergence of the solver of the weighted ℓ_{1-2} minimization, we demonstrate the significance of general weights, $w \in (0, 1)$, empirically through some applications, including the reconstruction of magnetic resonance images. In particular, we show that the general weights attain significance when we do not have fully accurate or fully corrupt information about the support of the signal to be reconstructed from its linear measurements. We conclude the work by discussing a numerical scheme that chooses the partial support and the weights iteratively.

1. Introduction

Compressed Sensing (CS) deals with recovering a sparse signal from a small set of its linear projections. In CS, one obtains the sparsest solution of an under-determined linear system via the ℓ_1 minimization. The applications of CS are far and wide in diverse fields [1][2][3][4]. The standard CS techniques, nevertheless, are non-adaptive in the sense that they do not take into account the properties of the signal to be reconstructed. It is possible to obtain a prior partial support estimate of the signal in many applied domains, [5][6] such as audio/video processing, dynamic Magnetic Resonance Imaging (MRI) etc. To accommodate such a prior support constraint into the reconstruction process, several researchers [5][6][7] have made novel contributions proposing the weighted ℓ_1 minimization problems and establishing the essential recovery guarantees. The results in these works have shown that the general weights can improve the reconstruction significantly when we have sufficiently accurate information about the support of the signal to be recovered.

In [8], Esser et al. have worked with the ℓ_{1-2} minimization in the context of finding structured sparse solutions to the non-negative least square problems. The basic idea behind this is that the level curves of the ℓ_{1-2} function given by $(\|.\|_1 - \|.\|_2)$ promote more sparsity than the standard ℓ_1 norm. P. Yin et al. [9] have established stable recovery results in terms of the RIP for the ℓ_{1-2} minimization problem and presented an iterative method for solving this based on the *Difference of Convex functions Algorithm* (DCA). They have also shown that, for highly coherent matrices, the ℓ_{1-2} minimization works better than the existing non-convex solvers available in the literature. Several authors have worked on certain variants of this problem [10][11], its numerical properties [12] and its theoretical recovery guarantees [13][14]. Recently, improved recovery guarantees for the ℓ_{1-2} minimization in terms of coherence [15], Restricted Isometry Property (RIP), [16] and Null Space Property (NSP) [17] have been established in the literature. These non-convex minimization techniques work well in sparse signal recovery, matrix completion, and reconstruction in MRI [12][18]. The recent sparse regularization techniques [19][20] using ℓ_{1-2} function and other non-convex sparse regularizers in deep neural networks have been shown to provide their compressed versions with comparable accuracy as their dense counterpart. The reason behind the popularity of ℓ_{1-2} minimization lies in its superior reconstruction quality in many applied domains even when the underlying matrix has coherence close to 1. In addition, the availability of efficient algorithms for its execution [9][10] has also contributed to its popularity.

The authors of [21][22] have generalized the ℓ_{1-2} minimization by considering general weights $w \in (0, 1)$ in ℓ_{1-2} minimization and proposed a solver along with the stable recovery conditions. Recently in [23], the authors have extended the notion of the weighted ℓ_{1-2} minimization to the block sparsity setup. To the

* Corresponding author.
E-mail addresses: najiyakz@gmail.com (K.Z. Najiya), csastry@math.iith.ac.in (C.S. Sastry).

https://doi.org/10.1016/j.dsp.2022.103833
1051-2004/

best of our knowledge, however, there have been no results that discuss (i) the convergence of the solver which is based on DCA and (ii) an empirical analysis that shows the significance of general weights in applications. Motivated by this, in this paper, we provide convergence guarantees of the DCA-based solver and demonstrate the role of general weights through some reconstruction problems. Consequently, this work complements the novel ideas presented in [21]. More specifically, we demonstrate the importance of general weights in ℓ_{1-2} minimization problem through extensive simulation work including the image reconstruction in MRI, and compare the results with those obtained through other state-of-the-art methods. Our empirical results show that the general weights in $(0, 1)$ can give smaller reconstruction errors than the standard weights 0 and 1, provided we have some partial information about the support of the sparse vector. Importantly, we observe that the role of the general weights becomes more significant for the coherent matrices in the sense that the reconstruction error falls significantly for a small deviation of weights from 0 or 1. To summarize, the novelty of the present work lies in establishing the relevance of the general weighted ℓ_{1-2} problem. The contributions of the present work may be summarized as follows:

- Demonstrating the role of general weights through applications
- Establishing the convergence of the associated solver
- Proposing an iterative method for choosing partial support and weights.

The paper is organized into several sections. In section 2, we discuss the relevant basics of CS and related results, and contribution of present work. In section 3, we discuss the convergence analysis of the solver and in section 4 we discuss simulation results providing a comparison with the other state-of-the-art methods and an application to reconstruction in MRI. We discuss a numerical scheme for iteratively determining the partial support and the weights in the ℓ_{1-2} minimization in section 5, followed by some concluding remarks in the last section.

2. Preliminaries and relevant works

In this section, we discuss the essential basics of CS, some relevant solvers, and a summary of present contributions.

2.1. Basics of compressed sensing

For a subset T of $[n] := \{1, 2, \cdots, n\}$, suppose $x_T \in \mathbb{R}^n$ denotes restriction of a vector x to the set T, which means $(x_T)_i = x_i$, for $i \in T$, and 0 else. The bold-symbol $\mathbf{0} \in \mathbb{R}^n$ denotes zero vector in $\mathbb{R}^n$. For $x \in \mathbb{R}^n$, support of x (denoted by $supp(x)$) is the set of all indices corresponding to the nonzero components of x and $\|x\|_0 = |supp(x)|$. A vector x is called k-sparse if $\|x\|_0 \leqslant k$. The mutual coherence [24] of a matrix $A \in \mathbb{R}^{m\times n}$ is defined as

$$\mu(A) = \max_{i\neq j, 1\leqslant i,j\leqslant n} \frac{|\langle a_i, a_j\rangle|}{\|a_i\|_2\|a_j\|_2},$$

where a_i denotes the i^{th} column of A. A matrix A is said to have the Restricted Isometry Property (RIP) of order k [24], if there exists a constant $\delta \in (0, 1)$ such that for any k-sparse vector x, we have

$$(1-\delta)\|x\|_2^2 \leqslant \|Ax\|_2^2 \leqslant (1+\delta)\|x\|_2^2. \tag{1}$$

The least such δ (denoted by δ_k) satisfying (1) is called the Restricted Isometry Constant (RIC) of order k.

In Compressed Sensing (CS), one considers the following ℓ_1 minimization problem [24] for finding a solution of $b = Ax + \eta$ with $\|\eta\|_2 \leqslant \epsilon$ that possesses a very few nonzero components

$$(\mathcal{P}_1): \quad \min_{x\in\mathbb{R}^n} \|x\|_1 \quad \textit{subject to} \quad \|Ax - b\|_2 \leqslant \epsilon, \tag{2}$$

where A is an under-determined matrix. The recovery guarantees [24] in compressed sensing via the standard ℓ_1 norm state that a higher (or smaller) value for the coherence translates to more (or less) number of measurements for faithful reconstruction. In applications such as computed tomography [25], poor coherence of discrete Radon matrix implies a large number of X-ray projections, which is to be avoided. The ℓ_1 minimization problem in (2) is not signal adaptive as it does not include any information about the underlying signal to be recovered.

2.2. Weighted ℓ_1-norm based CS

The weighted CS techniques are signal adaptive as they incorporate prior partial support information of the signal into the recovery process. The weighted ℓ_1 minimization problem $(\ell_{1,w})$ [5] is defined as

$$(\mathcal{P}_{1,w}): \quad \min_{x\in\mathbb{R}^n} \|x\|_{1,w} \quad \textit{subject to} \quad \|Ax - b\|_2 \leqslant \epsilon, \tag{3}$$

where $\|x\|_{1,w} = \|wx_T + x_{T^c}\|_1$ with $w \in [0, 1]$, and T is the prior partial support information. Here wx_T denotes the point-wise multiplication of the scalar w and the vector x_T, and $x_{T^c} \in \mathbb{R}^n$ is x restricted to T^c in line with the definition of x_T.

In the weighted ℓ_1 minimization problem given in (3), the entries from the partial support of the signal are less penalized than the rest so that the recovered signal is more likely to contain the entries from the partial support set. Such partial support information of many signals is available in signal processing [5][6]. The authors of [6][7][26] have dealt with the sparse recovery of signals using the weighted ℓ_1 minimization. The authors of [5] have proposed a stable recovery bound for the $(\mathcal{P}_{1,w})$ problem. It is worth mentioning here that the re-weighted ℓ_1 minimization methods [27] are different from the weighted ℓ_1 minimization methods as the former ones deal with approximating non-convex sparsity promoting functions such as ℓ_p, for $p \in [0, 1)$, using weighted ℓ_1 functions through the iteratively chosen weights.

2.3. ℓ_{1-2} based CS

By observing that the level curves of the ℓ_{1-2} function via $(\|.\|_1 - \|.\|_2)$ promote more sparsity than the standard ℓ_1 norm, Esser et al. [8] have worked with the following ℓ_{1-2} minimization:

$$(\mathcal{P}_{1-2,1}): \quad \min_{x\in\mathbb{R}^n} \left[\|x\|_1 - \|x\|_2\right] \quad \textit{subject to} \quad \|Ax - b\|_2 \leqslant \epsilon, \tag{4}$$

whose unconstrained version is defined as

$$\min_{x\in\mathbb{R}^n} \left[\frac{1}{2}\|Ax - b\|_2^2 + \lambda(\|x\|_1 - \|x\|_2)\right]. \tag{5}$$

The stable recovery guarantees of the afore-stated problem have been established in [9]. In [14], the authors have provided the coherence-based stable recovery bounds for the unconstrained ℓ_{1-2} minimization problem (5). In [21], for $w \in [0, 1]$, the authors have considered the following weighted ℓ_{1-2} $(\ell_{1-2,w})$ minimization problem

$$(\mathcal{P}_{1-2,w}): \quad \min_{x\in\mathbb{R}^n} \left[\|x\|_{1,w} - \|x\|_{2,w}\right] \quad \textit{subject to} \quad \|Ax - b\|_2 \leqslant \epsilon, \tag{6}$$

Algorithm 1: Weighted DCA (WDCA) for solving (8).

Define $\epsilon_0 > 0$. Initialize $x^0 = 0$.
for $k = 0, 1, \cdots, M_1$ **do**
$x^{k+1} = \arg\min_x \left[\frac{1}{2}\|Ax - b\|_2^2 + \frac{c}{2}\|x\|_2^2 + \lambda\|x\|_{1,w} - \langle x, \lambda e_w^k \rangle \right]$;
end
M_1: Maximum number of iterations for DCA.

and established its stable recovery conditions. In the above problem,

$$\|x\|_{1,w} = \|wx_T + x_{T^c}\|_1 \text{ and } \|x\|_{2,w} = \|wx_T + x_{T^c}\|_2. \tag{7}$$

A DCA-based solver for the weighted ℓ_{1-2} minimization has also been proposed in [21].

2.4. Solver for the weighted ℓ_{1-2} minimization

In this subsection, we discuss a solver for the following unconstrained $(\ell_{1-2,w})$ problem [21]

$$\min_{x\in\mathbb{R}^n} \left[f(x) := \frac{1}{2}\|Ax - b\|_2^2 + \lambda(\|x\|_{1,w} - \|x\|_{2,w}) \right]. \tag{8}$$

For the sake of convergence analysis and for fixing the notation, we present herein the ideas of the adapted algorithm. We take $f(x) = g(x) - h(x)$, with $g(x) = \frac{1}{2}\|Ax - b\|_2^2 + \lambda\|x\|_{1,w} + \frac{c}{2}\|x\|_2^2$ and $h(x) = \lambda\|x\|_{2,w} + \frac{c}{2}\|x\|_2^2$, where $c > 0$ is considered to ensure the strong convexity of functions g and h, which is needed for the convergence analysis. The DCA method iteratively computes x^k and y^k which approximate the optimal solution for the primal and dual problems at each iteration. One considers

$$\begin{aligned} &y^k \in \partial h(x^k) \\ &x^{k+1} = arg\min_{x\in\mathbb{R}^n} \left[g(x) - \big(h(x^k) + \langle y^k, x - x^k\rangle\big) \right], \end{aligned} \tag{9}$$

where $\partial h(x^k)$ is the subgradient of $h(x)$ at x^k. It may be noted that

$$\begin{aligned} f(x^{k+1}) &\leqslant g(x^{k+1}) - \big(h(x^k) + \langle y^k, x^{k+1} - x^k\rangle\big) \\ &\leqslant g(x^k) - \big(h(x^k) + \langle y^k, x^k - x^k\rangle\big) \\ &= f(x^k), \end{aligned} \tag{10}$$

which follows from the definition of the subgradient and (9). Since $f(x) \geqslant 0$, $\forall\, x \in \mathbb{R}^n$, the monotonically decreasing sequence $\{f(x^k)\}$ is convergent.

At iteration k, it follows that

$$x^{k+1} = \arg\min_x \left[\frac{1}{2}\|Ax - b\|_2^2 + \frac{c}{2}\|x\|_2^2 + \lambda\|x\|_{1,w} - \langle x, \lambda e_w^k \rangle \right], \tag{11}$$

where

$$e_w^k = \begin{cases} \frac{x_{w^2}^k}{\|x_w^k\|_2} + cx^k & \text{if } x^k \neq 0 \\ 0 & \text{else,} \end{cases}$$

with $x_{w^2} = w^2 x_T + x_{T^c}$. It is proved in section 3 that x^k converges to a stationary point. As a result, the stopping criterion for this algorithm may be considered as $\frac{\|x^{k+1} - x^k\|_2}{max\{\|x^k\|_2, 1\}} < \epsilon_0$, for some given parameter $\epsilon_0 > 0$. We summarize the weighted DC algorithm (WDCA) in Algorithm 1. The *Alternating Direction Method of Multipliers* (ADMM) algorithm [28] for solving the sub-problems mentioned in (11) is given in Algorithm 2, wherein $\lambda_w \in \mathbb{R}^n$ is defined as

Algorithm 2: ADMM for (11).

Initialize x^0, z^0 and y^0.
for $l = 0, 1, \cdots, M_2$ **do**
$x^{l+1} = (A^TA + (c+\delta)I)^{-1}(A^Tb - v + \delta z^l - y^l)$
$z^{l+1} = \mathcal{S}(x^{l+1} + y^l/\delta, \lambda_w/\delta))$
$y^{l+1} = y^l + \delta(x^{l+1} - z^{l+1})$
end
M_2: Maximum number of iterations for ADMM.

$$(\lambda_w)_i = \begin{cases} \lambda w, & i \in T \\ \lambda, & \text{else,} \end{cases} \tag{12}$$

and $\mathcal{S}(x, r)$ is the soft thresholding operator which is defined as $(\mathcal{S}(x, r))_i = sgn(x_i)max(|x_i| - r_i, 0)$. The stopping criterion [29] for Algorithm 2 is given by

$$\begin{aligned} &\|x^l - z^l\|_2 \leqslant \sqrt{n}\epsilon^{abs} + \epsilon^{rel} max\{\|x^l\|_2, \|z^l\|_2\}, \\ &\|\delta(z^l - z^{l-1})\|_2 \leqslant \sqrt{n}\epsilon^{abs} + \epsilon^{rel}\|y^l\|_2. \end{aligned}$$

2.5. Present contribution

The authors of [21] have provided the recovery guarantees of the problem in (6). But, to our knowledge, there is no convergence analysis established for the solver of (6). In addition, there appears to be no study available for determining the weights other than the trial and error method. Motivated by these points, we demonstrate the nontrivial role played by the general weights both in the cases of incoherent and highly coherent matrices. Besides establishing the convergence of the solver, we discuss a possible procedure for determining the weights iteratively. Furthermore, through reconstruction in MRI, we show that the general weights assume significance when we do not have fully accurate or fully corrupt information about the support of the underlying signal.

3. Convergence guarantees of the weighted DCA solver

We now prove that the iterates $\{x^k\}$ in Algorithm 1 converge to a stationary point, which obeys the first order optimality condition.

Theorem 3.1. *Suppose the sequence of iterates $\{x^k\}$ is defined as in (11). Then*

1. *If $rank(A) > |T|$, then the sequence of iterates $\{x^k\}$ is bounded.*
2. *$\|x^{k+1} - x^k\|_2 \to 0$ as $k \to \infty$.*
3. *Any non-zero limit point x^* of $\{x^k\}$ obeys the first order optimality condition $\mathbf{0} \in \big[A^T(Ax^* - b) + \lambda\big(\partial(\|x^*\|_{1,w}) - \nabla(\|x^*\|_{2,w})\big)\big]$, implying that x^* is a stationary point of f.*

Proof: We prove the afore-stated points in the same order.

1. We first prove that $f(\nu x) \to \infty$ as $\nu \to \infty$ for any $x \in \mathbb{R}^n/\{0\}$ which implies that $f(x) \to \infty$ as $\|x\|_2 \to \infty$. Since $\{x^k\} \subseteq \{x : f(x) \leqslant f(x^0)\}$, the above result implies that $\{x^k\}$ is bounded. We have $f(\nu x) = \frac{1}{2}\|\nu Ax - b\|_2^2 + \nu\lambda(\|x\|_{1,w} - \|x\|_{2,w})$ for $\nu \geqslant 0$. If $x \in Ker(A)$, then $\|x\|_0 \geqslant rank(A) + 1 > |T| + 1$ which implies that $\|x_{T^c}\|_0 > 1$. Thus $\|x\|_{1,w} - \|x\|_{2,w} > 0$ for all w and $f(\nu x) \to \infty$ as $\nu \to \infty$ as $\nu\lambda(\|x\|_{1,w} - \|x\|_{2,w}) \to \infty$ as $\nu \to \infty$. It may be noted that $f(\nu x) \geqslant \frac{1}{2}(\nu\|Ax\|_2 - \|b\|_2)^2 + \nu\lambda(\|x\|_{1,w} - \|x\|_{2,w})$. If $x \notin Ker(A)$, then $f(\nu x) \to \infty$ as $\nu \to \infty$, since $\nu\|Ax\|_2 \to \infty$ as $\nu \to \infty$.
2. Since both $g - \frac{c}{2}\|.\|_2^2$ and $h - \frac{c}{2}\|.\|_2^2$ are convex, by the Proposition A.1 in [30], we have $f(x^k) - f(x^{k+1}) \geqslant c\|x^{k+1} - x^k\|_2$. From the convergence of $\{f(x^k)\}$, it follows that $\|x^{k+1} - x^k\|_2 \to 0$ as $k \to 0$.

Table 1
Noiseless case: Reconstruction errors (denoted by $Err(w)$) at 0 and optimal w (denoted by w_{opt}), provided by $\ell_{1-2,w}$ solver for Gaussian (incoherent) matrix. For each ρ, the shaded cell gives error at $w = 1$ which is same for all values of α. Here ρ and α satisfy $\alpha\rho \leqslant 1$. The errors being small for some $w \in (0, 1)$, when $\alpha \neq 0, 1$, signify the importance of $\ell_{1-2,w}$ problem.

α	$\rho = 0.75$			$\rho = 1$			$\rho = 1.25$		
	w_{opt}	$Err(w_{opt})$	$Err(0)$	w_{opt}	$Err(w_{opt})$	$Err(0)$	w_{opt}	$Err(w_{opt})$	$Err(0)$
0	1	0.5830	1.5071	1	0.5865	2.3182	1	0.5866	4.6116
0.25	0.8	0.5731	0.9389	0.8	0.5786	1.2005	0.8	0.5716	2.4294
0.5	0.4	0.4517	0.5649	0.4	0.4226	0.6197	0.3	0.3716	1.0392
0.75	0.2	0.2792	0.2970	0.1	0.1770	0.2000	0.1	0.0529	0.0910
1	0	0.0917	0.0917	0.1	5.1158×10^{-05}	5.7706×10^{-05}	×	×	×

3. Let $\{x^{n_k}\}$ be the sub-sequence of $\{x^k\}$ that converges to x^*. Then the optimality condition at n_k^{th} iteration implies the following:

$$\mathbf{0} \in \left[A^T(Ax^{n_k} - b) + \lambda\left(\partial(\|x^{n_k}\|_{1,w}) - \nabla(\|x^{n_k-1}\|_{2,w})\right)\right]. \quad (13)$$

By the Proposition 3.1(c) of [9], we have $\partial(\|x^{n_k}\|_1) \subseteq \partial(\|x^*\|_1)$, for large n_k. This ensures that $\partial(\|x^{n_k}\|_{1,w}) \subseteq \partial(\|x^*\|_{1,w})$, for large n_k. This gives

$$-\left[A^T(Ax^{n_k} - b) - \lambda\nabla(\|x^{n_k-1}\|_{2,w})\right] \in \lambda\partial(\|x^*\|_{1,w}). \quad (14)$$

Further,

$$\begin{aligned}
&\lim_{k\to\infty}\left[A^T(Ax^{n_k} - b) - \lambda\nabla(\|x^{n_k-1}\|_{2,w})\right] \\
&= \lim_{k\to\infty}\left[A^T(Ax^{n_k} - b) - \lambda\nabla(\|x^{n_k}\|_{2,w}) + \lambda\nabla(\|x^{n_k}\|_{2,w})\right. \\
&\quad \left. - \nabla(\|x^{n_k-1}\|_{2,w})\right] \\
&= \lim_{k\to\infty}\left[A^T(Ax^{n_k} - b) - \lambda\nabla(\|x^{n_k}\|_{2,w})\right. \\
&\quad \left. + \frac{x_{w^2}^{n_k}}{\|x_w^{n_k}\|_2} - \frac{x_{w^2}^{n_k-1}}{\|x_w^{n_k-1}\|_2}\right] \\
&= A^T(Ax^* - b) - \lambda\nabla(\|x^*\|_{2,w})
\end{aligned} \quad (15)$$

as $\|x^{k+1} - x^k\|_2 \to 0$ as $k \to 0$. Thus (14) and (15) imply that

$$\mathbf{0} \in \left[A^T(Ax^* - b) + \lambda\left(\partial(\|x^*\|_{1,w}) - \nabla(\|x^*\|_{2,w})\right)\right].$$

Remark 3.1. It is worth mentioning here that the crucial part of this proof is the condition $rank(A) > |T|$, which ensures that $\|x_{T^c}\|_0 > 1$ for all $x \in Ker(A)$. It may be noted that this condition is not restrictive in the sense that, in general, the cardinality of the support estimate is taken less than the rank of the measurement matrix. This ensures that the iterates are bounded, which helps in proving the convergence of the iterates to a stationary point.

Remark 3.2. The novelty in the above proof occurs due to the presence of general weights w for the entries in the partial support T. The sufficient condition that ensures the boundedness of the iterates is different from that of the DCA solver for the standard ℓ_{1-2} minimization. It may be observed that, for proving the convergence of the iterates in the ℓ_2 norm sense, we incorporate the ideas from [9]. While for showing that the stationary point satisfies the first-order optimality condition, we extend the proof techniques from [30].

Remark 3.3. The weights introduced into the optimization problem result in $M_1M_2|T|$ number of additional multiplications (which may be ascertained from Algorithms 1 and 2), where $|T|$, M_1 and M_2 respectively stand for the cardinality of the partial support set T, number of DCA iterations and number of ADMM iterations. Nevertheless, the order of the complexity remains same as the matrix inversion step in Algorithm 2 is the dominating part in calculating the complexity of the algorithm.

4. Empirical analysis of weights

In this section, we demonstrate the significance of general weights in the following lines: (i). Recovery of the sparse signals from their randomly sampled linear measurements, (ii). Comparison of results obtained via (6) with those obtained through other methods and (iii). Reconstruction in MRI.

4.1. On the recovery of sparse signals

We test our weighted DCA algorithm on two sets of matrices, viz, Gaussian matrices and randomly over-sampled partial discrete cosine transform (DCT) matrices both with and without noise. Though the Gaussian matrices are incoherent possessing small mutual coherence, the over-sampled DCT matrices are not (due to their high coherence values). Consequently, in a way, these two represent complementary cases. Throughout the simulation work, unless stated, we report an average behaviour over 100 realizations when dealing with random matrices, and the signal x is drawn from the normal distribution $\mathcal{N}(0, 1)$. We report simulations for various values of ρ and α, where ρ represents the relative size of the partial support compared to the size of the actual support of the signal. Here α represents the proportion of the partial support that intersects with the actual support of the signal, which is a measure of the accuracy of the partial support information so that $\alpha\rho \leqslant 1$. Since the partial support information so obtained need not be fully accurate, analysis of reconstruction error with respect to w for different values of α attains significance. The following simulations show that the performance of the ℓ_{1-2} minimization improves when we introduce general weights. For the WDCA solver, we have fixed $\lambda = 10^{-06}$, $c = 10^{-09}$. The remaining parameters have been fixed as in [9]. All the simulations have been performed in Matlab R2021a on a laptop with 8 GB RAM and 2.50 GHz Intel(R) Core(TM) i5-7200U CPU under Windows 10 operating system.

The simulation results for the incoherent case (that is, with Gaussian matrices) are shown in Table 1. Here the matrix A has been taken as Gaussian possessing mean 0 and standard deviation $\frac{1}{\sqrt{m}}$ with $m = 128, n = 256$. The average coherence of this matrix is 0.36. In our simulation work, as an example, the sparsity k of the vector has been set to 100. The size of partial support has been taken as ρk for 3 different values of ρ = 0.75, 1 and 1.25. For each size of partial support, we have taken 0, 0.25, 0.5, 0.75 and 1 as possible values of α. We have performed $\ell_{1-2,w}$ minimization for each w in the set $\{0, 0.1, 0.2, ..., 1\}$. For different choices of (α, ρ), Tables 1 and 2 provide average relative ℓ_2 norm reconstruction errors and optimal values of $w \in [0, 1]$ that result in least

Table 2
Noise-case: Reconstruction errors with respect to the weighted DCA solver for Gaussian matrices.

α	$\rho = 0.75$			$\rho = 1$			$\rho = 1.25$		
	w_{opt}	Err(w_{opt})	Err(0)	w_{opt}	Err(w_{opt})	Err(0)	w_{opt}	Err(w_{opt})	Err(0)
0	1	0.5989	1.4891	1	0.5958	2.2408	1	0.5977	4.4653
0.25	0.8	0.5890	0.9462	0.8	0.5873	1.2143	0.8	0.5787	2.3521
0.5	0.4	0.4675	0.5872	0.3	0.4142	0.6247	0.3	0.3642	1.1319
0.75	0.2	0.2740	0.2911	0.1	0.1741	0.1967	0.1	0.0467	0.0819
1	0	0.0768	0.0768	0.1	6.4081×10^{-04}	6.5774×10^{-04}	×	×	×

Table 3
Noiseless case: Reconstruction errors given by the WDCA solver for randomly over-sampled partial DCT (coherent) matrix. The errors being significantly small even when w_{opt} remains close to 0 with $\alpha \neq 0, 1$ signify the importance of general weighted $\ell_{1-2,w}$ problem.

α	$\rho = 0.75$			$\rho = 1$			$\rho = 1.25$		
	w_{opt}	Err(w_{opt})	Err(0)	w_{opt}	Err(w_{opt})	Err(0)	w_{opt}	Err(w_{opt})	Err(0)
0	1	0.9356	1.6723	1	0.9420	1.8422	1	0.9464	1.9393
0.25	0.2	0.5835	1.1850	0.2	0.5065	0.9379	0.2	0.3659	0.7890
0.5	0.2	0.2526	0.7354	0.2	0.1128	0.3700	0.1	0.0430	0.1463
0.75	0.2	0.0401	0.3131	0.3	0.0097	0.0920	0.1	0.0214	0.0772
1	0	1.6058×10^{-04}	1.6058×10^{-04}	0.1	5.7538×10^{-05}	6.7603×10^{-05}	×	×	×

Table 4
Noise-case: Reconstruction errors for randomly over-sampled partial DCT (coherent) matrix. When $\alpha = 0$, reconstruction error at $w = 1$ is 0.9699.

α	$\rho = 0.75$			$\rho = 1$			$\rho = 1.25$		
	w_{opt}	Err(w_{opt})	Err(0)	w_{opt}	Err(w_{opt})	Err(0)	w_{opt}	Err(w_{opt})	Err(0)
0	0.7	0.9622	1.7641	1	0.9350	1.8232	1	0.9367	2.0777
0.25	0.2	0.5586	1.2856	0.2	0.4748	0.9977	0.1	0.3767	0.7519
0.5	0.1	0.2488	0.7016	0.1	0.1265	0.3569	0.1	0.0379	0.1893
0.75	0.2	0.0616	0.3759	0.1	0.0248	0.1324	0.2	0.0078	0.1148
1	0.1	0.0047	0.0052	0.1	0.0022	0.0023	×	×	×

error both in noise as well as noiseless cases. The relative ℓ_2 norm reconstruction error has been calculated via $\frac{\|x_r - x_0\|_2}{\|x_0\|_2}$, where x_0 and x_r are the original and reconstructed signals respectively. It may be observed in the tables that the optimal w takes 1 when $\alpha = 0$, implying thereby that the performance of the weighted ℓ_{1-2} problem reduces to that of standard ℓ_{1-2} problem. Further, it may also be noted that the reconstruction error at $w = 1$ (shown in the shaded cells in Tables 1 to 4) is independent of α for each ρ. Driven by this, we have avoided reporting these errors in a separate column.

It is clear from the Tables 1 and 2 that for the incoherent Gaussian matrix case, when $\alpha \neq 0, 1$, $w \in (0, 1)$ provides least reconstruction errors than the cases where $w = 0,\ 1$. The average behaviour is that, for $\alpha \geqslant 0.5$, optimal w is smaller than 0.5 and vice versa.

The average relative ℓ_2 norm reconstruction errors for the coherent matrix are shown in Tables 3 and 4 for the noiseless and noise cases respectively. Here the matrix A has been taken as the randomly over-sampled partial DCT matrix whose columns satisfy

$$a_i = \frac{1}{\sqrt{m}} cos\left(\frac{2\pi i v}{F}\right), \quad i = 1, \cdots, n, \tag{16}$$

where $v \sim \mathcal{U}([0, 1]^m)$ and $F \in \mathbb{N}$ is a refinement factor. In our simulations, we have considered $m = 100$, $n = 2000$ and $F = 20$. The average mutual coherence of this matrix is 0.9999. Despite this, a signal x can be recovered from its linear projections that are generated as Ax, if it has sufficient minimum separation [31]. A signal x has minimum separation L if $\min_{j,k \in supp(x)} |j - k| \geqslant L$. In our simulations, we have considered the signals of sparsity 36 with $L = 2F$ and $k = 36$. It is clear from Tables 3 and 4 that for the case of coherent DCT matrix, where $\alpha \neq 0, 1$, the weights in $(0, 1)$ give least reconstruction errors compared to the cases where $w = 0,\ 1$. Importantly, in this case, despite optimal w being close to 0, the corresponding error is smaller than that provided by the $w = 0$ case. Also here the average behaviour is that, for $\alpha = 0$, the optimal w is 1 and, for other values of α, the optimal w is at most 0.5.

To summarize, the stated simulations on the recovery of sparse signals in the cases of coherent and incoherent matrices provide the least reconstruction errors with weights in $(0, 1)$, when we do not have fully accurate or fully corrupt information about support of the signal. In some of the cases, even when the optimal values of w are close to 0 or 1, there is a significant difference between the corresponding reconstruction errors. Therefore the efficiency of $(\mathcal{P}_{1-2,1})$ problem can be improved by considering general weights in $\ell_{1-2,w}$ problem. The simulation results also show that bigger ρ, corresponding to a bigger support estimate, gives finer reconstruction and that the value of α has a better bearing on reconstruction than ρ. To conclude, $\ell_{1-2,w}$ gives the better reconstruction results for $w \leqslant 0.5$ when $\alpha > 0.5$. It is worth mentioning here that the finer range of values of w depends on many factors including the properties of the matrix and the signal to be recovered.

The overall observations of the behaviour of the optimal reconstruction error via weighted ℓ_{1-2} minimization problem for the incoherent and coherent sensing matrices are as follows:

- For each fixed value of ρ, the reconstruction error decreases with increasing α.
- For each fixed value of α, the reconstruction error decreases with increasing ρ.
- For a fixed ρ, when $\alpha = 0$ (that is, when the partial support is completely inaccurate), better reconstruction is obtained around the standard weight $w = 1$.

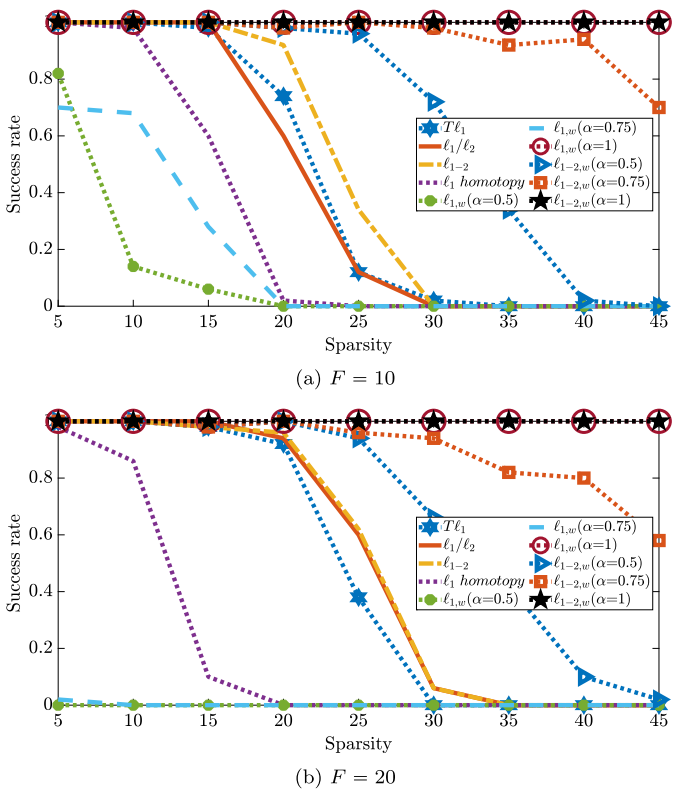
- For the values of w lying closer to 0, better reconstructions are obtained when $\alpha\rho$ is sufficiently big.
- For fixed ρ, when $\alpha = 1$ (that is, when the partial support is fully accurate) better reconstruction is obtained around $w = 0$.
- For a fixed ρ, when $\alpha \neq 0$ or 1 (that is, when the partial support is neither fully inaccurate nor accurate, better reconstruction is obtained for the general values of the weights in $w \in (0, 1)$.
- For the incoherent case, the behaviour is that when $\alpha < 0.5$, the optimal weights are bigger than 0.5, and when $\alpha > 0.5$, the optimal weights are less than 0.5.
- For the coherent case, the behaviour is that when $\alpha \neq 0$ and 1, the optimal weights take values which are at most 0.5.
- In some of the cases, even when the optimal values of w are close to 0 or 1, there is a significant difference between the corresponding reconstruction errors.

4.2. Comparison with other solvers

In this subsection we compare the WDCA solver for the unconstrained ($\ell_{1-2,w}$) problem in (8) with other solvers. It has already been demonstrated in [9][12] that for highly coherent matrices ℓ_{1-2} minimization works better than the existing solvers like ADMM-Lasso [32], IRLS-ℓ_p [33], re-weighted ℓ_1 [27], CoSAMP [34] and half thresholding [35]. We consider here a comparison with other methods such as ℓ_1 homotopy, [36], $T\ell_1$ [37] and ℓ_1/ℓ_2 [38], $(\ell_1 - \nu\ell_2)$ [10], weighted ℓ_1 ($\ell_{1,w}$) [5] along with ℓ_{1-2} [9] over the highly coherent partial DCT matrices. It is to be mentioned here that the weighted ℓ_p [39] minimization is shown to give inferior performance than the weighted ℓ_{1-2} minimization for the highly coherent matrices, and hence it is not considered here. We have considered matrices of size 100×2000 for $F = 10$ and 20 cases whose average mutual coherences are 0.9982 and 0.9999 respectively. Further, we have generated x randomly as a k sparse vector with minimum separation $2F$ where k takes the values 5 to 45 and the size of the partial support T is taken as k itself. For $\ell_{1,w}$ and $\ell_{1-2,w}$ problems, we have considered three different cases where α takes 0.5, 0.75 and 1. We have taken 50 trials for each setup, where a trial has been considered successful if the relative ℓ_2 norm reconstruction error is less than 10^{-03}. For the weighted DCA solver we have fixed $\lambda = 10^{-6}$, $\delta = 10\lambda$, $M_1 = 10$, $M_2 = 5000$, $\epsilon_0 = 10^{-2}$, $\epsilon^{abs} = 10^{-7}$ and $\epsilon^{rel} = 10^{-5}$. For $T\ell_1$ problem, we have fixed $\gamma = 10^{-06}$, $c = 0$, $a = 1$, $maxoit = 10$, $maxit = 5000$ and $tol = 10^{-08}$. For ℓ_1/ℓ_2 problem, we have fixed $restol = 10^{-05}$. For the ℓ_1-homotopy we have set delx-mode to 'qr'. For ℓ_{1-2}, $(\ell_1 - \nu\ell_2)$ (with $\nu = 0.5$) and WDCA solvers we have fixed $\lambda = 10^{-06}$. The rest of the parameters in all these solvers have been set to their default values. Since performance of $(\ell_1 - \nu\ell_2)$ is almost the same as ℓ_{1-2}, we have excluded the $(\ell_1 - \nu\ell_2)$ case from the plots. For $\ell_{1,w}$ and $\ell_{1-2,w}$ solvers, success rate corresponding to the optimal w has been considered at each trial. From Fig. 1, it can be seen that the WDCA solver has better success rates than all the other solvers. For $\alpha = 1$, $\ell_{1,w}$ solver has equal performance as $\ell_{1-2,w}$. Nevertheless, when $\alpha \neq 1$, success rates of $\ell_{1,w}$ are too low especially for the $F = 20$ case. It can be concluded that the WDCA solver performs well even over less-sparse vectors.

4.3. MRI reconstruction using prior support estimate

We now turn to the application of MRI, which involves reconstructing an image from its projections. The goal of MRI is to recover images from less number of projections. The authors of [9][18] have demonstrated the applicability of the ℓ_{1-2} minimization to the reconstruction in MRI in a very novel way, demonstrating that a very highly accurate reconstruction of Shepp-Logan



(a) $F = 10$

(b) $F = 20$

Fig. 1. Success rates with the coherent matrices as discussed in subsection 4.2. This plot demonstrates the superior performance shown by the weighted ℓ_{1-2} solver.

phantom image of size 256×256 is possible from merely 7 projections through ℓ_{1-2} minimization problem. In order to further reduce the number of projections without distorting the quality of MRI reconstruction, one can incorporate the idea of the weighted ℓ_{1-2} minimization using a prior support estimate. For instance, when patients need to take routine MRI scans, the image from the previous scan can be used to find a support estimate, which helps us to reduce the number of projection samples needed for the current scan. In this subsection, we use this idea for the MRI reconstruction. In particular, we show that 3 samples are enough for faithful reconstruction of the Shepp-Logan phantom image of size 256×256 with a proper choice of weights.

Since images are sparse in the gradient domain in general, we take minimization in the gradient domain. We pose the weighted minimization problem for the MRI reconstruction as

$$\min_u \left[\|\partial_x(u)\|_{1,w} + \|\partial_x(u)\|_{1,w} - \sqrt{\|\partial_x(u)\|_{2,w}^2 + \|\partial_x(u)\|_{2,w}^2} \right] \quad s.t.\ \mathcal{R}\mathcal{F}u = f,$$

where u is a 2-dimensional image, $\partial_x(u)$, $\partial_y(u)$ are the partial derivatives of u in the x and y directions respectively, $\mathcal{R}$ denotes the sampling matrix in the frequency domain, $\mathcal{F}$ denotes 2D Fourier transform and f is the data. Here $\|\partial_x(u)\|_{1,w}$, $\|\partial_x(u)\|_{2,w}$ are the ℓ_1 and ℓ_2 norms of $(\partial_x(u))_w$ where $(\partial_x(u))_w = w(\partial_x(u))_T + (\partial_x(u))_{T^c}$ and T is the partial support information known a priori. Now we solve the following unconstrained problem:

$$\min_u \left[\frac{\mu}{2}\|\mathcal{R}\mathcal{F}u - f\|_2^2 + \|\partial_x(u)\|_{1,w} + \|\partial_x(u)\|_{1,w} - \sqrt{\|\partial_x(u)\|_{2,w}^2 + \|\partial_y(u)\|_{2,w}^2} \right]. \tag{17}$$

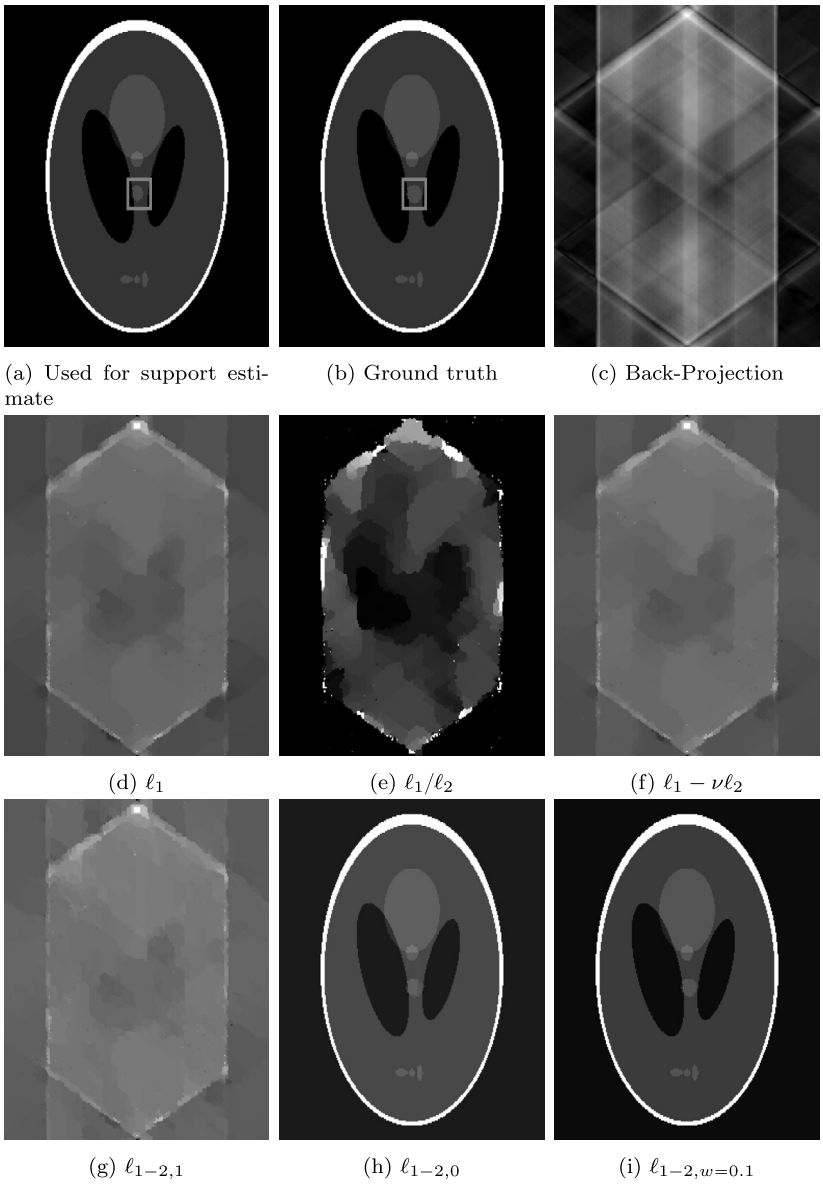


Fig. 2. Reconstruction of phantom image from 3 projections with different methods. We use the image (a) to obtain the prior support estimate to recover the new image (b). It is clear that, among the methods used, $\ell_{1-2,w}$ gives accurate reconstruction results at $w = 0.1$. In the result obtained through the $\ell_{1-2,0}$ solver, it may be noted that the marked portion is blurred at the corners. This figure shows the importance of general weights.

In order to solve the afore-stated problem, we combine the DCA techniques (discussed in Section 2.4) along with the split Bregman method in [40]. Then the iterative steps become

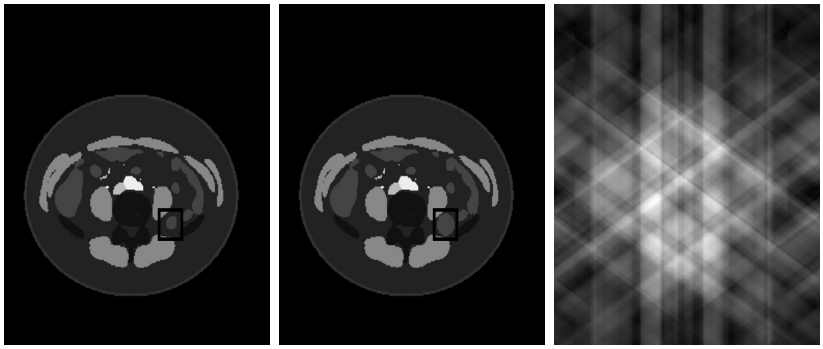
$$u^{k+1} = \min_{u,d_x.d_y} \Bigg[\frac{\mu}{2}\|\mathcal{R}\mathcal{F}u - f\|_2^2 + \|d_x\|_{1,w} + \|d_y\|_{1,w} - \frac{(d_x, d_y)^T)(\partial_x(u^k))_{w^2}, (\partial_y(u^k))_{w^2})}{\sqrt{\|\partial_x(u^k)\|_{2,w}^2 + \|\partial_y(u^k)\|_{2,w}^2}} + \frac{\lambda}{2}\|d_x - \partial_x(u) - b_x\|_2^2 + \frac{\lambda}{2}\|d_y - \partial_y(u) - b_y\|_2^2 \Bigg]. \tag{18}$$

Here $(\partial_x(u))_{w^2} = w^2(\partial_x(u))_T + (\partial_x(u))_{T^c}$. The values of the variables b_x and b_y are chosen through the Bregman iteration. In the methodology summarized in Algorithm 3, Δ is the Laplacian operator, $\mathcal{F}^T$ is the inverse Fourier transform, $\mathcal{S}$ is the soft thresholding operator, D and D^T are the forward and backward difference operators respectively. Here M_1 and M_2 are the number of outer (DCA) and inner (split Bregman) iterations respectively. Further, in Algorithm 3, $(t_x^k, t_y^k) = \frac{((\partial_x(u^k))_{w^2}, (\partial y(u^k))_{w^2})}{\sqrt{\|\partial_x(u^k)\|_{2,w}^2 + \|\partial_y(u^k)\|_{2,w}^2}}$ and the vector λ_w is defined as

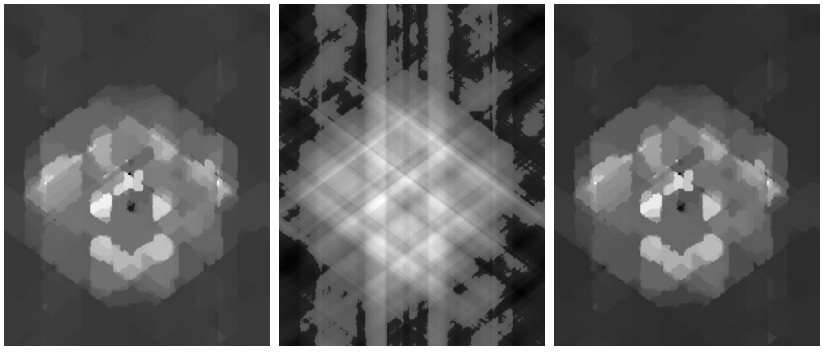
$$(\lambda_w)(i,j) = \begin{cases} \frac{w}{\lambda}, & (i,j) \in T. \\ \frac{1}{\lambda}, & else. \end{cases} \tag{19}$$

For the reported results, we have considered $M_1 = 50$, $M_2 = 50$, $\mu = 10^6$ and $\lambda = 10$.

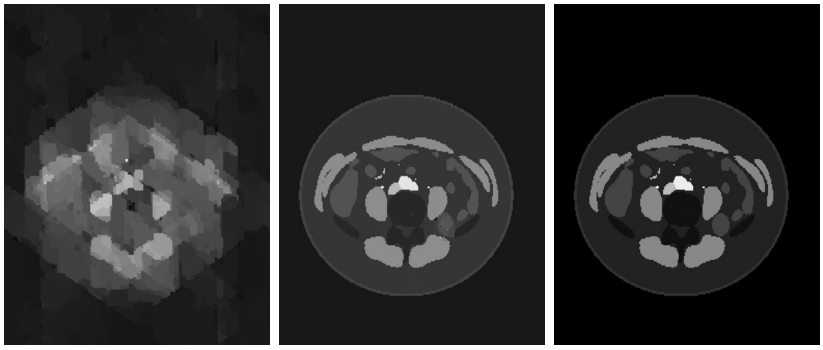
With a view to demonstrating the applicability of (17) and its implementation scheme in (18), we have considered three test images: Shepp-Logan phantom image, abdomen image [41], both are of size 256×256 and a non-phantom MRI (MATLAB sample) image of size 128 × 128. The image (b) has a deformation compared to the image (a) in Figs. 2, 3 and 4. The support of image (a) is considered as the prior support estimate for the reconstruction of image (b) for $\ell_{1-2,w}$. We have minimized $\ell_{1-2,w}$ for discrete weights



(a) Used for support estimate (b) Ground truth (c) Back-Projection



(d) ℓ_1 (e) ℓ_1/ℓ_2 (f) $\ell_1 - \nu\ell_2$



(g) $\ell_{1-2,1}$ (h) $\ell_{1-2,0}$ (i) $\ell_{1-2,w=0.1}$

Fig. 3. Reconstruction of an abdomen image from 3 projections with different methods. Clearly better reconstruction is provided by the $\ell_{1-2,w}$ method at $w = 0.1$.

Algorithm 3: The split Bregman method for (18).

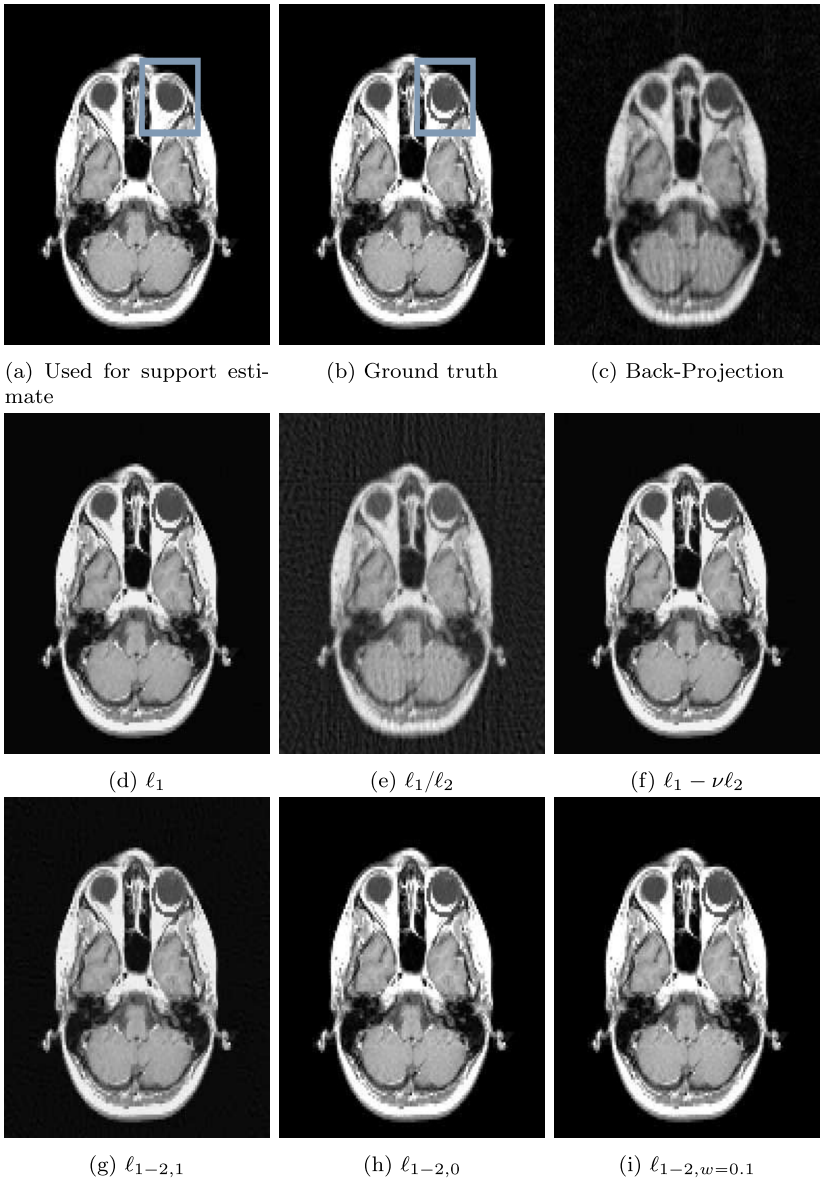
Initialize $u = d_x = d_y = b_x = b_y = 0$, $z = f$.
for *1 to* M_1 **do**
  for *1 to* M_2 **do**
    $u = (\mu\mathcal{R}^T\mathcal{R} - \lambda\mathcal{F}\Delta\mathcal{F}^T)^{-1}(\mu\mathcal{F}^T\mathcal{R}z + \lambda D_x^T(d_x - b_x) + \lambda D_y^T(d_y - b_y))$
    $d_x = \mathcal{S}(D_x u + b_x + \frac{t_x^k}{\lambda}, \lambda_w)$
    $d_y = \mathcal{S}(D_y u + b_y + \frac{t_y^k}{\lambda}, \lambda_w)$
    $b_x = b_x + D_x u - d_x$
    $b_y = b_y + D_y u - d_y$
  end
  $z = z + f - \mathcal{R}\mathcal{F}u$
end

$\{0, 0.1, \ldots, 1\}$ and obtained the optimal weight among them which gives the bigger structural similarity (SSIM) [42] index value. Its values range between 0 and 1, and bigger SSIM values correspond to the reconstruction of good quality.

We have compared our solver with the Back-Projection (BP), ℓ_1 [40], ℓ_1/ℓ_2 [38], $(\ell_1 - \nu\ell_2)$ [43], ℓ_{1-2} [9] minimization solvers in the gradient domain. In our simulation work, we have set the parameters to their default values. We have repeated the reconstruction with varying number of projections and the reconstruction results corresponding to the least number of projections that provide higher SSIM values are reported in Table 5. It may be noted that we have obtained better reconstruction results for the image (b) of Fig. 2 and Fig. 3 merely from 3 projections at $w = 0.1$ using $\ell_{1-2,w}$, while the quality of reconstruction by the rest of the methods is below par. Here for $w = 0$, however, in the reconstructed image, the marked portion (where there is growth) is hazy at the borders. The results in [9][38] and [44] have shown faithful MRI reconstruction of Shepp-Logan phantom image with 6, 7, and 8 projections respectively, implying thereby that the $\ell_{1-2,w}$, fares better than the state-of-the-art solvers. It is important to notice here that there is a significant difference between the reconstruction errors using optimal weights and those with the particular values of the weights at 0 and 1. We have also done the comparison for a non-phantom image given in image (b) of Figs. 4 along with the analysis of varying number of projections. The reconstruction results of this third image are given in Fig. 4 and Table 6. It can be observed that, the reconstruction quality of images increases with the number of projections and for each value of number of

Table 5
SSIM values for the reconstructed images (phantom images).

Image	L	Opt. w	SSIM						
			BP	ℓ_1	ℓ_1/ℓ_2	$\ell_1 - \nu\ell_2$	$\ell_{1-2,1}$ $(w=1)$	$\ell_{1-2,0}$ $(w=0)$	$\ell_{1-2,w_{opt}}$
Phantom	3	0.1	0.5261	0.5883	0.0118	0.1169	0.5487	0.9995	0.9998
Abdomen	3	0.1	0.5597	0.2980	0.0394	0.2891	0.2227	0.9921	1



(a) Used for support estimate (b) Ground truth (c) Back-Projection

(d) ℓ_1 (e) ℓ_1/ℓ_2 (f) $\ell_1 - \nu\ell_2$

(g) $\ell_{1-2,1}$ (h) $\ell_{1-2,0}$ (i) $\ell_{1-2,w=0.1}$

Fig. 4. For the test image in (b), reconstruction of a head image from 60 projections with different methods. Optimal reconstruction is obtained when $w = 0.1$. Notwithstanding the near similar appearance, the reconstruction error at optimal w is significantly small as reported in Table 5.

Table 6
Effect of varying number of projections on the reconstruction quality for the image in Fig. 4(b).

L	Opt. w	SSIM						
		BP	ℓ_1	ℓ_1/ℓ_2	$\ell_1 - \nu\ell_2$	$\ell_{1-2,1}$ $(w=1)$	$\ell_{1-2,0}$ $(w=0)$	$\ell_{1-2,w_{opt}}$
40	0	0.0392	0.5370	0.4626	0.5353	0.5230	0.7733	0.7733
50	0	0.0538	00.5518	0.4873	0.5502	0.5402	0.9447	0.9447
60	0.1	0.0666	0.5611	0.5027	0.5589	0.5480	0.9999	1

Table 7
Reconstruction errors provided by the weighted DCA solver, where the weights are determined iteratively using the proposed scheme. The entries in the last column have been obtained experimentally, as discussed in Section 4. This table implies the significance of the scheme proposed in Section 5.

k	Relative ℓ_2-error				
	$w^{(0)}$	$w^{(1)}$	$w^{(2)}$	$w^{(3)}$	Constant weight
20	0.0287	0.0166	0.0163	**0.0163**	0.0248
22	0.0544	0.0228	0.0218	**0.0216**	0.0504
24	0.1338	0.0837	**0.0820**	0.0830	0.1248
26	0.3801	0.3172	**0.3136**	0.3146	0.3469
28	0.5424	**0.4761**	0.4775	0.4797	0.4948
30	0.6673	0.6187	0.6200	0.6218	**0.6135**

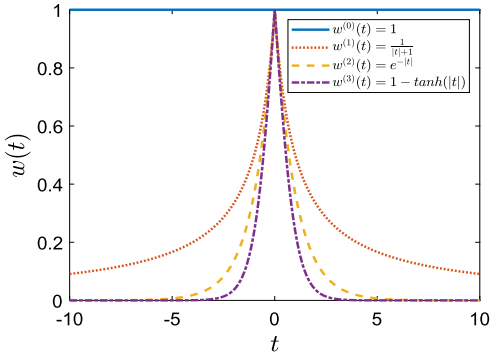


Fig. 5. Plots of the functions used for obtaining the weights iteratively.

projections, $\ell_{1-2,w}$ gives better reconstruction results with optimal weights. When using 60 projections, we have obtained better reconstruction at $w = 0.1$ for this non-phantom image. The SSIM values of the reconstructions provided by other methods are significantly low. From our simulation results, we conclude that the weighted norm setup - a generalization of novel ℓ_{1-2} minimization - has a prominent implication in MRI reconstruction.

Remark 4.1. For the test images in Figs. 2(b) and 3(b), our results provide better reconstruction for some smaller value of w. However, in applications where reconstruction from a small projection set is of prime importance, one may use $\ell_{1-2,w}$ minimization in the following way: Using the image from the previous reconstruction for the prior support estimate and repeating the $\ell_{1-2,w}$ minimization with a few projections by setting weights to 0, and then incrementing them by a small factor till the reconstruction of better quality is obtained. It is also worth highlighting here that our numerical simulations give better reconstruction results for bigger values of w when the image to be reconstructed differs drastically from the image used to generate the partial support estimate. We do not discuss this part herein to avoid presenting too many results.

5. An iterative scheme for updating partial support and weights

Finding the accurate relation between the weights and partial support is rather a challenging problem. Nevertheless, in this section, we provide an iterative scheme for updating the partial support and the weights in the weighted DCA solver. In this scheme, we update the partial support iteratively as a union of the $p\%$ supports of the recovered vectors in the previous two iterations. Here, the $p\%$ support of x, given by a set $T \subset supp(x)$, satisfies $\|x_T\|_2 \geqslant \frac{p}{100}\|x\|_2$ for $p \in [0, 100]$.

Our sparse recovery results in the previous section have shown that the optimal weights are inversely proportional to the accuracy of the partial support, which in turn is proportional to the magnitude of the entries in the partial support. In view of these observations from the previous section, we consider three different iterative schemes for updating the weights in the partial support, for which the weights corresponding to the entries in the partial support remain inversely proportional to their magnitudes. The three functions we consider for updating the weights are

- $w^{(1)}(t) = \frac{1}{|t|+1}$
- $w^{(2)}(t) = e^{-|t|}$
- $w^{(3)}(t) = 1 - tanh(|t|)$.

We denote $w^{(0)}(t) = 1$ to represent the standard weight. It can be seen that $0 \leqslant w^{(3)}(t) \leqslant w^{(2)}(t) \leqslant w^{(1)}(t) \leqslant w^{(0)}(t) = 1,\ \forall\, t \in \mathbb{R}$, which is shown in Fig. 5.

The proposed iterative scheme is as follows: at the k^{th} iteration of the weighted DC algorithm, we update the weights w_i at the i^{th} component of the partial support T based on the values of the components in $x^{(k-1)}$, which is recovered at the $(k-1)^{th}$ iteration. That is, for a given choice of a weight function $w^{(j)}$ for some $j \in \{0,\ 1,\ 2,\ 3\}$, at the k^{th} iteration, the weight w in T is updated as $w^{(j)}(x_*^{(k-1)})$, where $x_*^{(k-1)} = \max_{i \in T} |x_i^{(k-1)}|$. All components of $(x_*^{(k-1)})_T$ are equal to the maximum component of $(x^{(k-1)})_T$ in magnitude. We fix the weight as 1 on T^c as in (7).

Remark 5.1. The iterative re-weighted schemes [27] in the literature are used to approximate the non-convex sparsity promoting functions such as ℓ_p, for $p \in [0, 1)$ in terms of the convex ℓ_1 norm using weights. The objective of the weighted minimization in this section is different as the goal is to incorporate a partial support estimate T into the recovery process by means of weights. Driven by our observations (that the optimal weights are proportional to the accuracy of the partial support information which in turn is proportional to the magnitude of the entries in the partial support) we provide the stated iterative schemes for the updation of weights. Thus a comparison of re-weighted algorithms with the weighted minimization problem discussed here does not arise.

5.1. Numerical results with iterative scheme

We have evaluated the performance of the proposed iterative scheme on the randomly over-sampled DCT matrix of size 100×2000 with $F = 20$, whose average mutual coherence is 0.9999. We have drawn the vectors from the normal distribution with a minimum separation of $2F$ and various sparsity values k ranging from 20 to 30. We have fixed the $p\%$ support as 90%. The Table 7 shows the average of the relative ℓ_2 norm reconstruction errors over 100 trials for different weight functions. We have also repeated the experiment with constant weights as in Section 4 while keeping an iterative selection for the partial support described in this current subsection. It is clear from the Table 7 that the proposed iterative scheme gives better results than the standard ℓ_{1-2} minimization which corresponds to the weight $w^{(0)} = 1$. For smaller sparsity, $w^{(3)}$ provides good reconstruction results and with the increase in sparsity, $w^{(2)}$ and $w^{(1)}$ give better recon-

struction results. When sparsity is 30, the reconstruction using the constant weights provides better reconstruction results. It can be observed that the exponential weights $w^{(2)} = e^{-|.|}$ give the least or comparable reconstruction errors at all sparsity levels.

6. Conclusion

The work reported in this paper has addressed issues such as the convergence of the iterative solver and the significance of general weights pertaining to the weighted ℓ_{1-2} minimization problem. While highlighting the importance of general weights, we have demonstrated their relevance through extensive simulation work, including reconstruction in MRI. We have shown that the weights different from 0 and 1 result in the reconstruction of better quality when we do not have fully accurate or fully corrupt prior partial support information. From our simulation work, we have drawn observations on the dependence of optimal w on the values of associated parameters like accuracy and relative size of prior support information with respect to the actual support of the signal, sparsity of the signal to be recovered, and properties of underlying sensing matrix (such as its coherence, size, etc). Alongside, we have proposed a numerical scheme for updating the partial support and weights iteratively, which is shown to provide good reconstruction results for highly coherent matrices. This approach for the selection of weights is different from the standard trial and error method for finding optimal weights in the weighted minimization problems in CS. On the theoretical front, however, the convergence of the scheme proposed in Section 5 is needed to be established, which requires further analysis.

Our future works will include further analysis and improvement of the proposed iterative scheme for updating the partial support and weights, and for establishing its convergence guarantees. Further, we will also work on the advantage of the weighted minimization techniques in other applied domains where some prior support estimates can be obtained.

CRediT authorship contribution statement

Both K. Z. Najiya and C. S. Sastry have equal contributions in this paper.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: K. Z. Najiya reports financial support was provided by UGC INDIA. C. S. Sastry reports financial support was provided by CSIR INDIA.

Data availability

Data will be made available on request.

Acknowledgment

KZN and CSS gratefully acknowledge the supports received, respectively, from UGC (No. 409284), India, and CSIR (No. 25(0309)/20/EMR-II), India. The authors thank Prof. Penghang Yin for providing us with the code of their work in [9].

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.dsp.2022.103833.

References

[1] M. Iliadis, L. Spinoulas, A. Katsaggelos, Deepbinarymask: learning a binary mask for video compressive sensing, Digit. Signal Process. 96 (2020) 102591.
[2] Z.-X. Cui, Q. Fan, A nonconvex nonsmooth regularization method for compressed sensing and low-rank matrix completion, Digit. Signal Process. 62 (2017) 110–111.
[3] R.R. Naidu, P. Jampana, C.S. Sastry, Deterministic compressed sensing matrices: construction via Euler squares and applications, IEEE Trans. Signal Process. 64 (14) (2016) 3566–3575, https://doi.org/10.1109/TSP.2016.2550020.
[4] B. Tan, Y. Li, S. Ding, I. Paik, A. Kanemura, Dc programming for solving a sparse modeling problem of video key frame extraction, Digit. Signal Process. 83 (2018) 214–222, https://doi.org/10.1016/j.dsp.2018.08.005.
[5] M.P. Friedlander, H. Mansour, R. Saab, Ö. Yilmaz, Recovering compressively sampled signals using partial support information, IEEE Trans. Inf. Theory 58 (2) (2012) 1122–1134.
[6] N. Vaswani, W. Lu, Modified-CS: modifying compressive sensing for problems with partially known support, in: 2009 IEEE International Symposium on Information Theory, 2009, pp. 488–492.
[7] R. von Borries, C.J. Miosso, C. Potes, Compressed sensing using prior information, in: 2007 2nd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2007, pp. 121–124.
[8] E. Esser, Y. Lou, J. Xin, A method for finding structured sparse solutions to non-negative least squares problems with applications, SIAM J. Imaging Sci. 6 (2013) 2010–2046, https://doi.org/10.1137/13090540X.
[9] P. Yin, Y. Lou, Q. He, J. Xin, Minimization of ℓ_{1-2} for compressed sensing, SIAM J. Sci. Comput. 37 (2015) A536–A563.
[10] Y. Lou Yan, Fast L_1-L_2 minimization via a proximal operator, J. Sci. Comput. 74 (2018) 767–785, https://doi.org/10.1007/s10915-017-0463-2.
[11] W. Wang, J. Wang, Z. Zhang, Robust signal recovery with highly coherent measurement matrices, IEEE Signal Process. Lett. 24 (3) (2017) 304–308.
[12] Y. Lou, P. Yin, Q. He, J. Xin, Computing sparse representation in a highly coherent dictionary based on difference of l_1 and l_2, J. Sci. Comput. 64 (2014) 178–196, https://doi.org/10.1007/s10915-014-9930-1.
[13] H. Ge, J. Wen, W. Chen, The null space property of the truncated ℓ_{1-2}-minimization, IEEE Signal Process. Lett. 25 (8) (2018) 1261–1265.
[14] P. Geng, W. Chen, Unconstrained $\ell_1 - \ell_2$ minimization for sparse recovery via mutual coherence, Math. Found. Comput. 3 (2020) 65–79, https://doi.org/10.3934/mfc.2020006.
[15] Z. He, H. He, X. Liu, J. Wen, An improved sufficient condition for sparse signal recovery with minimization of $L1 - L2$, IEEE Signal Process. Lett. 29 (2022) 907–911.
[16] W. Wang, J. Zhang, Performance guarantees of regularized ℓ_{1-2}-minimization for robust sparse recovery, Signal Process. 201 (2022) 108730.
[17] N. Bi, W.-S. Tang, A necessary and sufficient condition for sparse vector recovery via $\ell_1 - \ell_2$ minimization, Appl. Comput. Harmon. Anal. 56 (2022) 337–350.
[18] T.-H. Ma, Y. Lou, T.-Z. Huang, Truncated ℓ_{1-2} models for sparse recovery and rank minimization, SIAM J. Imaging Sci. 10 (2017) 1346–1380, https://doi.org/10.1137/16M1098929.
[19] K. Bui, F. Park, S. Zhang, Y. Qi, J. Xin, Structured sparsity of convolutional neural networks via nonconvex sparse group regularization, Front. Appl. Math. Stat. (2021) 62.
[20] K. Bui, F. Park, S. Zhang, Y. Qi, J. Xin, Improving network slimming with nonconvex regularization, IEEE Access 9 (2021) 115292–115314.
[21] J. Zhang, S. Zhang, W. Wang, Robust signal recovery for ℓ_{1-2} minimization via prior support information, Inverse Probl. 37 (11) (2021) 115001.
[22] H. Ge, J. Wen, W. Chen, The null space property of the truncated ℓ_{1-2} - minimization, IEEE Signal Process. Lett. 25 (2018) 1779–1783, https://doi.org/10.1109/LSP.2018.2852138.
[23] J. Zhang, S. Zhang, Recovery analysis for ℓ_2/ℓ_{1-2} minimization via prior support information, Digit. Signal Process. 121 (2021) 103315, https://doi.org/10.1016/j.dsp.2021.103315.
[24] S. Foucart, H. Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhäuser, Basel, 2013.
[25] P. Theeda, P. Kumar, C.S. Sastry, P.V. Jampana, Reconstruction of sparse-view tomography via preconditioned radon sensing matrix, J. Appl. Math. Comput. 59 (1) (2019) 285–303.
[26] M.A. Khajehnejad, W. Xu, A.S. Avestimehr, B. Hassibi, Weighted ℓ_1 minimization for sparse recovery with prior information, in: 2009 IEEE International Symposium on Information Theory, 2009, pp. 483–487.
[27] E. Candès, M. Wakin, S. Boyd, Enhancing sparsity by reweighted ell_1 minimization, J. Fourier Anal. Appl. 14 (2007) 877–905, https://doi.org/10.1007/s00041-008-9045-x.
[28] B.S. He, X. Yuan, On the $o(1/n)$ convergence rate of the Douglas-Rachford alternating direction method, SIAM J. Numer. Anal. 50 (2012) 700–709, https://doi.org/10.1137/110836936.
[29] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (1) (2011) 1–122.

[30] T.P. Dinh, H.L. Thi, A dc optimization algorithm for solving the trust-region subproblem, SIAM J. Optim. 8 (1998) 476–505, https://doi.org/10.1007/s10107-018-1236-x.
[31] A. Fannjiang, W. Liao, Coherence pattern-guided compressive sensing with unresolved grids, SIAM J. Imaging Sci. 5 (2012) 179–202.
[32] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (2011) 1–122, https://doi.org/10.1561/2200000016.
[33] M.-J. Lai, Y. Xu, W. Yin, Improved iteratively reweighted least squares for unconstrained smoothed ℓ_q minimization, SIAM J. Numer. Anal. 51 (2013) 927–957, https://doi.org/10.1137/110840364.
[34] D. Needell, J. Tropp, Cosamp: iterative signal recovery from incomplete and inaccurate samples, Appl. Comput. Harmon. Anal. 26 (2009) 301–321.
[35] Z. Xu, X. Chang, F. Xu, H. Zhang, $L_{1/2}$ regularization: a thresholding representation theory and a fast solver, IEEE Trans. Neural Netw. Learn. Syst. 23 (2012) 1013–1027, https://doi.org/10.1109/TNNLS.2012.2197412.
[36] M.S. Asif, J. Romberg, Sparse recovery of streaming signals using ℓ_1-homotopy, IEEE Trans. Signal Process. 62 (16) (2014) 4209–4223, https://doi.org/10.1109/TSP.2014.2328981.
[37] S. Zhang, J. Xin, Minimization of transformed l_1 penalty: theory, difference of convex function algorithm, and robust application in compressed sensing, Math. Program. 169 (2018) 307–336, https://doi.org/10.1007/s10107-018-1236-x.
[38] Y. Rahimi, C. Wang, H. Dong, Y. Lou, A scale-invariant approach for sparse signal recovery, SIAM J. Sci. Comput. 41 (2019) A3649–A3672, https://doi.org/10.1137/18M123147X.
[39] N. Feng, J. Wang, W. Wang, Sparse signal recovery with prior information by iterative reweighted least squares algorithm, J. Inverse Ill-Posed Probl. 26 (2018) 171–184.
[40] T. Goldstein, S. Osher, The split Bregman method for ℓ_1-regularized problems, SIAM J. Imaging Sci. 2 (2) (2009) 323–343.
[41] W.C. Lo, Y. Chen, Y. Jiang, J. Hamilton, R. Grimm, M. Griswold, V. Gulani, N. Seiberlich, Realistic 4d mri abdominal phantom for the evaluation and comparison of acquisition and reconstruction techniques, Magn. Reson. Med. 81 (2019) 1863–1875, https://doi.org/10.1002/mrm.27545.
[42] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[43] Y. Lou, T. Zeng, S. Osher, J. Xin, A weighted difference of anisotropic and isotropic total variation model for image processing, SIAM J. Imaging Sci. 8 (2015) 1798–1823, https://doi.org/10.1137/14098435X.
[44] W. Guo, W. Yin, Edge guided reconstruction for compressive imaging, SIAM J. Imaging Sci. 5 (2012) 809–834.

K.Z. Najiya is currently a postdoctoral fellow at the Indian Institute of Technology Hyderabad, India. She received Ph.D. degree in Mathematics from the Indian Institute of Technology Hyderabad, India, in 2022. Her research interests include Compressed Sensing, Neural Networks, Computed Tomography, and Wavelets.

C.S. Sastry received Ph.D. in Applied Mathematics from the Indian Institute of Technology, Kanpur. He is currently on the faculty at the Indian Institute of Technology, Hyderabad. His research interests include Wavelets, Inverse Problems, and Sparse Optimization Theory.