Block-sparse compressed sensing: non-convex model and iterative re-weighted algorithm

Haitao Yinab, Shutao Lib* and Leyuan Fangb

aThe State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha 410082, China; bCollege of Electrical and Information Engineering, Hunan University, Changsha 410082, China

(Received 17 January 2011; final version received 15 March 2012)

Compressed sensing is a new sampling technique which can exactly reconstruct sparse signal from a few measurements. In this article, we consider the block-sparse compressed sensing with special structure assumption about the signal. A novel non-convex model is proposed to reconstruct the block-sparse signals. In addition, the conditions of the proposed model for recovering the block-sparse noise or noise-free signals are presented. The experimental results demonstrate that the proposed non-convex method surpasses the convex method (the mixed $\ell_2$/\ell_1\norm$ optimization) and some algorithms without considering the block-sparse structure (the $\ell_1$- and $\ell_p$-norm optimization).

Keywords: block-sparse compressed sensing; non-convex model; block-sparse recovery; iterative re-weighted algorithm

1. Introduction

The traditional signal acquisition must follow the Nyquist–Shannon sampling theorem. Based on this theorem, the sample rate is at least twice of the signal bandwidth for exact reconstruction. In many recent applications, the Nyquist rate may result in high storage and transmission pressure. In addition, the traditional signal acquisition model does not consider the signal structure known as the prior information. A breakthrough in data acquisition, namely compressed sensing, was brought by Donoho [1] and Candès et al. [2,3], which has been widely applied in magnetic resonance imaging [4] and radar imaging [5]. Compressed sensing assumes that the original signal is sparse or approximately sparse under some bases. Without loss of generality, we assume that the original signal is sparse under the canonical basis. Then, the $K$-sparse signal means that there are no more than $K$ nonzero entries in the signal. Let $x \in \mathbb{R}^N$ be the $K$-sparse signal. The task of compressed sensing is to recover $x$ from the measurements $y$ calculated as

$$y = \Phi x.$$  (1)

Here, $\Phi \in \mathbb{R}^{M \times N} (M < N)$ is the measurement matrix with more columns than rows. Obviously, Equation (1) is underdetermined. However, the theory of compressed sensing...
indicates that the sparse or compressible signal can be exactly reconstructed from a few measurements $y$ through the following optimization problem:

$$\min_x \|x\|_0 \quad \text{s.t.} \quad y = \Phi x,$$

where $\| \cdot \|_0$ denotes the $\ell_0$-norm counting the number of nonzero entries. However, the problem (2) is a combinatorial optimization problem. Solving it directly is general NP-hard. Therefore, some sub-optimization strategies are proposed to approximate the exact solution.

Popular sub-optimization strategies include greedy algorithms (such as the orthogonal matching pursuit [6]), and relaxation methods (such as convex and non-convex relaxation methods [7–9]). The representative convex method is the $\ell_1$-optimization problem which is called basis pursuit (BP) [7]

$$\min_x \|x\|_1 \quad \text{s.t.} \quad y = \Phi x,$$

where $\|x\|_1 = \sum_{i=1}^N |x_i|$ is the $\ell_1$-norm of the vector $x$. In the noise case, the problem (3) can be modified as the basis pursuit denoising (BPDN) [7]

$$\min_x \lambda \|x\|_1 + \frac{1}{2} \|y - \Phi x\|_2^2,$$

where $\lambda > 0$ is a tradeoff parameter. The non-convex relaxation method has been studied in [8,9], i.e. the following non-convex optimization problem:

$$\min_x \|x\|_p \quad \text{s.t.} \quad y = \Phi x,$$

where $\|x\|_p = (\sum_{i=1}^N |x_i|^p)^{1/p}$, $0 < p < 1$. In [8], the experimental results show that the non-convex problem (5) can reconstruct the sparse signal from fewer measurements than the convex method acquiring. The restricted isometry property (RIP) [3] plays a crucial role in the theory of compressed sensing. The previous results show that if the measurement matrix has the RIP with a certain parameters, the convex and non-convex optimization methods can recover the sparse signal exactly.

The existed analysis of compressed sensing does not consider the special inherent structures of signals. This article considers a variation of the compressed sensing problem named the block-sparse compressed sensing. In block-sparse compressed sensing, the block structure of sparse signal is investigated, i.e. the nonzero entries of the signal occur in clusters. Various practical signals exhibit the block structure, such as DNA microarrays [10], multiband signal [11] and magnetoencephalography (MEG) [12]. The framework of block-sparse compressed sensing was introduced and studied first in refs [13–15]. The mixed $\ell_2/\ell_1$-norm optimization is developed to recover the block-sparse signal in [13], which is a convex method. In this article, we propose a novel non-convex model termed as the mixed $\ell_2/\ell_p$ ($0 < p < 1$)-norm optimization for block-sparse signal recovery, which can greatly improve the performance of the mixed $\ell_2/\ell_1$-norm optimization.

The remainder of this article is organized as follows. In Section 2, we briefly review the block-sparse compressed sensing. In Section 3, the mixed $\ell_2/\ell_p$-norm optimization problem is proposed, and the stability and uniqueness of the proposed model is investigated. In Section 4, the re-weighted $\ell_2/\ell_1$-algorithm is used to solve mixed $\ell_2/\ell_p$-norm optimization problem, and some numerical experiments are reported. Finally, Section 5 concludes this article.
2. Block-sparse compressed sensing and RIP

In this section, we briefly review block sparsity and the block-sparse compressed sensing. In the block-sparse compressed sensing, the signal $x \in \mathbb{R}^N$ is the concatenation of some blocks. Without loss of generality, we assume that there are $n$ blocks with the block size $d = N/n$ in $x$. Then, the signal $x$ can be rewritten as

$$x = (x^T[1], x^T[2], \ldots, x^T[n])^T,$$

where $x[i] = (x_{(i-1)d+1}, x_{(i-1)d+2}, \ldots, x_{id})^T$, $i \in \mathcal{I} \equiv \{1, 2, \ldots, n\}$. Then, the signal $x$ is called $K$-block-sparse if no more than $K$ blocks over $\mathcal{I}$ are nonzero. In [13], a convex optimization problem termed as the mixed $\ell_2/\ell_1$-norm optimization is proposed to reconstruct the block-sparse signal, i.e. the following model:

$$\min_x \|x\|_{(2,1)} \quad \text{s.t. } y = \Phi x,$$

where $\|x\|_{(2,1)} = \sum_{i=1}^{n} \|x[i]\|_2$. In noisy measurements case, the measurements can be calculated as

$$y = \Phi x + e,$$

where $e \in \mathbb{R}^M$ is unknown noise. In this article, we assume $e \sim \mathcal{N}(0, \sigma^2)$, and the standard deviation $\sigma$ denotes the noise level. In this noise case, the block-sparse signal can be reconstructed via the following optimization problem:

$$\min_x \|x\|_{(2,1)} \quad \text{s.t. } \|y - \Phi x\|_2 \leq \varepsilon.$$

To investigate the theory of compressed sensing, Candès and Tao [3] introduce the RIP. The matrix $\Phi$ is said to have the RIP with parameter $\delta_K \in (0, 1)$, if the inequality

$$(1 - \delta_K)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_K)\|x\|_2^2$$

holds for every $K$-sparse signal $x$. Many random matrices have been proven to have the RIP with overwhelming probability, such as Gaussian, Bernoulli and partial Fourier matrices [16]. To study the mixed $\ell_2/\ell_1$-norm optimization problems (7) and (9), Eldar and Mishali [13] introduce the block RIP (B-RIP) which is the block-sparse version of RIP. The matrix $\Phi$ has the B-RIP with parameter $\delta_{K|\mathcal{I}} \in (0, 1)$, if the inequality

$$(1 - \delta_{K|\mathcal{I}})\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_{K|\mathcal{I}})\|x\|_2^2$$

holds for every $K$-block-sparse signal $x$, where $\delta_{K|\mathcal{I}}$ is the B-RIP constant over $\mathcal{I}$. In [13] Eldar and Mishali show that the Gaussian random matrix has the B-RIP with high probability.

3. Non-convex recovery model

In this section, the motivation of proposed non-convex method is stated first. Then, the theoretical results about the uniqueness and robust of the proposed non-convex model to recover the block-sparse noised and noise-free signals are studied. At last, an iterative re-weighted algorithm is presented to solve the non-convex method.
3.1. Motivation

In [13], Eldar and Mishali claim that the block-sparse signal $x$ can be exactly reconstructed from the measurements $y = \Phi x$ by the following optimization problem:

$$\min_x \|x\|_{(2,0)} \quad \text{s.t. } y = \Phi x,$$

(12)

where

$$\|x\|_{(2,0)} = \sum_{i=1}^{n} \varphi(\|x[i]\|_2) \quad \text{and} \quad \varphi(\theta) = \begin{cases} 1, & \text{if } \theta \neq 0 \\ 0, & \text{if } \theta = 0 \end{cases}.$$ 

By introducing auxiliary variables $z_i = \|x[i]\|_2$ ($i = 1, 2, \ldots, n$) and $z = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$, the objective function of problem (12) can be equivalently transformed as $\|z\|_0$. It is most accurate to measure the sparsity by the $\ell_0$-norm. Unfortunately, the $\ell_0$-norm optimization problem is combinatorial optimization, which cannot be solved directly. An effective relaxation strategy of the $\ell_0$-norm to measure the sparsity is the $\ell_p$-norm, i.e. $\|z\|_p^p$. Figure 1 illustrates the one-dimensional trend of $\|z\|_p^p$ for varying $p$. Some previous results have demonstrated that the $\ell_1$-norm can approximate the $\ell_0$-norm effectively [3,13]. However, with $p \to 0$, the $\ell_p$-norm tends to the $\ell_0$-norm. Obviously, the non-convex $\ell_p$ ($0 < p < 1$)-norm can give more powerful sparsity measure than the $\ell_1$-norm. Therefore, we focus on non-convex method in our work. Corresponding to the block-sparse signal $x$, we study the following non-convex method termed as the mixed $\ell_2/\ell_p$-norm optimization:

$$\min_x \|x\|_{(2,p)} \quad \text{s.t. } y = \Phi x,$$

(13)

where

$$\|x\|_{(2,p)} = \sum_{i=1}^{n} \|x[i]\|_2^p, \quad 0 < p < 1.$$ 

Figure 1. The one-dimensional trend of $\|z\|_p^p$ for varying $p$. 

H. Yin et al.
3.2. Exact recovery

In this section, we show that the proposed non-convex model (13) can recover the block-sparse signal \( \mathbf{x} \) exactly in the noise-free case with a certain condition of the measurement matrix \( \Phi \).

**Theorem 1** Let \( \mathbf{y} = \Phi \mathbf{x} \) be the measurements of the \( K \)-block-sparse signal \( \mathbf{x} \). If the measurement matrix \( \Phi \) has the B-RIP with the condition

\[
\tau^{\frac{1}{p}} \delta_{(K+L)|I} + \delta_{L|I} < \tau^{\frac{1}{p}} - 1,
\]

where \( L = \tau K \), \( (\tau > 1) \), then \( \mathbf{x} \) is the unique solution of the problem (13).

3.3. Robust recovery

This section considers the noise case. Assume \( \mathbf{x} \) be an arbitrary signal, and the measurements are computed with noise, i.e.

\[
\mathbf{y} = \Phi \mathbf{x} + \mathbf{e},
\]

where \( \mathbf{e} \in \mathbb{R}^M \) is the Gaussian noise. In this case, the optimization problem (13) must be modified as

\[
\min_{\mathbf{x}} \| \mathbf{x} \|_{(2,p)} \quad \text{s.t.} \quad \| \mathbf{y} - \Phi \mathbf{x} \|_2 \leq \varepsilon. \tag{16}
\]

The following Theorems 2 and 3 state the robustness of the proposed non-convex model. In addition, we prove that the reconstruction error of problem (16) is the linear function of the noise energy \( \varepsilon \).

First, the case of the block-sparse signal with noise is studied.

**Theorem 2** Let \( \mathbf{y} = \Phi \mathbf{x} + \mathbf{e} \) be the noisy measurements of the \( K \)-block-sparse signal \( \mathbf{x} \). If the condition (17) holds

\[
\tau^{\frac{1}{p}} \delta_{(K+L)|I} + \delta_{L|I} < \tau^{\frac{1}{p}} - 1
\]

holds, where \( L = \tau K \), \( (\tau > 1) \). Assume \( \mathbf{x}' \) be the solution of the problem (16). Then,

\[
\| \mathbf{x} - \mathbf{x}' \|_2 \leq C_1^{1/p} \varepsilon,
\]

where

\[
C_1 = \frac{2\left[ 1 + \tau^{p/2-1}(2/p - 1)^{-p/2} \right]}{(1 - \delta_{(K+L)|I})^{p/2} - (1 + \delta_{L|I})^{p/2} \tau^{p/2-1}}.
\]

Afterwards, let \( \mathbf{x} \) and \( \mathbf{x}_{T_0} \) be an arbitrary signal and the \( K \)-block-sparse approximation of \( \mathbf{x} \), respectively. Assume \( T_0 \) be the index set corresponding to the first \( K \) largest blocks in terms of \( \ell_2 \)-norm. Then, we have the following theory for general case.

**Theorem 3** Let \( \mathbf{y} = \Phi \mathbf{x} + \mathbf{e} \) be the noisy measurements of any signal \( \mathbf{x} \). If the condition (17) holds, then the solution \( \mathbf{x}' \) of problem (16) obeys

\[
\| \mathbf{x} - \mathbf{x}' \|_2 \leq \left( C_1 \varepsilon^p + C_2 \left\| \mathbf{x}_{T_0} \right\|_{2,p} \right)^{1/p},
\]
where

\[
C_1 = \frac{2[1 + \tau^{p/2-1}(2/p - 1)^{-p/2}]}{(1 - \delta_{(K+L)J})^{p/2} - (1 + \delta_{LJ})^{p/2} \tau^{p/2-1}},
\]

\[
C_2 = \frac{4[1 + \tau^{p/2-1}(2/p - 1)^{-p/2}][1 + \delta_{LJ}]^{p/2}}{(1 - \delta_{(K+L)J})^{p/2} L^{1-p/2} - (1 + \delta_{LJ})^{p/2} K^{1-p/2}} + \frac{2}{L^{1-p/2}(2/p - 1)^{p/2}}.
\]

**Proof** Following the method of proof stated in [9], let \( h = x' - x \) be the difference between the true signal and the reconstructed signal. The set \( T_0 \) can be divided as \( T_0 = \cup_{j=1}^J T_j \), where \( T_j \) (\( j \geq 1 \)) is the indices set which is make up of \( L = \tau K, (\tau > 1) \) blocks, chosen such that the \( \ell_2 \)-norm of \( h_{T_0} \) over \( T_1 \) is largest, and the \( \ell_2 \)-norm of \( h_{T_0} \) over \( T_2 \) is the second largest, and so on. So, we can decompose \( h \) as \( h = \sum_{j=0}^J h_j \).

Based on the B-RIP and the triangle inequality of \( \| \cdot \|_2^p (p \in (0, 1)) \), we obtain the lower bound of \( \| \Phi h \|_2^p \) as follows:

\[
\| \Phi h \|_2^p \geq \| \Phi_{T_0 \cup T_1} h_{T_0 \cup T_1} \|_2^p - \sum_{j=2}^J \| \Phi_{T_j} h_{T_j} \|_2^p.
\]

\[
\geq \| \Phi_{T_0 \cup T_1} h_{T_0 \cup T_1} \|_2^p - \sum_{j=2}^J \| \Phi_{T_j} h_{T_j} \|_2^p
\geq (1 - \delta_{(K+L)J})^{p/2} \| h_{T_0 \cup T_1} \|_2^p - (1 + \delta_{LJ})^{p/2} \sum_{j=2}^J \| h_{T_j} \|_2^p.
\]

(18)

Next, we will give the bounds of \( \| h_{T_0 \cup T_1} \|_2^p \) and \( \sum_{j=2}^J \| h_{T_j} \|_2^p \).

For all \( t \in T_j, s \in T_{j-1}, j \geq 2 \), \( \| h[t] \|_2 \leq \| h[s] \|_2 \) holds. Following this inequality, we have

\[
\| h[t] \|_2^p \leq \| h[s] \|_2^p \leq \| h_{T_{j-1}} \|_{2,p}/L,
\]

\[
\| h_{T_j} \|_2^p \leq \frac{1}{L^{2/p-1}} \| h_{T_{j-1}} \|_{2,p}^{2/p}, \quad \| h_{T_j} \|_2^p \leq \frac{1}{L^{1-p/2}} \| h_{T_{j-1}} \|_{2,p}.
\]

Thus, we obtain

\[
\sum_{j=2}^J \| h_{T_j} \|_2^p \leq \frac{1}{L^{1-p/2}} \sum_{j=1}^{J-1} \| h_{T_j} \|_{2,p} \leq \frac{1}{L^{1-p/2}} \sum_{j=1}^J \| h_{T_j} \|_{2,p}.
\]

Due to \( h = x' - x \),

\[
\| x \|_{2,p} \geq \| x' \|_{2,p} = \| x + h \|_{2,p} = \| x_{T_0} + h_{T_0} \|_{2,p} + \| x_{T_0} + h_{T_0} \|_{2,p} \geq \| x_{T_0} \|_{2,p} - \| h_{T_0} \|_{2,p} + \| h_{T_0} \|_{2,p} - \| x_{T_0} \|_{2,p}.
\]
Inverse Problems in Science and Engineering

Therefore,
\[
\left\| hT_0 \right\|_{(2,p)} \leq \left\| hT_0 \right\|_{(2,p)} + 2 \left\| xT_0 \right\|_{(2,p)}.
\]

Applying the Holder inequality, the \(\|hT_0\|_{2,p}\) is bounded as the \(\ell_2\)-norm, i.e.
\[
\|hT_0\|_{(2,p)} = \sum_{j \epsilon T_0} \|h[j]\|_2^p \cdot 1
\leq \left( \sum_{j \epsilon T_0} \|h[j]\|_2^{2/p} \right)^{p/2} \cdot \left( \sum_{j \epsilon T_0} 1 \right)^{1-p/2}
= K^{1-p/2} \left( \sum_{j \epsilon T_0} \|h[j]\|_2^2 \right)^{p/2}
= K^{1-p/2} \left( \|hT_0\|_2^2 \right)^{p/2} = K^{1-p/2} \|hT_0\|_2^p.
\]

Thus, we have
\[
\sum_{j=2}^{J} \|hT_j\|_2^p \leq \frac{1}{L^{1-p/2}} \|hT_0\|_{(2,p)}
\leq \frac{1}{L^{1-p/2}} \left( \|hT_0\|_{(2,p)} + 2 \left\| xT_0 \right\|_{(2,p)} \right)
\leq \frac{1}{L^{1-p/2}} \left( K^{1-p/2} \|hT_0\|_2^2 + 2 \left\| xT_0 \right\|_{(2,p)} \right)
\leq \left( \frac{K}{L} \right)^{1-p/2} \left\| hT_0 \right\|_{(2,p)} + \frac{2}{L^{1-p/2}} \left\| xT_0 \right\|_{(2,p)}.
\]

(19)
Combining (18) and (19), the following inequality holds
\[
\|\Phi h\|_2^p \geq C \left\| hT_{0 \cup T_1} \right\|_2^p - e_0,
\]
where
\[
C \equiv \left( 1 - \delta_{(K+L)I} \right)^{p/2} - \left( 1 + \delta_{L|I} \right)^{p/2} \left( \frac{K}{L} \right)^{1-p/2},
\]
\[
e_0 \equiv \left( 1 + \delta_{L|I} \right)^{p/2} \frac{2}{L^{1-p/2}} \left\| x_{T_0} \right\|_{(2,p)}.
\]

Furthermore, \(\|\Phi h\|_2^p\) has the upper bound as
\[
\|\Phi h\|_2^p \leq \|\Phi (x' - x)\|_2^p \leq \|y - \Phi x'\|_2^p + \|y - \Phi x\|_2^p \leq 2\varepsilon^p.
\]

Then, we have \(C \left\| hT_{0 \cup T_1} \right\|_2^p - e_0 \leq 2\varepsilon^p\). That is, the upper bound of \(\|hT_{0 \cup T_1}\|_2^p\) is
\[
\|hT_{0 \cup T_1}\|_2^p \leq [2\varepsilon^p + e_0] / C.
\]

(20)
By the condition of (17), we know that \(C > 0\). Hence, the inequality (20) holds.
In this section, we propose the re-weighted algorithm. For $\mathbf{h}(T_0 \cup T_1)'[V] = \mathbf{h}_T[I + L]$, the $\|\mathbf{h}(T_0 \cup T_1)'\|_2$ is bounded as

$$\|\mathbf{h}(T_0 \cup T_1)'\|_2^2 = \sum_i \|\mathbf{h}_T[i]\|_2^2 \leq \left( \frac{\|\mathbf{h}_T\|_{(2,p)}}{L^{1-p/2}(2/p - 1)^{p/2}} \right)^{2/p}.$$ 

Finally, we have

$$\|\mathbf{h}\|_2^2 \leq \|\mathbf{h}(T_0 \cup T_1)\|_2^2 + \|\mathbf{h}(T_0 \cup T_1)'\|_2^2 \leq \left( \|\mathbf{h}_T\|_{(2,p)} \right)^{2/p} + \left( \frac{\|\mathbf{h}_T\|_{(2,p)} + 2\|\mathbf{x}_T\|_{(2,p)}}{L^{1-p/2}(2/p - 1)^{p/2}} \right)^{2/p} \leq \left( \frac{K^{1-p/2}\|\mathbf{h}_T\|_2 + 2\|\mathbf{x}_T\|_2}{L^{1-p/2}(2/p - 1)^{p/2}} \right)^{2/p} \leq \left( \frac{K^{1-p/2}\|\mathbf{T_0 \cup T_1}\|_2 + 2\|\mathbf{x}_T\|_{(2,p)}}{L^{1-p/2}(2/p - 1)^{p/2}} \right)^{2/p} \leq \left\{ 1 + \left( \frac{K}{L} \right)^{1-p/2}(2/p - 1)^{-p/2} \right\} \|\mathbf{T_0 \cup T_1}\|_2^2 + \left( \frac{2}{L^{1-p/2}(2/p - 1)^{p/2}} \right)^{2/p} \|\mathbf{x}_T\|_{(2,p)}^2 \leq C_1 \gamma^p + C_2 \|\mathbf{x}_T\|_{(2,p)}^2 \leq \left( \frac{2}{L^{1-p/2}(2/p - 1)^{p/2}} \right)^{2/p} \|\mathbf{x}_T\|_{(2,p)}^2 \right\}^{2/p}.$$ 

**Remark 1** Theorems 1 and 2 are the special cases of Theorem 3. When the original signal $x$ is a $K$-block-sparse signal, then $\|\mathbf{x}_T\|_{(2,p)} = 0$ yields Theorem 2. Furthermore, Theorem 1 is the $K$-block-sparse signal and noise-free case of Theorem 3.

3.4. **The iterative re-weighted $\ell_2/\ell_1$-algorithm**

In this section, we propose the re-weighted $\ell_2/\ell_1$-algorithm for solving the $\ell_2/\ell_p$-norm optimization deriving from the re-weighted $\ell_1$-algorithm [17]. Due to

$$\|\mathbf{x}[\delta]\|_2^p = \lim_{\delta \to 0} \frac{\|\mathbf{x}[\delta]\|_2}{(\|\mathbf{x}[\delta]\|_2 + \delta)^{1-p}},$$
the main steps of the iterative re-weighted $\ell_2/\ell_1$-algorithm are summarized in Algorithm 1.

**Algorithm 1** The iterative re-weighted $\ell_2/\ell_1$-algorithm

**Input:** $k = 0$, $\omega_j^{(0)} = 1$, $i = 1, 2, \ldots, n$, $\delta = 1$, $\xi$, $\text{Num}$.

**For** $i=1: \text{Num}$

**While** $\|x^{(k+1)} - x^{(k)}\|_2 / \|x^{(k)}\|_2 > \xi$, **do**

(1) Solve the weighted $\ell_2/\ell_1$-norm optimization

$$x^{k+1} \leftarrow \min_x \sum_{i=1}^{n} \omega_i \|x[i]\|_2$$  
subject to $y - \Phi x \leq \varepsilon$.

(2) Update the weights

$$\omega_i^{k+1} = \frac{1}{\left(\|x^{k+1}[i]\|_2 + \delta\right)^{1-p}},$$

**End while**

$$\delta \leftarrow \delta / 10.$$

**End for**

**Output:** The block-sparse signal $x$.

The weighted $\ell_2/\ell_1$-norm optimization

$$\min_x \sum_{i=1}^{n} \omega_i \|x[i]\|_2$$  
subject to $y - \Phi x \leq \varepsilon$

can be equivalently transformed to a second-order cone programming (SOCP):

$$\min_{c, t_i} \sum_{i=1}^{n} t_i$$

subject to $y - \Phi c \leq \varepsilon,$

$$t_i \geq \omega_i \|c[i]\|_2, \quad 1 \leq i \leq n.$$

Furthermore, the ‘cvx’ software packages (the code is available at http://cvxr.com/cvx/) can be used to solve the above SOCP.

4. Experiments

In this section, some numerical experiments are presented to evaluate the performance of the mixed $\ell_2/\ell_p$-norm optimization. First, the experimental settings are presented. Then, the performance of the mixed $\ell_2/\ell_p$-norm optimization is compared with three popular
methods, including the mixed $\ell_2/\ell_1$-norm optimization, $\ell_1$-norm optimization and the $\ell_p$-norm optimization.

4.1. Experimental settings

Three numerical experiments are performed to assess the capacity of the mixed $\ell_2/\ell_p$-norm optimization: (a) the noise-free case; (b) the noise case; (c) the large scale problem. Therefore, in this section, we describe the experiments setting, including the generation of the experiment data and the parameters setting.

We assume that the synthetic signal has $n$ blocks. Then, the synthetic block-sparse signals are generated as follows: (1) the $K$ blocks are selected from these $n$ blocks randomly; (2) the entries restricted to these $K$ blocks are generated from a standard Gaussian distribution and the entries out of the selected $K$ blocks are set to zero. In our experiments, the Gaussian random matrix is used as the measurement matrix. For different experiment cases, we set the parameters of the re-weighted $\ell_2/\ell_1$-algorithm listed in Table 1. In addition, the parameters for other tested algorithms are set according to refs [7,8,13]. The relative error and probability of exact reconstruction are used as the evaluation indices. The relative error between the reconstructed signal $x'$ and the original signal $x$ is calculated as

$$\text{Error} = \frac{\|x' - x\|_2}{\|x\|_2}.$$ 

The signal is said to be exactly reconstructed, if the relative error between the reconstructed signal $x'$ and the original signal $x$ obeys $\text{Error} \leq 10^{-5}$.

4.2. Experimental results

Firstly, we consider the noise-free case. The dimension of signal is $N = 150$, and the block size is $d = 5$. The comparison experiments about recovering the block-sparse signal from the $M = 80$ measurements are performed. The active blocks vary from 1 to 18. For each block-sparsity level, the tested algorithms are all repeated 500 times. The probability of exact reconstruction is used to valuate the performance of the algorithms. The proposed method is compared with the mixed $\ell_2/\ell_1$-norm optimization, BP and the $\ell_p(p = 0.5)$-method. The exact reconstruction performances are presented in Figure 2. The horizontal axis of Figure 2 denotes the block-sparsity level $K$, and the vertical axis is the probability of exact reconstruction. When the block-sparsity level is greater than 6, the BP, $\ell_p(p = 0.5)$ and mixed $\ell_2/\ell_1$-norm optimization methods cannot recover the signal exactly, but the proposed method can recover the signal exactly up to 9-block-sparse level. In addition, when the block-sparsity level is greater than 11, the BP, $\ell_p(p = 0.5)$ and mixed $\ell_2/\ell_1$-norm

Table 1. The parameters of the re-weighted $\ell_2/\ell_1$-algorithm for three different experiments.

<table>
<thead>
<tr>
<th></th>
<th>Noise-free case</th>
<th>Noise case</th>
<th>Large scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi$</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td></td>
<td>$10^{-3}$</td>
<td>$10^{-3}$</td>
</tr>
</tbody>
</table>

H. Yin et al.
optimization methods cannot recover the signal, but the proposed non-convex method can also recover the signal exactly with high probability. Therefore, from Figure 2, it can be seen that the proposed non-convex method is more accurate and effective than other three methods. Especially, the capacity of the mixed $\ell_2/\ell_p$-norm optimization is improved with the decreasing $p$.

Secondly, the case of noisy measurements is studied. The dimension of the simulation signal is $N = 1024$, and the block number is $n = 32$. The relative error is used to evaluate the performance of the test algorithms from $M = 512$ measurements. The noise is assumed as i.i.d. Gaussian with zero mean and standard deviation $\sigma$. The experimental results about

![Figure 2. The exact reconstruction performance of four algorithms in noise-free case.](image)

Table 2. The comparison results ($N = 1024$, $M = 512$; average of 20 runs).

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>$\ell_2/\ell_1$</th>
<th>$\ell_2/\ell_1$ ($p = 0.5$)</th>
<th>$\ell_2/\ell_1$ ($p = 0.3$)</th>
<th>$\ell_2/\ell_1$ ($p = 0.5$)</th>
<th>$\ell_2/\ell_1$ ($p = 0.7$)</th>
<th>$\ell_2/\ell_1$ ($p = 0.9$)</th>
<th>$\ell_0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>1.18e$-3$</td>
<td><strong>9.11e$-5$</strong></td>
<td><strong>9.11e$-5$</strong></td>
<td><strong>9.12e$-5$</strong></td>
<td><strong>9.15e$-5$</strong></td>
<td><strong>9.70e$-5$</strong></td>
<td>4.40e$-2$</td>
</tr>
<tr>
<td>0.01</td>
<td>1.30e$-3$</td>
<td>9.52e$-4$</td>
<td>9.50e$-4$</td>
<td><strong>9.49e$-4$</strong></td>
<td>9.51e$-4$</td>
<td>1.00e$-3$</td>
<td>6.00e$-3$</td>
</tr>
<tr>
<td>0.05</td>
<td>6.20e$-3$</td>
<td>6.40e$-3$</td>
<td>5.70e$-3$</td>
<td>5.00e$-3$</td>
<td><strong>4.90e$-3$</strong></td>
<td>5.20e$-3$</td>
<td>1.47e$-2$</td>
</tr>
<tr>
<td>0.1</td>
<td>1.24e$-2$</td>
<td>1.36e$-2$</td>
<td>1.23e$-2$</td>
<td>1.13e$-2$</td>
<td><strong>1.02e$-2$</strong></td>
<td>1.04e$-2$</td>
<td>2.61e$-2$</td>
</tr>
<tr>
<td>$K = 7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>1.48e$-4$</td>
<td><strong>8.80e$-5$</strong></td>
<td><strong>8.80e$-5$</strong></td>
<td><strong>8.81e$-5$</strong></td>
<td><strong>8.87e$-5$</strong></td>
<td><strong>9.87e$-5$</strong></td>
<td>2.14e$-1$</td>
</tr>
<tr>
<td>0.01</td>
<td>1.40e$-3$</td>
<td>8.77e$-4$</td>
<td>8.75e$-4$</td>
<td><strong>8.72e$-4$</strong></td>
<td><strong>8.74e$-4$</strong></td>
<td>9.62e$-4$</td>
<td>1.93e$-1$</td>
</tr>
<tr>
<td>0.05</td>
<td>7.30e$-3$</td>
<td>6.00e$-3$</td>
<td>5.30e$-3$</td>
<td>4.70e$-3$</td>
<td><strong>4.61e$-3$</strong></td>
<td>5.00e$-3$</td>
<td>2.22e$-1$</td>
</tr>
<tr>
<td>0.1</td>
<td>1.43e$-2$</td>
<td>1.26e$-2$</td>
<td>1.18e$-2$</td>
<td>1.07e$-2$</td>
<td><strong>9.70e$-3$</strong></td>
<td>1.02e$-2$</td>
<td>1.96e$-1$</td>
</tr>
<tr>
<td>$K = 9$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>8.90e$-3$</td>
<td>9.03e$-5$</td>
<td><strong>9.02e$-5$</strong></td>
<td><strong>9.02e$-5$</strong></td>
<td><strong>9.08e$-5$</strong></td>
<td>1.08e$-4$</td>
<td>4.06e$-1$</td>
</tr>
<tr>
<td>0.01</td>
<td>5.80e$-3$</td>
<td>9.16e$-4$</td>
<td>9.14e$-4$</td>
<td><strong>9.13e$-4$</strong></td>
<td>9.19e$-4$</td>
<td>1.10e$-3$</td>
<td>3.74e$-1$</td>
</tr>
<tr>
<td>0.05</td>
<td>1.72e$-2$</td>
<td>6.00e$-3$</td>
<td>5.40e$-3$</td>
<td>4.70e$-3$</td>
<td><strong>4.60e$-3$</strong></td>
<td>5.40e$-3$</td>
<td>3.87e$-1$</td>
</tr>
<tr>
<td>0.1</td>
<td>3.18e$-2$</td>
<td>1.25e$-2$</td>
<td>1.17e$-2$</td>
<td>1.06e$-2$</td>
<td><strong>9.90e$-3$</strong></td>
<td>1.12e$-2$</td>
<td>3.93e$-1$</td>
</tr>
</tbody>
</table>

optimization methods cannot recover the signal, but the proposed non-convex method can also recover the signal exactly with high probability. Therefore, from Figure 2, it can be seen that the proposed non-convex method is more accurate and effective than other three methods. Especially, the capacity of the mixed $\ell_2/\ell_p$-norm optimization is improved with the decreasing $p$. Secondly, the case of noisy measurements is studied. The dimension of the simulation signal is $N = 1024$, and the block number is $n = 32$. The relative error is used to evaluate the performance of the test algorithms from $M = 512$ measurements. The noise is assumed as i.i.d. Gaussian with zero mean and standard deviation $\sigma$. The experimental results about
Figure 3. The simulation results of the large scale problem: (a) the original signal and the reconstructed signal by $\ell_{0,5}$, (b) the original signal and the reconstructed signal by BPDN, (c) the original signal and the reconstructed signal by $\ell_2/\ell_1$ and (d) the original signal and the reconstructed signal by $\ell_2/\ell_{0.5}$. 
three block-sparsity levels and four noise levels are listed in Table 2. The best results are labeled in bold. The experimental results indicate that the mixed $\ell_2/\ell_p$-norm optimization provides lower relative error than the $\ell_p(p = 0.5)$, BPDN and the mixed $\ell_2/\ell_1$-norm methods in the presence of noise. Especially, for the less block sparse $K = 5$ and $K = 7$, when the noise level is small, the performance of mixed $\ell_2/\ell_p$-norm optimization is more efficient with a little smaller $p$. However, if the noise level is large, the reconstruction capacity decreases with small $p$. For block-sparsity level $K = 9$, when $p$ is around 0.5, the mixed $\ell_2/\ell_p$-norm optimization have better reconstruction performance. Especially, comparing with the convex method (the mixed $\ell_2/\ell_1$-norm optimization), the proposed method generates smaller reconstruction error. These experimental results demonstrate that the proposed method is more robust than the mixed $\ell_2/\ell_1$-norm optimization in the case of noisy measurements.

Lastly, the large-scale problem is studied. The length of the block-sparse signal is set to $N = 4096$, the number of blocks is 64, and the block-sparsity level is fixed as $K = 8$. Figure 3 illustrates the original true signal and the reconstructed signal of different algorithms from $M = 960$ measurements with the noise level $\sigma = 0.05$. The black diamond is the original true signal and the blue cross is the reconstructed signal. Figure 3(a)–(d) shows the results of $\ell_p(p = 0.5)$, BPDN, $\ell_2/\ell_1$ and $\ell_2/\ell_p(p = 0.5)$, respectively. The relative errors are listed simultaneously. Obviously, the mixed $\ell_2/\ell_p(p = 0.5)$-norm optimization generates the lowest relative error, which shows the effectiveness of proposed non-convex method for large scale block-sparse signal recovery.

5. Conclusion

In this article, we propose a non-convex model for block-sparse signals recovery termed as the mixed $\ell_2/\ell_p(0 < p < 1)$-norm optimization. Based on the B-RIP, we prove that this non-convex method can recover the block-sparse signals effectively in noise-free and noise cases. The re-weighted $\ell_2/\ell_1$-algorithm is developed to solve the proposed non-convex optimization problem. The experimental results show that the performance of the mixed $\ell_2/\ell_p$-norm optimization surpasses the mixed $\ell_2/\ell_1$-norm optimization and some reconstruction algorithms without considering the block structure. Our future works will focus on the applications of block-sparse compressed sensing including the DNA microarrays, multiband signals and MEG.

Acknowledgements

This article is supported by the National Natural Science Foundation of China (no. 61172161), the Fundamental Research Funds for the Central Universities, Hunan University, the Scholarship Award for Excellent Doctoral Student granted by Chinese Ministry of Education, the Independent Research Funds of the State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body (Hunan University, no. 71165002) and the Hunan Provincial Innovation Foundation For Postgraduate (no. CX2011B131).

References