View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Crossref

Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2014, Article ID 104709, 8 pages
http://dx.doi.org/10.1155/2014/104709

Research Article

A Simple Gaussian Measurement Bound for Exact Recovery of Block-Sparse Signals

Zhi Han,[1] Jianjun Wang,[2] Jia Jing,[2] and Hai Zhang[3]

[1] *State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Science, Shenyang 110016, China*
[2] *School of Mathematics and Statistics, Southwest University, Chongqing 400715, China*
[3] *Department of Mathematics, Northwest University, Xi'an 710069, China*

Correspondence should be addressed to Hai Zhang; zhanghai0306@gmail.com

Received 9 May 2014; Revised 10 September 2014; Accepted 25 September 2014; Published 26 November 2014

Academic Editor: Manuel De la Sen

Copyright © 2014 Zhi Han et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a probabilistic analysis on conditions of the exact recovery of block-sparse signals whose nonzero elements appear in fixed blocks. We mainly derive a simple lower bound on the necessary number of Gaussian measurements for exact recovery of such block-sparse signals via the mixed l_2/l_q $(0 < q \leq 1)$ norm minimization method. In addition, we present numerical examples to partially support the correctness of the theoretical results. The obtained results extend those known for the standard l_q minimization and the mixed l_2/l_1 minimization methods to the mixed l_2/l_q $(0 < q \leq 1)$ minimization method in the context of block-sparse signal recovery.

1. Introduction and Main Results

The problem of block-sparse signal recovery naturally arises in a number of genetics, image processing, and machine learning tasks. Prominent examples include DNA microarrays [1], wavelet sparsity modeling [2], color imaging [3], and wideband spectrum sensing [4]. In these contexts, we often require to recover an unknown signal $\mathbf{x} \in \mathbb{R}^N$ from an underdetermined system of linear equations $\mathbf{y} = \Phi\mathbf{x}$, where $\mathbf{y} \in \mathbb{R}^M$ are available measurements and Φ is a $M \times N$ $(M < N)$ measurement matrix. Unlike previous works in compressed sensing (CS) [5–7], the unknown signal $\mathbf{x}$ not only is sparse but also exhibits additional structure in the form where the nonzero coefficients appear in some fixed blocks. We refer to such a structured sparse vector as block-sparse signal in this paper. Following [8–11], we only consider the nonoverlapping case in the present study. Thus, from mathematical point, a block signal $\mathbf{x} \in \mathbb{R}^N$ can be viewed as concatenation of $\mathbf{x}$ in m blocks of length d; that is,

$$\mathbf{x} = \Big[\underbrace{x_1 \cdots x_d}_{\mathbf{x}[1]} \underbrace{x_{d+1} \cdots x_{2d}}_{\mathbf{x}[2]} \cdots \underbrace{x_{N-d+1} \cdots x_N}_{\mathbf{x}[m]}\Big]^T, \quad (1)$$

where $\mathbf{x}[i]$ denotes the ith block of $\mathbf{x}$ and $N = md$ $(m, d \in Z^+)$. In these terms, we say that $\mathbf{x}$ is block k-sparse if $\mathbf{x}[i]$ has nonzero Euclidean norm for at most k blocks. In this paper, we furthermore assume that each element in these k nonzero blocks has nonzero coefficient. Obviously, if $d = 1$, the block-sparse signal degenerates to the conventional sparse signal well studied in compressed sensing.

Denote

$$\|\mathbf{x}\|_{2,0} = \sum_{i=1}^{m} I\left(\|\mathbf{x}[i]\|_2 > 0\right), \quad (2)$$

where $I(\|\mathbf{x}[i]\|_2 > 0)$ is an indicator function; that is, $I(\|\mathbf{x}[i]\|_2 > 0) = 1$, if $\|\mathbf{x}[i]\|_2 > 0$; $I(\|\mathbf{x}[i]\|_2 > 0) = 0$, otherwise. Thus a block k-sparse signal $\mathbf{x}$ can be defined as a signal that satisfies $\|\mathbf{x}\|_{2,0} \leq k$. It is known that, under certain conditions on measurement matrix Φ (i.e., [8]), there is a unique block-sparse signal that obeys the observation $\mathbf{y} = \Phi\mathbf{x}$ and can be exactly recovered by solving the problem

$$\min_{\mathbf{x}} \ \|\mathbf{x}\|_{2,0} \quad \text{s.t.} \quad \mathbf{y} = \Phi\mathbf{x}. \quad (3)$$

Similar to the standard l_0 minimization problem, (3) is NP-hard and computationally intractable except for very small size. Motivated by the study of CS, one then commonly uses the strategy to replace the l_2/l_0 norm with its closest convex surrogate l_2/l_1 norm, thus to solve a mixed l_2/l_1 norm minimization problem:

$$\begin{aligned} &\min_{\mathbf{x}} \ \|\mathbf{x}\|_{2,1} \\ &\text{s.t.} \quad \mathbf{y} = \Phi\mathbf{x}, \end{aligned} \tag{4}$$

where $\|\mathbf{x}\|_{2,1} = \sum_{i=1}^{m} \|\mathbf{x}[i]\|_2$. This model can be treated as a second-order cone program (SOCP) problem and many standard software packages can be used for the solutions very efficiently. In many practical cases, the measurements $\mathbf{y}$ are corrupted by bounded noise; then we can apply the modified SOCP or the group version of basis pursuit denoising (BPDN, [12]) program as the following:

$$\min_{\mathbf{x}} \left\|\mathbf{y} - \Phi\mathbf{x}\right\|_2^2 + \lambda \|\mathbf{x}\|_{2,1}, \tag{5}$$

where λ is a tuning parameter, which controls the tolerance of the noise term. There are also many methods to solve this optimization problem efficiently, such as the block-coordinate descent technique [13] and the Landweber iterations technique [14]. Conditions under which solving problem (4) can successfully recover a block-sparse signal $\mathbf{x}$ have been extensively studied [8, 10, 15]. For example, Eldar and Mishali [8] generalized the conventional RIP notion to the block-sparse setting and showed that if Φ satisfies the block-RIP (see Definition 1) with constant $\delta_{2k} < 0.414$, solving (4) can accurately get any block-sparse solution of $\mathbf{y} = \Phi\mathbf{x}$.

Among the latest researches in compressed sensing, many authors [16–21] have showed that l_q minimization with $0 < q < 1$ allows the exact recovery of conventional sparse signals from much fewer linear measurements than that by l_1 minimization. Naturally, it would be interesting to make an ongoing effort to extend l_q $(0 < q < 1)$ minimization to the setting of block-sparsity. Specifically, the following mixed l_2/l_q $(0 < q \le 1)$ norm minimization problem is proposed (see [11])

$$\begin{aligned} &\min_{\mathbf{x}} \ \|\mathbf{x}\|_{2,q}^q \\ &\text{s.t.} \quad \mathbf{y} = \Phi\mathbf{x}, \end{aligned} \tag{6}$$

for block-sparse signal recovery, as a generalization of the standard l_q minimization, where $\|\mathbf{x}\|_{2,q} = (\sum_{i=1}^{m} \|\mathbf{x}[i]\|_2^q)^{1/q}$. Similar to the standard l_q minimization problem, (6) is also a nonconvex problem for any $0 < q < 1$, and finding its global minimizer is in general computationally impossible. However, it is well known that there are several efficient heuristic methods to compute local minimizers of the standard l_q $(0 < q \le 1)$ minimization problem; say, for example, [18, 22]. One can generalize those approaches to solve the mixed l_2/l_q minimization problem (6). In particular, we will adopt the iteratively reweighted least squares techniques in this paper.

In [17], Chartrand and Staneva conducted a detailed analysis of the l_q $(0 < q \le 1)$ minimization approach for the nonblock sparse recovery problem. They derived a lower bound of Gaussian measurement for exact recovery of a nonblock sparse signal. Furthermore, Eldar and Mishali [8] also provided a lower bound on block-sparse signal recovery in the Gaussian measurement ensemble for the mixed l_2/l_1 minimization. Along this line, we will provide in this paper a lower bound of Gaussian measurements for exact recovery of block-sparse signal through the mixed l_2/l_q $(0 < q \le 1)$ norm minimization. The obtained results will complement the results of [8, 17] and demonstrate particularly that the block version of l_q $(0 < q \le 1)$ minimization can reduce the number of Gaussian measurements necessary for the exact recovery as q decreases.

To introduce our results, we first state the definition of block-RIP [8] as follows.

Definition 1 (see [8]). Let $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ be an $M \times N$ measurement matrix. One says that Φ has the block-RIP over $\mathcal{I} = \{d_1 = d, \ldots, d_m = d\}$ (then $N = md$) with constant $\delta_{k|\mathcal{I}}$ if for every block k-sparse signal $\mathbf{x} \in \mathbb{R}^N$ over $\mathcal{I}$ such that

$$\left(1 - \delta_{k|\mathcal{I}}\right) \|\mathbf{x}\|_2^2 \le \|\Phi\mathbf{x}\|_2^2 \le \left(1 + \delta_{k|\mathcal{I}}\right) \|\mathbf{x}\|_2^2. \tag{7}$$

This new definition of block-RIP is crucial for our analysis of the mixed l_2/l_q $(0 < q \le 1)$ minimization method. For convenience, we still use δ_k instead of $\delta_{k|\mathcal{I}}$, henceforth to represent the block-RIP constant of order k. With this notion, we can prove the following sufficient condition for exact recovery of a block k-sparse signal.

Theorem 2. *Let $\mathbf{x} \in \mathbb{R}^N$ be an arbitrary block k-sparse signal. For any given $q \in (0, 1]$, if the measurement matrix $\Phi \in \mathbb{R}^{M\times N}$ has the block-RIP (7) with constant*

$$\delta_{2k} < \frac{1}{2} - \frac{q}{4}\left(\frac{2}{3} - \frac{q}{3}\right)^{2/q-1}, \tag{8}$$

then the mixed l_2/l_q minimization problem (6) has a unique solution, given by the original signal $\mathbf{x}$.

Condition (8) generalizes the condition of the mixed l_2/l_1 case in [8] to the mixed l_2/l_q $(0 < q \le 1)$ case. Note that the right-hand side of (8) is monotonically decreasing with q, which shows that decreasing q can get weaker recovery condition. This shows further that fewer measurements might be needed to recover a block k-sparse signal whenever the mixed l_2/l_q $(0 < q < 1)$ minimization methods are used instead of the mixed l_2/l_1 minimization method. For clarifying this issue more precisely, we can adopt a similar probabilistic method of [17] to derive a simple lower bound on how many random Gaussian measurements are sufficient for (8) to hold with high probability. The result to be verified is in Section 3.

Theorem 3. *Let Φ be a $M \times N$ matrix with i.i.d. zero-mean unit variance Gaussian entries and $N = md$ for some integer m. For any given $q \in (0, 1]$, if*

$$M \ge C_1(q)\,kd + C_2(q)\,kd \log\frac{m}{k}, \tag{9}$$

then the subsequent conclusion is true with probability exceeding $1 - 1/\binom{m}{k}$: Φ *satisfies (8); therefore, any block k-sparse signal* $\mathbf{x} \in \mathbb{R}^N$ *can be recovered exactly by the solution of the mixed* l_2/l_q *minimization problem (6).*

Theorem 3 shows that a block k-sparse signal $\mathbf{x}$ can be recovered exactly with high probability by solving the mixed l_2/l_q minimization (6) provided the number of Gaussian measurements M satisfies

$$M \approx C_2(q)\, k \log \frac{m}{k}. \tag{10}$$

More detailed remarks on this bound will be presented in Section 3.

The rest of the paper is organized as follows. In Section 2, we present several key lemmas needed for the proofs of our main results. All these lemmas can be regarded as generalizations of the standard non-block-sparse case to the block-sparse case. The proofs of Theorems 2 and 3 are given in Section 3. Numerical experiments are provided in Section 4 to demonstrate the correctness of the theoretical results. We conclude the paper then in Section 5 with some useful remarks.

2. Fundamental Lemmas

In this section, we establish several lemmas necessary for the proofs of the main results.

Lemma 4 (see [8]). *Consider*

$$\left|\langle \Phi\mathbf{x}, \Phi\mathbf{x}'\rangle\right| \le \delta_{k+k'} \|\mathbf{x}\|_2 \left\|\mathbf{x}'\right\|_2, \tag{11}$$

for all $\mathbf{x}$, $\mathbf{x}'$ *supported on disjoint subsets* $T, T' \subseteq \{1,2,\ldots,N\}$ *with* $|T| < k$, $|T'| < k'$.

In the next lemma, we show that the probability that block restricted isometry constant δ_k exceeds a certain scope decays exponentially in certain length of $\mathbf{x}$.

Lemma 5 (see [8]). *Suppose* Φ *is an* $M \times N$ *matrix from the Gaussian ensemble; namely,* $\Phi_{ij} \sim N(0, 1/M)$. *Let* $\delta_{k|\mathcal{I}}$ *be the smallest value satisfying the block-RIP of* Φ *over* $I = \{d_1 = d, d_2 = d, \ldots, d_m = d\}$ *and* $N = md$ *for some integer m. Then, for every* $\epsilon > 0$, *the block restricted isometry constant* $\delta_{k|\mathcal{I}}$ *satisfies the following inequality:*

$$\sqrt{1+\delta_{k|\mathcal{I}}} > 1 + (1+\epsilon)\sqrt{\frac{N}{M}}\left(\sqrt{\frac{kd}{N}} + \sqrt{\frac{2}{d}H\left(\frac{kd}{N}\right)}\right) \tag{12}$$

which holds with high probability $2e^{-mH(kd/N)\epsilon}$, *where* $H(p) = -p\log(p) - (1-p)\log(1-p)$ $(0 < p < 1)$.

Lemma 6. *Suppose* Φ *is an* $M \times N$ *matrix from the Gaussian ensemble; namely,* $\Phi_{ij} \sim N(0, 1/M)$. *For any* $0 < \delta < 1$, *the following estimation*

$$(1-\delta)\|\mathbf{x}\|_2^2 \le \|\Phi\mathbf{x}\|_2^2 \le (1+\delta)\|\mathbf{x}\|_2^2 \tag{13}$$

holds uniformly for $\mathbf{x} \in \mathbb{R}^L$ *with probability exceeding* $1 - 2(12/\delta)^L e^{-M(1+\delta/2)mH(k/m)/k}$.

Proof. Note that [23] has verified the subsequent conclusion: if Φ is a random matrix of size $M \times N$ drawn according to a distribution that satisfies the concentration inequality $P(\left|\|\Phi\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2\right| \ge \beta\|\mathbf{x}\|_2^2) \le 2e^{-Mc_0(\beta)}$, where $c_0(\beta)$ is a constant dependent only on β such that $c_0(\beta) > 0$ for $0 < \beta < 1$, then, for any set T with $|T| = k < n$ and any $0 < \delta < 1$, there holds the estimation

$$(1-\delta)\|\mathbf{x}\|_2 \le \|\Phi\|_2 \le (1+\delta)\|\mathbf{x}\|_2 \quad (\mathbf{x} \in T) \tag{14}$$

with probability exceeding $1 - 2(12/\delta)^k e^{-Mc_0(\delta/2)}$. By Lemma 5, the concentration inequality is clearly true for Gaussian measurement matrix. Thus Lemma 6 follows. □

We also need the following inequality.

Lemma 7 (see [24]). *For any fixed* $q \in (0,1)$ *and* $\mathbf{x} \in \mathbb{R}^N$,

$$\|\mathbf{x}\|_2 \le \frac{\|\mathbf{x}\|_q}{N^{1/q-1/2}} + \sqrt{N}\left(\max_{1\le i\le N}|\mathbf{x}_i| - \min_{1\le i\le N}|\mathbf{x}_i|\right). \tag{15}$$

3. Proofs of Theorems

With the preparations made in the last section, we now prove the main results of the paper. Henceforth, we let Φ denote an $M \times N$ matrix whose elements are i.i.d. random variables; specifically, $\Phi_{ij} \sim N(0, 1/M)$. We prove Theorem 2 by using the block-RIP and Lemmas 4 and 7.

Proof of Theorem 2. First notice that assumption (8) implies $\delta_{2k} < 1$ and further implies the uniqueness of $\mathbf{x}$ satisfying the observation $\mathbf{y} = \Phi\mathbf{x}$. Let $\mathbf{x}^* = \mathbf{x} + \mathbf{h}$ be a solution of (6). Our goal is then to show that $\mathbf{h} = 0$. For this purpose, we let $\mathbf{x}_T$ be the signal that is identical to $\mathbf{x}$ on the index set T and to zero elsewhere. Let T_0 be the block index set over the k nonzero blocks of $\mathbf{x}$, and we decompose $\mathbf{h}$ into a series of vectors $\mathbf{h}_{T_0}, \mathbf{h}_{T_1}, \mathbf{h}_{T_2}, \ldots, \mathbf{h}_{T_J}$ such that

$$\mathbf{h} = \sum_{i}^{J} \mathbf{h}_{T_i}. \tag{16}$$

Here $\mathbf{h}_{T_i}$ is the restriction of $\mathbf{h}$ onto the set T_i and each T_i consists of k blocks (except possibly T_J). Rearrange the block indices such that $\left\|\mathbf{h}_{T_j}[1]\right\|_2 \ge \left\|\mathbf{h}_{T_j}[2]\right\|_2 \ge \cdots \ge \left\|\mathbf{h}_{T_j}[k]\right\|_2 \ge \left\|\mathbf{h}_{T_{j+1}}[1]\right\|_2 \ge \left\|\mathbf{h}_{T_{j+1}}[2]\right\|_2 \ge \cdots$, for any $j \ge 1$.

From [9], $\mathbf{x}^*$ is the unique sparse solution of (6) being equal to $\mathbf{x}$ if and only if

$$\left\|\mathbf{h}_{T_0}\right\|_{2,q} < \left\|\mathbf{h}_{T_0^c}\right\|_{2,q} \tag{17}$$

for all nonzero signal $\mathbf{h}$ in the null space of Φ. This is the so-called the null space property (NSP). In order to characterize the NSP more precisely, we consider the following equivalent form: there exists a constant $\gamma(\mathbf{h},q)$ satisfying $0<\gamma(\mathbf{h},q)<1$ such that

$$\left\|\mathbf{h}_{T_0}\right\|_{2,q}=\gamma(\mathbf{h},q)\left\|\mathbf{h}_{T_0^c}\right\|_{2,q}. \quad (18)$$

We proceed by showing that $\gamma(\mathbf{h},q)<1$ under the assumption of Theorem 2.

In effect, we observe that

$$\begin{aligned}\|\mathbf{h}\|_2&=\left\|\mathbf{h}_{T_0\bigcup T_1}+\mathbf{h}_{(T_0\bigcup T_1)^c}\right\|_2\\&\le\left\|\mathbf{h}_{T_0\bigcup T_1}\right\|_2+\left\|\mathbf{h}_{(T_0\bigcup T_1)^c}\right\|_2.\end{aligned} \quad (19)$$

For any $j\ge 2$, if we denote $\mathbf{d}_j=(\|\mathbf{h}_{T_j}[1]\|_2,\|\mathbf{h}_{T_j}[2]\|_2,\ldots,\|\mathbf{h}_{T_j}[k]\|_2)$, then we have

$$\begin{aligned}\left\|\mathbf{d}_j\right\|_2&=\left(\sum_{i=1}^k\left\|\mathbf{h}_{T_j}[i]\right\|_2^2\right)^{1/2}=\left\|\mathbf{h}_{T_j}\right\|_2,\\\left\|\mathbf{d}_j\right\|_q&=\left(\sum_{i=1}^k\left\|\mathbf{h}_{T_j}[i]\right\|_2^q\right)^{1/q}=\left\|\mathbf{h}_{T_j}\right\|_{2,q},\\\max_{1\le i\le k}\left|\mathbf{d}_j(i)\right|&=\max_{1\le i\le k}\left\|\mathbf{h}_{T_j}[i]\right\|_2=\left\|\mathbf{h}_{T_j}\right\|_{\infty,T},\\\min_{1\le i\le k}\left|\mathbf{d}_j(i)\right|&=\min_{1\le i\le k}\left\|\mathbf{h}_{T_j}[i]\right\|_2.\end{aligned} \quad (20)$$

By Lemma 7, we thus have

$$\sum_{j\ge 2}\left\|\mathbf{d}_j\right\|_2\le\frac{q^{1/2}(2-q)^{1/q-1/2}}{2^{1/q-1}k^{1/q-1/2}}\sum_{j\ge 1}\left\|\mathbf{d}_j\right\|_q; \quad (21)$$

that is,

$$\begin{aligned}\sum_{j\ge 2}\left\|\mathbf{h}_{T_j}\right\|_2&\le\frac{q^{1/2}(2-q)^{1/q-1/2}}{2^{1/q-1}k^{1/q-1/2}}\sum_{j\ge 1}\left\|\mathbf{h}_{T_j}\right\|_{2,q}\\&\le\frac{q^{1/2}(2-q)^{1/q-1/2}}{2^{1/q-1}k^{1/q-1/2}}\left(\sum_{j\ge 1}\left\|\mathbf{h}_{T_j}\right\|_{2,q}^q\right)^{1/q}.\end{aligned} \quad (22)$$

On the other hand, let

$$\begin{aligned}\left\|\mathbf{h}_{T_0}\right\|_{2,q}&=\gamma(\mathbf{h},q)\left\|\mathbf{h}_{T_0^c}\right\|_{2,q},\\\left\|\mathbf{h}_{T_1}\right\|_{2,q}^q&=t\sum_{i\ge 1}\left\|\mathbf{h}_{T_i}\right\|_{2,q}^q.\end{aligned} \quad (23)$$

It is easy to see that

$$\begin{aligned}&\sum_{i\ge 2}\left\|\mathbf{h}_{T_i}\right\|_2^2\\&\le\left\|\mathbf{h}_{T_2}[1]\right\|_{2,q}^{2-q}\sum_{j\ge 2}\left\|\mathbf{h}_{T_j}\right\|_{2,q}^q\\&\le\left(\frac{\left(\left\|\mathbf{h}_{T_1}[1]\right\|_2^q+\left\|\mathbf{h}_{T_1}[2]\right\|_2^q+\cdots+\left\|\mathbf{h}_{T_1}[k]\right\|_2^q\right)}{k}\right)^{(2-q)/q}\\&\quad\times\sum_{j\ge 2}\left\|\mathbf{h}_{T_j}\right\|_{2,q}^q\\&=\left(\frac{\left\|\mathbf{h}_{T_1}\right\|_{2,q}^q}{k}\right)^{(2-q)/q}\sum_{j\ge 2}\left\|\mathbf{h}_{T_j}\right\|_{2,q}^q\\&=\left(\frac{\left\|\mathbf{h}_{T_1}\right\|_{2,q}^q}{k}\right)^{(2-q)/q}\left(\sum_{j\ge 1}\left\|\mathbf{h}_{T_j}\right\|_{2,q}^q-\left\|\mathbf{h}_{T_1}\right\|_{2,q}^q\right)\\&=\left(\frac{\left\|\mathbf{h}_{T_1}\right\|_{2,q}^q}{k}\right)^{(2-q)/q}\left(\frac{1}{t}\left\|\mathbf{h}_{T_1}\right\|_{2,q}^q-\left\|\mathbf{h}_{T_1}\right\|_{2,q}^q\right)\\&=\frac{1-t}{tk^{(2-q)/q}}\left\|\mathbf{h}_{T_1}\right\|_{2,q}^2\\&=\frac{1-t}{t^{1-2/q}k^{(2-q)/q}}\left(\sum_{j\ge 1}\left\|\mathbf{h}_{T_j}\right\|_{2,q}^q\right)^{2/q}.\end{aligned} \quad (24)$$

Since $\Phi\mathbf{x}=\Phi\mathbf{x}^*$, we have $\Phi\mathbf{h}=0$, which means $\Phi(\mathbf{h}_{T_0}+\mathbf{h}_{T_1})=-\Phi(\sum_{j\ge 2}\mathbf{h}_{T_j})$. By the definition of block-RIP, Lemma 4, (22), and (24), it then follows that

$$\begin{aligned}&\left\|\Phi\left(\sum_{j\ge 2}\mathbf{h}_{T_j}\right)\right\|_2^2\\&=\sum_{i,j\ge 2}\left\langle\Phi\left(\mathbf{h}_{T_i}\right),\Phi\left(\mathbf{h}_{T_j}\right)\right\rangle\\&=\sum_{j\ge 2}\left\langle\Phi\left(\mathbf{h}_{T_j}\right),\Phi\left(\mathbf{h}_{T_j}\right)\right\rangle\\&\quad+2\sum_{i,j\ge 2,i<j}\left\langle\Phi\left(\mathbf{h}_{T_i}\right),\Phi\left(\mathbf{h}_{T_j}\right)\right\rangle\\&\le(1+\delta_k)\sum_{j\ge 2}\left\|\mathbf{h}_{T_j}\right\|_2^2+2\delta_{2k}\sum_{i>j\ge 2}\left\|\mathbf{h}_{T_i}\right\|_2\left\|\mathbf{h}_{T_j}\right\|_2\\&\le\sum_{j\ge 2}\left\|\mathbf{h}_{T_j}\right\|_2^2+\delta_{2k}\left(\sum_{i\ge 2}\left\|\mathbf{h}_{T_i}\right\|_2\right)^2\\&\le\left(\frac{1-t}{t^{1-2/q}k^{(2-q)/q}}+\delta_{2k}\frac{q(2-q)^{2/q-1}}{2^{2/q-2}k^{2/q-1}}\right)\left(\sum_{j\ge 1}\left\|\mathbf{h}_{T_j}\right\|_{2,q}^q\right)^{2/q}.\end{aligned} \quad (25)$$

By the definition of block-RIP δ_{2k} and by using Hölder's equality, we then get

$$\begin{aligned}
&\left\|\Phi\left(\mathbf{h}_{T_0}+\mathbf{h}_{T_1}\right)\right\|_2^2 \\
&\quad \geq \left(1-\delta_{2k}\right)\left\|\mathbf{h}_{T_0}+\mathbf{h}_{T_1}\right\|_2^2 \\
&\quad = \left(1-\delta_{2k}\right)\left(\left\|\mathbf{h}_{T_0}\right\|_2^2+\left\|\mathbf{h}_{T_1}\right\|_2^2\right) \\
&\quad \geq \left(1-\delta_{2k}\right)\left(k^{1-2/q}\left\|\mathbf{h}_{T_0}\right\|_{2,q}^2+k^{1-2/q}\left\|\mathbf{h}_{T_1}\right\|_{2,q}^2\right) \\
&\quad = \left(1-\delta_{2k}\right)k^{1-2/q} \\
&\qquad \times\left(\gamma\left(\mathbf{h},q\right)^2\left\|\mathbf{h}_{T_0^c}\right\|_{2,q}^{2/q}+t^{2/q}\left(\sum_{i\geq 1}\left\|\mathbf{h}_{T_i}\right\|_{2,q}^q\right)^{2/q}\right) \\
&\quad = \left(1-\delta_{2k}\right)k^{1-2/q}\left(\gamma\left(\mathbf{h},q\right)^2+t^{2/q}\right)\left(\sum_{i\geq 1}\left\|\mathbf{h}_{T_i}\right\|_{2,q}^q\right)^{2/q}.
\end{aligned} \tag{26}$$

This together with (24) and (25) implies

$$\begin{aligned}
&\gamma\left(\mathbf{h},q\right)^2 \\
&\quad \leq \frac{\left(\left(\left(1-t\right)/t^{1-2/q}\right)+\left(q\left(2-q\right)^{2/q-1}/2^{2/q-2}\right)\delta_{2k}\right)}{\left(1-\delta_{2k}\right)} \\
&\qquad - t^{2/q} \doteq f\left(t\right).
\end{aligned} \tag{27}$$

Through a straightforward calculation, one can check that the maximum of $f(t)$ occurs at $t_0=(1-q/2)/(2-\delta_{2k})$ and

$$f\left(t_0\right)=\frac{\left(1-q/2\right)^{2/q-1}}{1-\delta_{2k}}\left\{\frac{q}{2\left(2-\delta_{2k}\right)^{2/q-1}}+2q\delta_{2k}\right\}. \tag{28}$$

If $f(t_0)<1$, then we clearly have $\gamma(\mathbf{h},q)<1$ and finish the proof. However, $f(t_0)<1$ gives that

$$\delta_{2k}\frac{2-q}{2-\delta_{2k}}+q\left(\frac{1-q/2}{2-\delta_{2k}}\right)^{2/q}<\frac{2-q}{2-\delta_{2k}}\left(1-\delta_{2k}\right) \tag{29}$$

or, equivalently,

$$\left(2\delta_{2k}-1\right)\left(2-\delta_{2k}\right)^{2/q-1}+\left(\frac{1}{2}\right)^{2/q}q\left(2-q\right)^{2/q-1}<0. \tag{30}$$

Note that (30) implies $0<\delta_{2k}<1/2$ and

$$\begin{aligned}
&\left(2\delta_{2k}-1\right)\left(2-\delta_{2k}\right)^{2/q-1}+\left(\frac{1}{2}\right)^{2/q}q\left(2-q\right)^{2/q-1} \\
&\quad <\left(2\delta_{2k}-1\right)\left(\frac{3}{2}\right)^{2/q-1}+\left(\frac{1}{2}\right)^{2/q}q\left(2-q\right)^{2/q-1}.
\end{aligned} \tag{31}$$

By simple calculations, we can check that, for any given $q\in(0,1]$, inequality (30) is true as long as

$$\delta_{2k}<\frac{1}{2}-\frac{q}{4}\left(\frac{2}{3}-\frac{q}{3}\right)^{2/q-1}; \tag{32}$$

that is, condition (8) is satisfied. With this, the proof of Theorem 2 is completed. □

Now we adopt the probabilistic methods to analyze how condition (8) can be satisfied; particularly, how many measurements are needed for exact recovery of a block k-sparse signal.

Proof of Theorem 3. From Theorem 2, we only need to show that, under condition (9), (8) holds with high probability. In particular, we will proceed to determine how random Gaussian measurements are sufficient for exact recovery of block k-sparse signals with a failure probability at most $1/\binom{m}{k}$. To this end, we let $L=2kd$. Then by Lemma 6, an upper bound for the probability that any $M\times L$ submatrix of Φ fails to satisfy the block-RIP (30) is

$$\binom{m}{2k}2\left(\frac{12}{\delta_{2k}}\right)^L e^{-M(1+\delta_{2k}/2)mH(k/m)/k}. \tag{33}$$

By inequality $1/\binom{m}{k}\geq k^k/m^ke^k$, it suffices to show that we can have the following estimation:

$$\left(\frac{12}{\delta_{2k}}\right)^L e^{-M(1+\delta_{2k}/2)mH(k/m)/k}\leq\frac{k^k}{2m^ke^k}\frac{(2k)^{2k}}{m^{2k}e^{2k}}. \tag{34}$$

So by a simple calculation,

$$\begin{aligned}
M &\geq \frac{1}{\left(1+\delta_{2k}/2\right)\left(m/k\right)H\left(k/m\right)} \\
&\quad \times\left\{L\log\left(\frac{12}{\delta_{2k}}\right)+3k\log\left(\frac{m}{k}\right)+\left(3-2\ln 2\right)k+\ln 2\right\} \\
&\geq C_1\left(q\right)kd+C_2\left(q\right)k\log\frac{m}{k},
\end{aligned} \tag{35}$$

where $C_1(q)=(3-\log 2+d\log(12/(1/2-(q/4)(2/3-q/3)^{2/q-1})))/(1+1/4-(q/8)(2/3-q/3)^{2/q-1})(m/k)H(k/m)$, $C_2(q)=3/(1+1/4-(q/8)(2/3-q/3)^{2/q-1})(m/k)H(k/m)$. That is, (35) is satisfied as long as (9) is met. This justifies Theorem 3. □

Remark 8. From the right-hand side of (35), we can further see that, except the first term related to m/k, other terms in the numerator are only related to k and d, that is, to the numbers of nonzero blocks and the block size. Obviously, those terms have a smaller contribution to the number of measurements. Basically, Theorem 3 shows that the signal $\mathbf{x}$ with block-sparsity $\|\mathbf{x}\|_{2,0}\leq k$ can be recovered exactly by the solution of the mixed l_2/l_q minimization (6) with high probability provided roughly

$$M\approx C_2\left(q\right)k\log\frac{m}{k}. \tag{36}$$

Obviously, this result generalizes the well-known result in [8] on the Gaussian ensemble for the mixed l_2/l_1 minimization. We can further see from Theorem 3 that, when specified to $q=1$, the bound (36) has the same order as Eldar's result ($O(k\log(m/k))$, see [8]). It is also obvious that $C(q)$ is monotonically increasing with respect to $q \in (0,1]$, which then implies that decreasing q allows fewer necessary measurements for exact recovery of block-sparse signals by the mixed l_2/l_q minimization method. Since, when block size $d = 1$, the mixed l_2/l_q minimization method degenerates to the standard l_q minimization method, the presented result then provides a theoretical support to such experimental observation reported in the following section.

4. Numerical Experiments

In this section, we conduct two numerical experiments to support the correctness of the obtained theoretical results. Iteratively reweighted least squares (IRLS) method has been proved to be very efficient for the standard l_q minimization. Thus, in the experiments, we adopted IRLS method to solve the following unconstrained smoothed version of (6):

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{2,q}^{\epsilon,q} + \frac{1}{2\tau}\|\mathbf{y}-\Phi\mathbf{x}\|_2^2, \tag{37}$$

where τ is a regularization parameter and $\|\mathbf{x}\|_{2,q}^{\epsilon,q} = \sum_{i=1}^{m}(\|\mathbf{x}[i]\|_2^2+\epsilon^2)^{q/2}$. Note that, through defining a diagonal weighting matrix W as $W_i = \text{diag}(q^{1/2}(\epsilon^2+\|\mathbf{x}[i]\|_2^2)^{q/4-1/2})$, (37) can be transformed to the following weighted least squares minimization problem:

$$\frac{1}{2\tau}\|\mathbf{y}-\Phi\mathbf{x}\|_2^2+\frac{1}{2}\|W\mathbf{x}\|_2^2. \tag{38}$$

With this, the IRLS algorithm we used can be summarized as follows (more details can be found in [11]).

Step 1. Set the iteration count t to zero and $\epsilon_0 = 1$. Initialize $\mathbf{x}^{(0)} = \min\left\|\mathbf{y}-\Phi\mathbf{x}\right\|_2^2$.

Step 2. Let

$$W^{(t)} = \text{diag}\left(q^{1/2}\left(\epsilon_t^2+\left\|\mathbf{x}^{(t)}[i]\right\|_2^2\right)^{1/2-q/4}\right), \quad i=1,\dots,m, \tag{39}$$

and update

$$\begin{aligned}\mathbf{x}^{(t+1)} &= \left(W^{(t)}\right)^{-1}\left(\Phi\left(W^{(t)}\right)^{-1}\right)^{T} \\ &\quad\times\left(\Phi\left(W^{(t)}\right)^{-1}+\tau\mathbf{I}\right)^{-1}\left(\Phi\left(W^{(t)}\right)^{-1}\right)^{T}\mathbf{y}, \\ \epsilon_{t+1} &= \frac{\epsilon_t}{10}.\end{aligned} \tag{40}$$

Step 3. Terminate the iteration on convergence or when t attains a specified maximum number $t_{\max}$.

Otherwise, set $t = t+1$ and go to Step 2.

The above algorithm can be seen as a natural generalization of general sparse signal recovery algorithm to the block-sparse setting, which alternates between estimating $\mathbf{x}$ and redefining the weighting matrix W. Though this algorithm is for the unconstrained penalized problem (37), it can still give a good estimation on the minimum of the constrained problem (6) when we choose a sufficient small τ (such as $\tau = 10^{-6}$). In the following experiments, we set $\tau = 10^{-6}$ and terminate the algorithm if $\epsilon < 10^{-7}$ or $\left\|x^{(t+1)}-x^{(t)}\right\|_2 < 10^{-8}$.

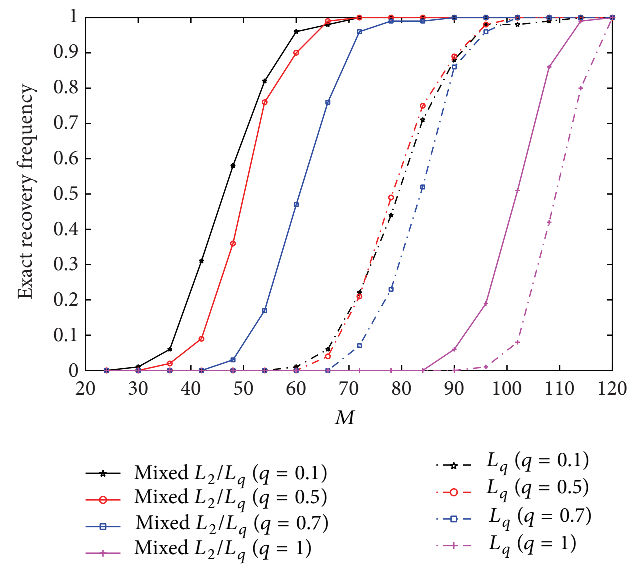


FIGURE 1: Plots of the exact recovery frequency versus the number of measurements, for the use of the mixed l_2/l_q and the standard l_q minimization methods with different values of q.

In our first experiment, we took the signal length $N = 128$, the block size $d = 4$, and the block-sparsity $k = 6$. For independent 100 trails, we first randomly generated the block-sparse signal $\mathbf{x}$ with values from a Gaussian distribution of mean 0 and standard deviation 1, and then we randomly drew a measurement matrix Φ from Gaussian ensemble. We also took the number of measurements varying from 24 to 124. The purpose of the experiment was then to check the correctness of Theorem 3. We considered four different values of $q = 0.1, 0.5, 0.7, 1$ for both the mixed l_2/l_q minimization method and the standard l_q minimization method.

The experiment result is shown in Figure 1. It is seen from Figure 1 that, for all the experiment runs, the smaller q requires the smaller number of measurements for exact recovery of block sparse signals. This observation is consistent with Theorem 3. On the other hand, for a fixed q, the mixed l_2/l_q method is clearly superior to the standard l_q method in this block-sparse setting. For instance, the mixed l_2/l_q method with $q = 0.1$ only uses 50 Gaussian measurements for the exact recovery, while the number of measurements needed for l_q method is around 90. This observation also partially supports the theoretical assertion of Theorem 3, namely, that incorporating the block structure into recovery method requires fewer measurements for exact recovery.

In our second experiment, we further studied the effect of block size d for block-sparse signal recovery. We set $N = 256$ and drew a measurement matrix of size 128×256

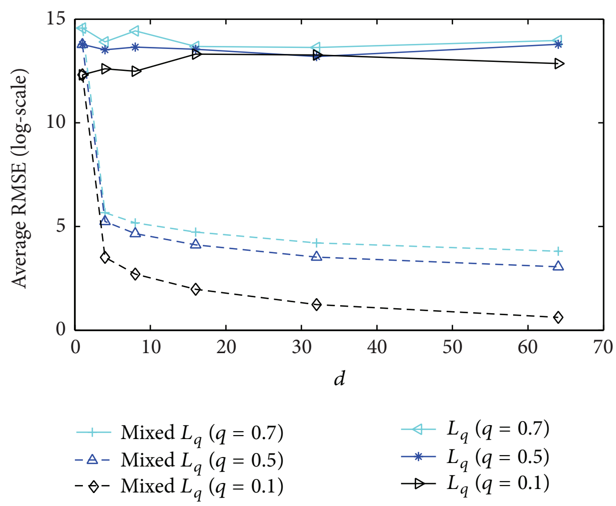


Figure 2: Plots of average RMSE (log-scale) versus block size, for the use of the mixed l_2/l_q and the standard l_q minimization methods with different values of q.

from Gaussian ensemble. In this experiment, the block size d was changed while keeping the total sparsity $kd = 64$ fixed. Figure 2 shows the average root mean squares error (RMSE). Here RMSR is defined as $\sqrt{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2/N}$ in the logarithmic scale over 100 independent random runs. One can easily see from Figure 2 that the recovery performance for the standard l_q method is independent of the active block number k, while the recovery errors for the mixed l_2/l_q method are significantly better than the standard l_q method when the active block number k is far smaller than the total signal sparsity kd. Since the total sparsity kd is fixed and larger d leads to smaller k, as predicted by Theorem 3, the necessary measurement number needed for exact recovery of a block k-sparse signal by the mixed method is also reduced. This also demonstrates the advantage of the mixed l_2/l_q method over the standard l_q method.

5. Conclusion

In this paper, the number of Gaussian measurements necessary for the exact recovery of a block-sparse signal by the mixed l_2/l_q $(0 < q \leq 1)$ norm minimization has been studied. The main contribution is the derivation of a lower bound on the necessary number of Gaussian measurements for the exact recovery based on a probabilistic analysis with block-RIP. The obtained results are helpful for understanding the recovery capability and algorithm development of the mixed l_2/l_q norm minimization approach for sparse recovery and, particularly, facilitate the applications in the development of block-sparse information processing.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work has been supported by the Natural Science Foundation of China under Grant nos. 61303168, 61273020, and 11171272.

References

[1] F. Parvaresh, H. Vikalo, S. Misra, and B. Hassibi, "Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays," *IEEE Journal on Selected Topics in Signal Processing*, vol. 2, no. 3, pp. 275–285, 2008.

[2] N. S. Rao, R. D. Nowak, S. J. Wright, and N. G. Kingsbury, "Convex approaches to model wavelet sparsity patterns," in *Proceedings of the 18th IEEE International Conference on Image Processing (ICIP '11)*, pp. 1917–1920, September 2011.

[3] A. Majumdar and R. K. Ward, "Compressed sensing of color images," *Signal Processing*, vol. 90, no. 12, pp. 3122–3127, 2010.

[4] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: compressed sensing for analog signals," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 993–1009, 2009.

[5] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[6] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[7] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[8] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.

[9] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3075–3085, 2009.

[10] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[11] Y. Wang, J. Wang, and Z. Xu, "On recovery of block-sparse signals via mixed $l_2 = lq(0 < q \leq 1)$ norm minimization," *Eurasip Journal on Advances in Signal Processing*, vol. 2013, no. 1, article 76, 2013.

[12] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[13] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society B*, vol. 68, no. 1, pp. 49–67, 2006.

[14] M. Fornasier and H. Rauhut, "Recovery algorithms for vector-valued data with joint sparsity constraints," *SIAM Journal on Numerical Analysis*, vol. 46, no. 2, pp. 577–613, 2008.

[15] J. Huang and T. Zhang, "The benefit of group sparsity," *The Annals of Statistics*, vol. 38, no. 4, pp. 1978–2004, 2010.

[16] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707–710, 2007.

[17] R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," *Inverse Problems*, vol. 24, no. 3, Article ID 035020, 14 pages, 2008.

[18] S. Foucart and M. J. Lai, "Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 395–407, 2009.

[19] Z. Xu, H. Zhang, Y. Wang, X. Chang, and Y. Liang, "$L_{1/2}$ regularization," *Science China. Information Sciences*, vol. 53, no. 6, pp. 1159–1169, 2010.

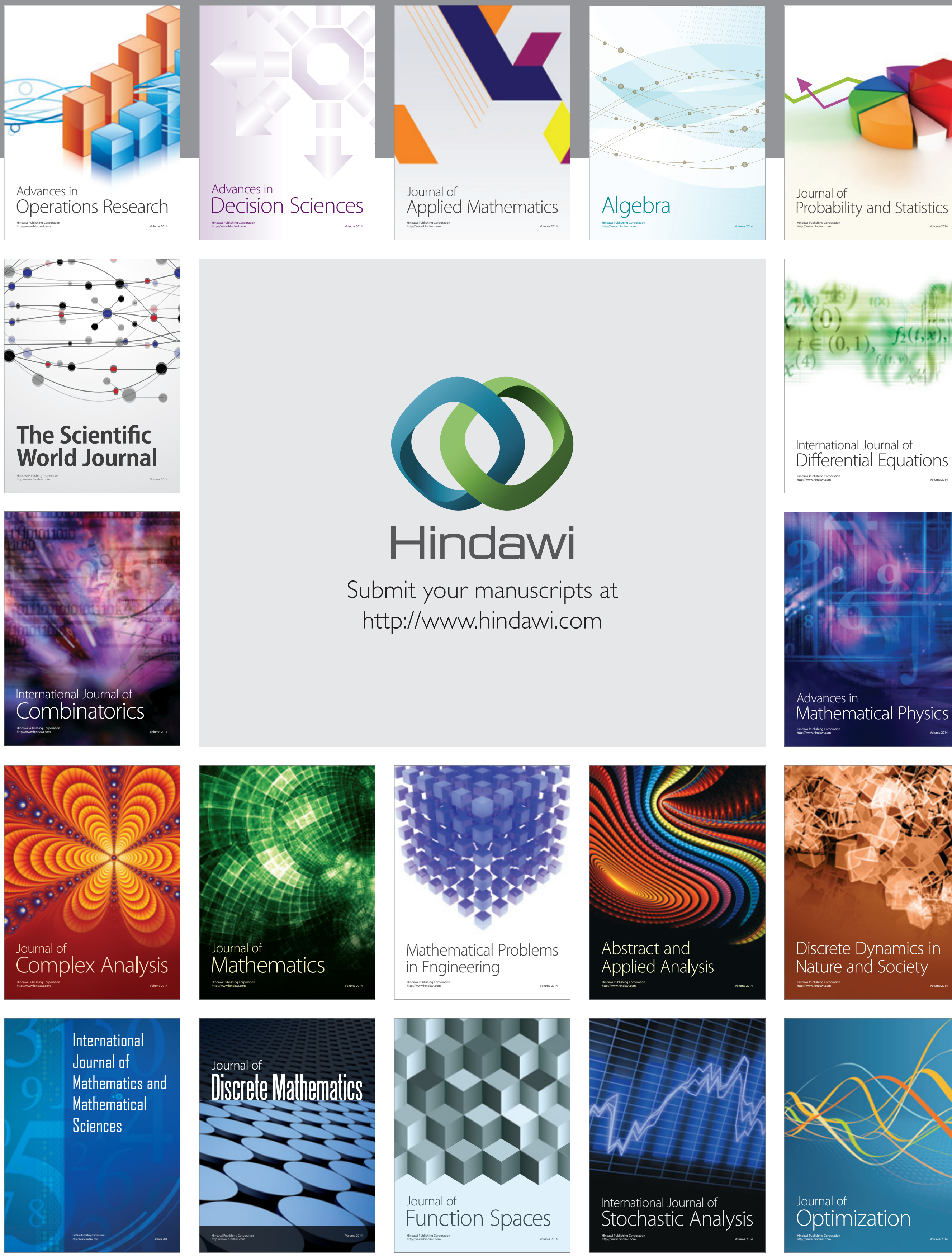
[20] Z. Xu, X. Chang, F. Xu, and H. Zhang, "$L_{1/2}$ regularization: a thresholding representation theory and a fast solver," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 7, pp. 1013–1027, 2012.

[21] Q. Sun, "Recovery of sparsest signals via ℓq-minimization," *Applied and Computational Harmonic Analysis*, vol. 32, no. 3, pp. 329–341, 2012.

[22] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Communications on Pure and Applied Mathematics*, vol. 63, no. 1, pp. 1–38, 2010.

[23] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[24] M. J. Lai and L. Y. Liu, "A new estimate of restricted isometry constants for sparse solutions," *Applied and Computational Harmonic Analysis*, vol. 30, pp. 402–406, 2011.


Advances in Operations Research
Advances in Decision Sciences
Journal of Applied Mathematics
Algebra
Journal of Probability and Statistics
The Scientific World Journal
Hindawi
Submit your manuscripts at
http://www.hindawi.com
International Journal of Differential Equations
International Journal of Combinatorics
Advances in Mathematical Physics
Journal of Complex Analysis
Journal of Mathematics
Mathematical Problems in Engineering
Abstract and Applied Analysis
Discrete Dynamics in Nature and Society
International Journal of Mathematics and Mathematical Sciences
Journal of Discrete Mathematics
Journal of Function Spaces
International Journal of Stochastic Analysis
Journal of Optimization