Bull. Aust. Math. Soc. **92** (2015), 496–507
doi:10.1017/S0004972715000933

THE RESTRICTED ISOMETRY PROPERTY FOR SIGNAL RECOVERY WITH COHERENT TIGHT FRAMES

FEN-GONG WU✉ and DONG-HUI LI

(Received 12 March 2015; accepted 21 March 2015; first published online 19 August 2015)

Abstract

In this paper, we consider signal recovery via l_1-analysis optimisation. The signals we consider are not sparse in an orthonormal basis or incoherent dictionary, but sparse or nearly sparse in terms of some tight frame D. The analysis in this paper is based on the restricted isometry property adapted to a tight frame D (abbreviated as D-RIP), which is a natural extension of the standard restricted isometry property. Assuming that the measurement matrix $A \in \mathbb{R}^{m\times n}$ satisfies D-RIP with constant δ_{tk} for integer k and $t > 1$, we show that the condition $\delta_{tk} < \sqrt{(t-1)/t}$ guarantees stable recovery of signals through l_1-analysis. This condition is sharp in the sense explained in the paper. The results improve those of Li and Lin [‘Compressed sensing with coherent tight frames via l_q-minimization for $0 < q \leq 1$’, Preprint, 2011, arXiv:1105.3299] and Baker [‘A note on sparsification by frames’, Preprint, 2013, arXiv:1308.5249].

2010 *Mathematics subject classification*: primary 94A12.

Keywords and phrases: compressed sensing, sparse recovery, l_1-analysis, D-restricted isometry property, tight frames.

1. Introduction

Compressed sensing is a new branch of signal recovery, distinct from sparse and redundant representations. By exploiting sparse representation of signals, their sampling can be made far more effective compared to the classical Nyquist–Shannon sampling (see, for example, [4, 12]).

In compressed sensing, the signal $f \in \mathbb{R}^n$ is acquired by collecting m linear measurements of the form $y_k = \langle a_k, f\rangle + z_k$, $1 \leq k \leq m$, or in matrix notation,

$$y = Af + z,$$

where A is a known $m \times n$ measurement matrix (with $m \ll n$) and $z \in \mathbb{R}^m$ is a vector of measurement errors. Sensing is nonadaptive in that A does not depend on f. The compressed sensing theory asserts that if the unknown signal f is sparse, or nearly sparse, it is possible to recover f under suitable conditions on the matrix A, by convex programming:

$$\min_{\tilde{f}\in R^n} \|\tilde{f}\|_1 \quad \text{subject to } \|A\tilde{f} - y\|_2 \leq \epsilon,$$

Supported by the NSF of China grant 11371154.

where $\|\cdot\|_2$ denotes the Euclidean norm, $\|f\|_1 = \sum_{i=1}^n |f_i|$ is the l_1-norm and $\epsilon \geq 0$ is a likely upper bound on the noise level $\|z\|_2$. We call $\epsilon = 0$ the noiseless case and $\epsilon > 0$ the noisy case.

Let $x \in \mathbb{R}^d$ be a column vector. The support of x is $\text{supp}(x) = \{i : x_i \neq 0, i = 1, \ldots, d\}$. For $k \in \mathbb{N}$, a vector x is said to be k-sparse if $|\text{supp}(x)| \leq k$. For an $m \times n$ measurement matrix A, we say that A obeys the restricted isometry property (RIP) [6] with constant $\delta_k \in (0, 1)$ if

$$(1 - \delta_k)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k)\|x\|_2^2$$

for all k-sparse signals x, where δ_k is called the restricted isometry constant (RIC) of order k of the measurement matrix A.

Various conditions on the RIC for sparse signal recovery have been introduced and studied in the literature. For example, sufficient conditions for exact recovery in the noiseless case include $\delta_{2k} < \sqrt{2} - 1 \approx 0.414$ in Candès and Tao [6], $\delta_{2k} < 0.4531$ in Foucart and Lai [13], $\delta_{2k} < 0.472$ in Cai *et al.* [8], $\delta_{2k} < 0.497$ in Mo and Li [17], $\delta_k < 0.307$ in [7], and $\delta_k < 1/3$ and $\delta_{2k} < 1/2$ in [10]. Recently Cai and Zhang in [11] have shown that for any given constant $t \geq 4/3$, the condition $\delta_{tk} < \sqrt{(t-1)/t}$ is sharp, both for exact recovery of all k-sparse signals in the noiseless case and stable recovery of approximately sparse signals in the noisy case. There are also other sufficient conditions that involve the RIC of different orders; for example, $\delta_{3k} + 3\delta_{4k} < 2$ in Candès *et al.* [5] and $\delta_k + \delta_{2k} < 1$ in Cai and Zhang [9].

The techniques above can hold for signals which are sparse in the standard coordinate basis or sparse with respect to some other orthonormal basis. However, in practice, there are many examples in which a signal of interest is not sparse in an orthonormal basis. More often than not, sparsity is not expressed in terms of an orthonormal basis, but in terms of an overcomplete dictionary. This means that the signal $f \in \mathbb{R}^n$ is now expressed as $f = Dx$, where $D \in \mathbb{R}^{n\times d}$ is some overcomplete dictionary and x is (nearly) sparse. We refer to [3] and the references therein for details.

Let $g_1, g_2, \ldots, g_d$ be the d columns of D. We say that D is a tight frame for $\mathbb{R}^n$ if, for any $f \in \mathbb{R}^n$, the following relations hold:

$$f = \sum_{i=1}^{d} \langle f, g_i \rangle g_i \quad \text{and} \quad \|f\|_2^2 = \sum_{i=1}^{d} |\langle f, g_i \rangle|^2,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard Euclidean inner product. So, for a tight frame D,

$$DD^* = I,$$

where D^* stands for the conjugate transpose of D. We wish to recover the unknown signal $f \in \mathbb{R}^n$, that is sparse or nearly sparse in terms of some tight frame D, from linear measurements $y = Af + z$. This means that D^*f is sparse or nearly sparse. This problem has been considered in [1–3, 14–16, 18, 19]. The methods introduced in [2, 18, 19] force incoherence on the dictionary D so that the matrix AD conforms to the standard compressed sensing result above. As a result, they are not suitable for a dictionary which is largely correlated. Candès *et al.* [3] proposed a new way of recovering signals of this kind from $y = Af + z$ by the use of l_1-analysis optimisation:

$$\hat{f} = \operatorname*{argmin}_{\tilde{f}\in\mathbb{R}^n} \|D^*\tilde{f}\|_1 \quad \text{subject to } \|A\tilde{f} - y\|_2 \le \epsilon, \tag{1.1}$$

where again ϵ is a likely upper bound on the noise level $\|z\|_2$. To this end, a new property called D-RIP was introduced in [14]. It is a natural extension to the standard RIP.

Definition 1.1 (D-RIP). Let D be a tight frame and Σ_k be the set of all k-sparse vectors in $\mathbb{R}^d$. A measurement matrix A is said to obey the restricted isometry property adapted to D (abbreviated as D-RIP) with constant δ_k if

$$(1 - \delta_k)\|Dv\|_2^2 \le \|ADv\|_2^2 \le (1 + \delta_k)\|Dv\|_2^2$$

holds for all $v \in \Sigma_k$. When k is not an integer, we define δ_k as $\delta_{\lceil k\rceil}$.

Candès *et al.* [3] noted that Gaussian matrices and other random compressed sensing matrices satisfy the D-RIP. In fact any $m \times n$ matrix A obeying

$$P((1 - \delta)\|v\|_2^2 \le \|Av\|_2^2 \le (1 + \delta)\|v\|_2^2) \le Ce^{-\gamma m},$$

for fixed $v \in \mathbb{R}^n$ (γ is an arbitrary positive numerical constant), will satisfy the D-RIP with overwhelming probability, provided that $m \gtrsim s\log(d/s)$. See [3] and the references therein for details.

In what follows, δ_k denotes the D-RIP constant with order k of the measurement matrix A. Throughout the paper, we denote by $v_{[k]}$ the vector consisting of the k largest entries of $v \in \mathbb{R}^n$ in magnitude, that is,

$$v_{[k]} = \operatorname*{argmin}_{\|\tilde{v}\|_0 \le k} \|v - \tilde{v}\|_2,$$

where $\|v\|_0 = |\{i : v_i \neq 0\}|$. The main result of Candès *et al.* [3] is that if the measurement matrix A satisfies D-RIP with $\delta_{2k} < 0.08$, then the solution $\hat{f}$ to (1.1) satisfies

$$\|\hat{f} - f\|_2 \le C_0'\epsilon + C_1' \frac{\|D^*f - (D^*f)_{[k]}\|_1}{\sqrt{k}},$$

where the constants C_0' and C_1' may only depend on δ_{2k}. Many researchers have tried to improve on this result. For example, the D-RIP condition $\delta_{2k} < 0.4931$ was used by Li *et al.* [14] and it can be improved to $\delta_{2k} < 0.656$ in some special cases. Recently, the D-RIP condition was improved to $\delta_{2k} < \sqrt{2}/2 \approx 0.7071$ by Baker [1].

The main goal of this paper is to establish a sharp D-RIP condition on δ_{tk} for the recovery of signals that are sparse or nearly sparse in terms of the tight frame D in (1.1). We use the new technique developed in [11] to prove our main result. We will show that our result is an improvement on those in [1, 14].

This paper is organised as follows. In Section 2 we introduce some lemmas and notation. In Section 3 we give our main result.

We conclude this section by giving some notation that will be used throughout the paper. We use $f \in \mathbb{R}^n$ to denote the unknown signal that we want to reconstruct and $A \in \mathbb{R}^{m\times n}$ the measurement matrix. For $v \in \mathbb{R}^n$, we set $\|v\|_1 = \sum_{i=1}^n |v_i|$ and $\|v\|_\infty = \sup_{1\le i\le n} |v_i|$. For a coherent tight frame $D \in \mathbb{R}^{n\times d}$ and $T \subset \{1, 2, \ldots, d\}$, we denote by D_T the matrix D restricted to the columns indexed by T. The notation D_T^* means $(D_T)^*$. The index set T^c stands for the complement of T in $\{1, 2, \ldots, d\}$. For a given vector $h \in \mathbb{R}^n$, we denote by $D_T^* h(i)$ the ith element of $D_T^* h$.

2. Some useful lemmas

In this section, we give some lemmas that will be very useful in the later parts of the paper.

The first lemma relates to the l_1-norm invariant convex s-sparse decomposition (see Wu and Xu [20] and also [11]).

Lemma 2.1. *For positive integers $s \le n$ and a positive constant C, let $v \in \mathbb{R}^n$ be a vector satisfying*

$$\|v\|_1 \le C \quad \text{and} \quad \|v\|_\infty \le \frac{C}{s}.$$

Then there are s-sparse vectors $u_1, u_2, \ldots, u_N$ with

$$\|u_i\|_1 = \|v\|_1 \quad \text{and} \quad \|u_i\|_\infty \le \frac{C}{s}, \quad \text{for } i = 1, 2, \ldots, N,$$

such that

$$v = \sum_{i=1}^N \lambda_i u_i$$

for some nonnegative real numbers $\lambda_1, \lambda_2, \ldots, \lambda_N$ satisfying $\sum_{i=1}^N \lambda_i = 1$.

The next l_2-norm identity is very important in the proof of our main result.

Lemma 2.2. *Let $\beta_i \in \mathbb{R}^d$, λ_i and c be nonnegative real numbers with $\sum_{i=1}^N \lambda_i = 1$. The following equality holds for any $m \times d$ matrix B:*

$$\sum_{i=1}^N \lambda_i \Big\| B\Big(\sum_{j=1}^N \lambda_j \beta_j - c\beta_i\Big)\Big\|_2^2 + (1-2c) \sum_{1\le i<j\le N} \lambda_i\lambda_j \|B(\beta_i - \beta_j)\|_2^2 = \sum_{i=1}^N \lambda_i (1-c)^2 \|B\beta_i\|_2^2.$$

Proof. Set $x_i = B\beta_i$. The desired equality can be written as

$$\sum_{i=1}^N \lambda_i \Big\| \sum_{j=1}^N \lambda_j x_j - c x_i \Big\|_2^2 + (1-2c) \sum_{1\le i<j\le N} \lambda_i\lambda_j \|x_i - x_j\|_2^2 = \sum_{i=1}^N \lambda_i (1-c)^2 \|x_i\|_2^2.$$

https://doi.org/10.1017/S0004972715000933 Published online by Cambridge University Press

By an elementary calculation, we have

$$\begin{aligned}
&\sum_{i=1}^{N}\lambda_i\Big\|\sum_{j=1}^{N}\lambda_j x_j - c x_i\Big\|_2^2\\
&\quad=\sum_{i=1}^{N}\lambda_i\Big\langle\sum_{j=1}^{N}\lambda_j x_j - c x_i, \sum_{j=1}^{N}\lambda_j x_j - c x_i\Big\rangle\\
&\quad=\sum_{i=1}^{N}\lambda_i\Big(\Big\|\sum_{i=1}^{N}\lambda_j x_j\Big\|_2^2 - 2c\Big\langle x_i, \sum_{i=1}^{N}\lambda_j x_j\Big\rangle + c^2\|x_i\|_2^2\Big)\\
&\quad=(1-2c)\Big\|\sum_{i=1}^{N}\lambda_j x_j\Big\|_2^2 + c^2\sum_{i=1}^{N}\lambda_i\|x_i\|_2^2\\
&\quad=(1-2c)\Big[\sum_{i=1}^{N}\lambda_j^2\|x_j\|_2^2 + 2\sum_{1\le i<j\le N}\lambda_i\lambda_j\langle x_i, x_j\rangle\Big] + c^2\sum_{i=1}^{N}\lambda_i\|x_i\|_2^2. \qquad (2.1)
\end{aligned}$$

We also have

$$(1-2c)\sum_{1\le i<j\le N}\lambda_i\lambda_j\|x_i - x_j\|_2^2 = (1-2c)\Big[\sum_{1\le i<j\le N}\lambda_i\lambda_j(\|x_i\|_2^2 - 2\langle x_i, x_j\rangle + \|x_j\|_2^2)\Big]. \qquad (2.2)$$

By adding Equations (2.1) and (2.2),

$$\begin{aligned}
&\sum_{i=1}^{N}\lambda_i\Big\|\sum_{j=1}^{N}\lambda_j x_j - c x_i\Big\|_2^2 + (1-2c)\sum_{1\le i<j\le N}\lambda_i\lambda_j\|x_i - x_j\|_2^2\\
&\quad=(1-2c)\Big[\sum_{i=1}^{N}\lambda_j^2\|x_j\|_2^2 + \sum_{1\le i<j\le N}\lambda_i\lambda_j(\|x_i\|_2^2 + \|x_j\|_2^2)\Big] + c^2\sum_{i=1}^{N}\lambda_i\|x_i\|_2^2\\
&\quad=(1-2c)\Big[\lambda_1\Big(\sum_{i=1}^{N}\lambda_i\Big)\|x_1\|_2^2 + \lambda_2\Big(\sum_{i=1}^{N}\lambda_i\Big)\|x_2\|_2^2 + \cdots + \lambda_N\Big(\sum_{i=1}^{N}\lambda_i\Big)\|x_N\|_2^2\Big]\\
&\qquad + c^2\sum_{i=1}^{N}\lambda_i\|x_i\|_2^2\\
&\quad=(1-2c)\sum_{i=1}^{N}\lambda_i\|x_i\|_2^2 + c^2\sum_{i=1}^{N}\lambda_i\|x_i\|_2^2 = \sum_{i=1}^{N}\lambda_i(1-c)^2\|x_i\|_2^2.
\end{aligned}$$

The proof is complete. □

Since DD^* and D^*D have the same nonzero eigenvalues, it is easy to establish the following inequality for a tight frame.

LEMMA 2.3. *Let D be a $n \times d$ tight frame. Then, for any $v \in \mathbb{R}^d$,*

$$\|Dv\|_2 \le \|v\|_2.$$

We conclude this section by stating a lemma from [10].

https://doi.org/10.1017/S0004972715000933 Published online by Cambridge University Press

Lemma 2.4. *Given positive integers d and k with $d \geq k$ and numbers $a_1 \geq a_2 \geq \cdots \geq a_d \geq 0$ satisfying $\sum_{i=1}^{k} a_i \geq \sum_{j=k+1}^{d} a_j$, then for any $\alpha \geq 1$,*

$$\sum_{i=1}^{k} a_i^{\alpha} \geq \sum_{j=k+1}^{d} a_j^{\alpha}.$$

More generally, if $a_1 \geq a_2 \geq \cdots \geq a_d \geq 0$, $\lambda \geq 0$ and $\sum_{i=1}^{k} a_i + \lambda \geq \sum_{j=k+1}^{d} a_j$, then for any $\alpha \geq 1$,

$$\sum_{j=k+1}^{d} a_j^{\alpha} \leq k\Big[\Big(\frac{1}{k}\sum_{i=1}^{k} a_i^{\alpha}\Big)^{1/\alpha} + \frac{\lambda}{k}\Big]^{\alpha}.$$

3. Main results

The main result of the paper is the following theorem.

Theorem 3.1. *Let D be an arbitrary $n \times d$ tight frame and A be an $m \times n$ measurement matrix satisfying the D-RIP with $\delta_{tk} < \sqrt{(t-1)/t}$ for some $t > 1$. Then the solution $\hat{f}$ of (1.1) satisfies*

$$\|\hat{f} - f\|_2 \leq \sqrt{2}C_1\epsilon + \frac{2(\sqrt{2}C_2 + 1)}{\sqrt{k}}\|D^*f - (D^*f)_{[k]}\|_1, \tag{3.1}$$

where

$$C_1 = \frac{2\sqrt{t(t-1)(1+\delta_{tk})}}{t(\sqrt{(t-1)/t} - \delta_{tk})}, \quad C_2 = \frac{2\delta_{tk} + \sqrt{2t(\sqrt{(t-1)/t} - \delta_{tk})\delta_{tk}}}{2t(\sqrt{(t-1)/t} - \delta_{tk})}.$$

Proof. We first assume that tk is an integer. Set $h = \hat{f} - f$. Let T_0 be the set of indices corresponding to the largest k components of D^*h in magnitude. Since $\hat{f}$ is the solution of (1.1), we have

$$\begin{aligned}\|D^*f\|_1 \geq \|D^*\hat{f}\|_1 &= \|D^*h + D^*f\|_1 = \|D^*_{T_0}h + D^*_{T_0^c}h + D^*_{T_0^c}f + D^*_{T_0}f\|_1\\ &= \|D^*_{T_0}h + D^*_{T_0}f\|_1 + \|D^*_{T_0^c}h + D^*_{T_0^c}f\|_1\\ &\geq \|D^*_{T_0}f\|_1 - \|D^*_{T_0}h\|_1 + \|D^*_{T_0^c}h\|_1 - \|D^*_{T_0^c}f\|_1.\end{aligned}$$

This implies

$$\|D^*_{T_0^c}h\|_1 \leq \|D^*_{T_0}h\|_1 + 2\|D^*_{T_0^c}f\|_1. \tag{3.2}$$

We also have

$$\|Ah\|_2 \leq \|y - A\hat{f}\|_2 + \|y - Af\|_2 \leq 2\epsilon.$$

Set $\alpha = (\|D^*_{T_0}h\|_1 + 2\|D^*_{T_0^c}f\|_1)/k$. We split $D^*_{T_0^c}h$ into $D^*_{T_0^c}h = D^*_{\Lambda_1}h + D^*_{\Lambda_2}h$, where the ith elements of $D^*_{\Lambda_1}h$ and $D^*_{\Lambda_2}h$ are defined by

https://doi.org/10.1017/S0004972715000933 Published online by Cambridge University Press

$$D^*_{\Lambda_1}h(i) = \begin{cases} D^*_{T_0^c}h(i) & \text{if } |D^*_{T_0^c}h(i)| > \alpha/(t-1), \\ 0 & \text{otherwise,} \end{cases}$$

and

$$D^*_{\Lambda_2}h(i) = \begin{cases} D^*_{T_0^c}h(i) & \text{if } |D^*_{T_0^c}h(i)| \le \alpha/(t-1), \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that $\|D^*_{\Lambda_1}h\|_1 \le \|D^*_{T_0^c}h\|_1 \le k\alpha$. Denote $|\text{supp}\{D^*_{\Lambda_1}h\}| = \|D^*_{\Lambda_1}h\|_0 = p$. From the definition of $D^*_{\Lambda_1}h$,

$$k\alpha \ge \|D^*_{\Lambda_1}h\|_1 = \sum_{i\in\text{supp}\{D^*_{\Lambda_1}h\}} |D^*_{\Lambda_1}h(i)| \ge \sum_{i\in\text{supp}\{D^*_{\Lambda_1}h\}} \frac{\alpha}{t-1} = \frac{p\alpha}{t-1}.$$

which gives $p \le k(t-1)$. Hence,

$$\|D^*_{\Lambda_2}h\|_1 = \|D^*_{T_0^c}h\|_1 - \|D^*_{\Lambda_1}h\|_1 \le k\alpha - \frac{p\alpha}{t-1} = (k(t-1)-p)\cdot\frac{\alpha}{t-1}.$$

By the definition of $D^*_{\Lambda_2}h$, it is obvious that

$$\|D^*_{\Lambda_2}h\|_\infty \le \frac{\alpha}{t-1}.$$

By Lemma 2.1, we can express $D^*_{\Lambda_2}h$ as a convex combination of sparse vectors:

$$D^*_{\Lambda_2}h = \sum_{i=1}^{N}\lambda_i u_i,$$

where u_i is $(k(t-1)-p)$-sparse, $\|u_i\|_1 = \|D^*_{\Lambda_2}h\|_1$ and $\|u_i\|_\infty \le \alpha/(t-1)$. Hence,

$$\|u_i\|_2 \le \sqrt{\|u_i\|_0}\|u_i\|_\infty \le \sqrt{k(t-1)-p}\|u_i\|_\infty \le \sqrt{k(t-1)}\|u_i\|_\infty \le \alpha\sqrt{k/(t-1)}.$$

Let $\mu \ge 0$, $c \ge 0$ be constants to be determined and set $\beta_i = D^*_{T_0}h + D^*_{\Lambda_1}h + \mu u_i$. Then

$$\begin{aligned}\sum_{j=1}^{N}\lambda_j\beta_j - c\beta_i &= D^*_{T_0}h + D^*_{\Lambda_1}h + \mu D^*_{\Lambda_2}h - c\beta_i \\ &= (1-\mu-c)(D^*_{T_0}h + D^*_{\Lambda_1}h) - c\mu u_i + \mu D^*h.\end{aligned}$$

https://doi.org/10.1017/S0004972715000933 Published online by Cambridge University Press

Taking $c = 1/2$ and applying Lemma 2 with $B = AD$, we obtain

$$
\begin{aligned}
0 &= \sum_{i=1}^{N} \lambda_i \Big\| AD\Big((D_{T_0}^* h + D_{\Lambda_1}^* h + \mu D_{\Lambda_2}^* h) - \frac{1}{2}(D_{T_0}^* h + D_{\Lambda_1}^* h + \mu u_i)\Big)\Big\|_2^2 - \sum_{i=1}^{N} \frac{\lambda_i}{4} \|AD\beta_i\|_2^2 \\
&= \sum_{i=1}^{N} \lambda_i \Big\| AD\Big(\Big(\frac{1}{2} - \mu\Big)(D_{T_0}^* h + D_{\Lambda_1}^* h) - \frac{\mu}{2} u_i + \mu D^* h\Big)\Big\|_2^2 - \sum_{i=1}^{N} \frac{\lambda_i}{4} \|AD\beta_i\|_2^2 \\
&= \sum_{i=1}^{N} \lambda_i \Big\| AD\Big(\Big(\frac{1}{2} - \mu\Big)(D_{T_0}^* h + D_{\Lambda_1}^* h) - \frac{\mu}{2} u_i\Big)\Big\|_2^2 + \mu^2 \|Ah\|_2^2 \\
&\quad + 2\Big\langle AD\Big(\frac{1}{2} - \mu\Big)(D_{T_0}^* h + D_{\Lambda_1}^* h) - \frac{\mu}{2} D_{\Lambda_2}^* h, \mu Ah\Big\rangle - \sum_{i=1}^{N} \frac{\lambda_i}{4} \|AD\beta_i\|_2^2. \\
&= \sum_{i=1}^{N} \lambda_i \Big\| AD\Big(\Big(\frac{1}{2} - \mu\Big)(D_{T_0}^* h + D_{\Lambda_1}^* h) - \frac{\mu}{2} u_i\Big)\Big\|_2^2 \\
&\quad + \mu(1-\mu)\langle AD((D_{T_0}^* h + D_{\Lambda_1}^* h), Ah\rangle - \sum_{i=1}^{N} \frac{\lambda_i}{4} \|AD\beta_i\|_2^2.
\end{aligned}
$$

Since $\|D_{T_0}^* h\|_0 \le k$, $\|D_{\Lambda_1}^* h\|_0 \le p$ and $\|u_i\|_0 \le k(t-1) - p$, the vectors

$$
\Big(\frac{1}{2} - \mu\Big)(D_{T_0}^* h + D_{\Lambda_1}^* h) - \frac{\mu}{2} u_i \quad \text{and} \quad \beta_i = D_{T_0}^* h + D_{\Lambda_1}^* h + \mu u_i
$$

are tk-sparse. By the definition of D-RIP and the triangle inequality,

$$
\begin{aligned}
0 \le (1 + \delta_{tk}) &\sum_{i=1}^{N} \lambda_i \Big[\Big(\frac{1}{2} - \mu\Big)^2 \|D(D_{T_0}^* h + D^* \Lambda_1 h)\|_2^2 + \frac{\mu^2}{4} \|Du_i\|_2^2\Big] \\
&+ \mu(1-\mu)\langle AD(D_{T_0}^* h + D_{\Lambda_1}^* h), Ah\rangle \\
&- (1 - \delta_{tk}) \sum_{i=1}^{N} \frac{\lambda_i}{4} (\|D(D_{T_0}^* h + D_{\Lambda_1}^* h)\|_2^2 + \mu^2 \|Du_i\|_2^2).
\end{aligned} \tag{3.3}
$$

Also, from Lemma 2.3 and the definition of D-RIP,

$$
\begin{aligned}
\langle AD(D_{T_0}^* h + D_{\Lambda_1}^* h), Ah\rangle &\le \|AD(D_{T_0}^* h + D_{\Lambda_1}^* h)\|_2 \|Ah\|_2 \\
&\le 2\epsilon \sqrt{1 + \delta_{tk}} \|D(D_{T_0}^* h + D_{\Lambda_1}^* h)\|_2 \\
&\le 2\epsilon \sqrt{1 + \delta_{tk}} \|D_{T_0}^* h + D_{\Lambda_1}^* h\|_2.
\end{aligned} \tag{3.4}
$$

https://doi.org/10.1017/S0004972715000933 Published online by Cambridge University Press

Substituting (3.4) into (3.3) and using Lemma 2.3 again,

$$
\begin{aligned}
0 &\le (1+\delta_{tk})\sum_{i=1}^{N}\lambda_i\Big[\Big(\frac{1}{2}-\mu\Big)^2\|D^*_{T_0}h + D^*\Lambda_1 h\|_2^2 + \frac{\mu^2}{4}\|u_i\|_2^2\Big] \\
&\quad + 2\epsilon\mu(1-\mu)\sqrt{1+\delta_{tk}}\|D^*_{T_0}h + D^*_{\Lambda_1}h\|_2 \\
&\quad - (1-\delta_{tk})\sum_{i=1}^{N}\frac{\lambda_i}{4}(\|D^*_{T_0}h + D^*_{\Lambda_1}h\|_2^2 + \mu^2\|u_i\|_2^2) \\
&= \sum_{i=1}^{N}\lambda_i\Big[(1+\delta_{tk})\Big(\frac{1}{2}-\mu\Big)^2 - \frac{(1-\delta_{tk})}{4}\|D^*_{T_0}h + D^*\Lambda_1 h\|_2^2 + \frac{\delta_{tk}\mu^2}{2}\|u_i\|_2^2\Big] \\
&\quad + 2\epsilon\mu(1-\mu)\sqrt{1+\delta_{tk}}\|D^*_{T_0}h + D^*_{\Lambda_1}h\|_2. \qquad (3.5)
\end{aligned}
$$

Set $x = \|D^*_{T_0}h + D^*_{\Lambda_1}h\|_2$ and $P = 2k^{-1/2}\|D^*_{T_0^c}f\|_1$. Since $\alpha = (\|D^*_{T_0}h\|_1 + 2\|D^*_{T_0^c}f\|_1)/k$,

$$
\begin{aligned}
\|u_i\|_2 \le \sqrt{k/(t-1)}\alpha &\le \frac{\|D^*_{T_0}h\|_2}{\sqrt{t-1}} + \frac{2\|D^*_{T_0^c}f\|_1}{\sqrt{k(t-1)}} \\
&\le \frac{\|D^*_{T_0}h + D^*_{\Lambda_1}h\|_2}{t-1} + \frac{2\|D^*_{T_0^c}f\|_1}{\sqrt{k(t-1)}} = \frac{x+P}{\sqrt{t-1}}.
\end{aligned}
$$

It then follows from (3.5) that

$$
\begin{aligned}
0 &\le \Big[(\mu^2-\mu) + \delta_{tk}\Big(\mu^2 - \mu + \frac{1}{2} + \frac{\mu^2}{2(t-1)}\Big)\Big]x^2 \\
&\quad + \Big(\frac{P\mu^2\delta_{tk}}{t-1} + 2\epsilon\mu(1-\mu)\sqrt{1+\delta_{tk}}\Big)x + \frac{\mu^2P^2\delta_{tk}}{2(t-1)}.
\end{aligned}
$$

If we let $\mu = \sqrt{t(t-1)} - (t-1)$, then the last inequality becomes

$$
\begin{aligned}
0 &\le -t[(2t-1) - 2\sqrt{t(t-1)}]\Big(\sqrt{\frac{t-1}{t}} - \delta_{tk}\Big)x^2 \\
&\quad + \Big[\frac{P\mu^2\delta_{tk}}{t-1} + 2\epsilon\mu^2\sqrt{\frac{t}{t-1}}\sqrt{1+\delta_{tk}}\Big]x + \frac{\mu^2P^2\delta_{tk}}{2(t-1)} \\
&= \frac{\mu^2}{t-1}\Big[-t\Big(\sqrt{\frac{t-1}{t}} - \delta_{tk}\Big)x^2 + (P\delta_{tk} + 2\epsilon\sqrt{t(t-1)(1+\delta_{tk})})x + \frac{P^2\delta_{tk}}{2}\Big].
\end{aligned}
$$

By assumption $\delta_{tk} \le \sqrt{(t-1)/t}$. So, we can solve the last inequality, which we abbreviate as $0 \le Ax^2 + Bx + C$ for x, to obtain

$$
\begin{aligned}
x &\le \frac{B + \sqrt{B^2 + 4AC}}{2A} \\
&\le \frac{2\sqrt{t(t-1)(1+\delta_{tk})}}{t(\sqrt{(t-1)/t} - \delta_{tk})}\epsilon + \frac{2\delta_{tk} + \sqrt{2t(\sqrt{(t-1)/t} - \delta_{tk})\delta_{tk}}}{2t(\sqrt{(t-1)/t} - \delta_{tk})}P.
\end{aligned}
$$

https://doi.org/10.1017/S0004972715000933 Published online by Cambridge University Press

Define C_1 and C_2 as in the statement of the theorem. The last inequality can be rewritten as

$$x \le C_1\epsilon + C_2 P. \tag{3.6}$$

Since D is a tight frame, we have

$$\|\hat{f} - f\|_2 = \|h\|_2 = \|D^*h\|_2 = \sqrt{\|D^*_{T_0}h\|_2^2 + \|D^*_{T_0^c}h\|_2^2}. \tag{3.7}$$

It follows from (3.2) that

$$\|D^*_{T_0^c}h\|_1 \le \|D^*_{T_0}h\|_1 + P\sqrt{k} \le \|D^*_{T_0^c}h\|_2 \le \|D^*_{T_0}h\|_2 + P, \tag{3.8}$$

where the last inequality follows from Lemma 2.4. From (3.8) together with (3.7),

$$\begin{aligned}
\|\hat{f} - f\|_2 &\le \sqrt{\|D^*_{T_0}h\|_2^2 + (\|D^*_{T_0}h\|_2 + P)^2} \\
&\le \sqrt{2\|D^*_{T_0}h\|_2^2} + P \\
&\le \sqrt{2}x + P \\
&\le \sqrt{2}C_1\epsilon + \frac{2(\sqrt{2}C_2 + 1)}{\sqrt{k}}\|D^*_{T_0^c}f\|_1 \\
&= \sqrt{2}C_1\epsilon + \frac{2(\sqrt{2}C_2 + 1)}{\sqrt{k}}\|D^*f - (D^*f)_{[k]}\|_1,
\end{aligned}$$

where the last inequality follows from (3.6) and the definition of P, and the equality follows from the definition of T_0. This establishes (3.1).

In the case where tk is not an integer, we work with $t' = \lceil tk \rceil / k$. It is clear that $t' > t$ and that the integer $t'k$ satisfies

$$\delta_{t'k} = \delta_{tk} < \sqrt{\frac{t-1}{t}} < \sqrt{\frac{t'-1}{t'}}.$$

Repeating the above process with δ_{tk} replaced by $\delta_{t'k}$, it is not difficult to derive (3.1). The proof is complete. □

REMARK 3.2.

(a) When $\|D^*_{T_0^c}f\|_1 = 0$ and $\epsilon = 0$, that is, $x = D^*f$ is k-sparse and the measurement error is zero, the condition $\delta_{tk} \le \sqrt{(t-1)/t}$ guarantees exact recovery of f from problem (1.1).

(b) When $D = I$, which corresponds to the case of standard compressed sensing, our result is consistent with [11, Theorem 2.1].

(c) Cai and Zhang [11] have shown that in the special case $D = I$, for any $t \ge 4/3$ the condition $\delta_{tk} \le \sqrt{(t-1)/t}$ (δ_{tk} is the RIC) is sharp for both exact recovery in the noiseless case and stable recovery in the noisy case. It is not difficult to show that for any tight frame D the condition $\delta_{tk} \le \sqrt{(t-1)/t}$ is also sharp when $t \ge 4/3$. In this sense, our result extends the result in [11].

(d) In the particular case $t = 2$, the condition in Theorem 3.1 can be simplified to $\delta_{2k} < 2^{-1/2}$, which is the same as that in [1]. However, the result in Theorem 3.1 is better than that in [1]. Indeed, if $\delta_{2k} \leq 1/2$, we have $\sqrt{2}C_1 \approx 11.8272$ and $2(\sqrt{2}C_2 + 1)k^{-1/2} \approx 7.6116k^{-1/2}$, which are strictly less than 16.7262 and $9.9360k^{-1/2}$, the respective coefficients in the result of [1]. Consequently, the result in Theorem 3.1 is an improvement on the comparable result in [1].

References

[1] C. Baker, 'A note on sparsification by frames', Preprint, 2013, arXiv:1308.5249.

[2] W. Bajwa, R. Calderbank and S. Jafarpour, 'Why Gabor frames? Two fundamental measures of coherence and their geometric significance', *J. Commun. Netw.* **12** (2010), 289–307.

[3] E. J. Candès, Y. C. Eldar, D. Needell and P. Randall, 'Compressed sensing with coherent and redundant dictionaries', *Appl. Comput. Harmon. Anal.* **31** (2011), 59–73.

[4] E. J. Candès, J. Romberg and T. Tao, 'Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information', *IEEE Trans. Inform. Theory* **52** (2006), 489–509.

[5] E. J. Candès, J. K. Romberg and T. Tao, 'Stable signal recovery from incomplete and inaccurate measurements', *Comm. Pure Appl. Math.* **59** (2006), 1207–1223.

[6] E. J. Candès and T. Tao, 'Decoding by linear programming', *IEEE Trans. Inform. Theory* **51** (2005), 4203–4215.

[7] T. Cai, L. Wang and G. Xu, 'New bounds for restricted isometry constants', *IEEE Trans. Inform. Theory* **56** (2010), 4388–4394.

[8] T. Cai, L. Wang and G. Xu, 'Shifting inequality and recovery of sparse signals', *IEEE Trans. Signal Process.* **58** (2010), 1300–1308.

[9] T. Cai and A. Zhang, 'Compressed sensing and affine rank minimization under restricted isometry', *IEEE Trans. Signal Process.* **61** (2013), 3279–3290.

[10] T. Cai and A. Zhang, 'Sharp RIP bound for sparse signal and low-rank matrix recovery', *Appl. Comput. Harmon. Anal.* **35** (2013), 74–93.

[11] T. Cai and A. Zhang, 'Sparse representation of a polytope and recovery of sparse signals and low-rank matrices', *IEEE Trans. Inform. Theory* **60** (2014), 122–132.

[12] D. Donoho, 'Compressed sensing', *IEEE Trans. Inform. Theory* **52** (2006), 1289–1306.

[13] S. Foucart and M. J. Lai, 'Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$', *Appl. Comput. Harmon. Anal.* **26** (2009), 395–407.

[14] S. Li and J. Lin, 'Compressed sensing with coherent tight frames via l_q-minimization for $0 < q \leq 1$', *Inverse Probl. Imaging* **8** (2011), doi:10.3934/ipi.2014.8761.

[15] J. Lin and S. Li, 'Sparse recovery with coherent tight frames via analysis Dantzig selector and analysis LASSO', *Appl. Comput. Harmon. Anal.* **37** (2014), 126–139.

[16] J. Lin, S. Li and Y. Shen, 'New bounds for restricted isometry constants with coherent tight frames', *IEEE Trans. Signal Process.* **61** (2013), 611–621.

[17] Q. Mo and S. Li, 'New bounds on the restricted isometry constant δ_{2k}', *Appl. Comput. Harmon. Anal.* **31** (2011), 460–468.

[18] H. Rauhut, K. Schnass and P. Vandergheynst, 'Compressed sensing and redundant dictionaries', *IEEE Trans. Inform. Theory* **54** (2008), 2210–2219.

[19] J. A. Tropp, 'Greed is good: Algorithmic results for sparse approximation', *IEEE Trans. Inform. Theory* **50** (2004), 2231–2242.

[20] G. Wu and Z. Xu, 'On the l_1-norm invariant convex k-sparse decomposition of signals', *J. Oper. Res. Soc. China* **1** (2013), 537–541.

FEN-GONG WU, School of Mathematical Sciences,
South China Normal University, Guangzhou 510631, China
e-mail: wufengong@gdei.edu.cn

DONG-HUI LI, School of Mathematical Sciences,
South China Normal University, Guangzhou 510631, China
e-mail: dhli@scnu.edu.cn