TELKOMNIKA, Vol.11, No.1, January 2013, pp. 213~226
ISSN: 2302-4046

Lebesgue-type Inequality for Orthogonal Matching Pursuit for μ-coherent Dictionaries

Ye Peixin*, Wei Xiujie
School of Mathematics and LPMC, Nankai University, Tianjin 300071, China
*corresponding author, e-mail: yepx@nankai.edu.cn

Abstract

In this paper, we investigate the efficiency of some kind of Greedy Algorithms with respect to dictionaries from Hilbert spaces. We establish ideal Lebesgue-type inequality for Orthogonal Matching Pursuit also known in literature as the Orthogonal Greedy Algorithm for μ -coherent dictionaries. We show that the Orthogonal Matching Pursuit provides an almost optimal approximation on the first $[1/20\mu]$.

***Keywords**: orthogonal matching pursuit (OMP), orthogonal greedy algorithms, best m -term, Lebesgue-type inequalities, dictionaries*

1. Introduction

In this paper we continue to study the convergence of greedy algorithms with regards to dictionaries with small coherence (see [1-9]). The purpose of the research of approximation with regard to incoherent dictionaries was to apply in compressed sensing. In [1, 8, 10, 11], it was shown that the Orthogonal Matching Pursuit is effective for signal recovering. In this paper, we will discuss this problem from the point of view of Approximation Theory. We obtain upper estimate for the error of approximation by OMP in terms of the error of the best m-term approximation. This article is a development of recent results obtains by Eugene Livshitz in [12].

Let us recall the standard definitions of Greedy Algorithms. We say that a set D from a Hilbert Space H is a dictionary if

$$\phi \in \mathrm{D} \Rightarrow \| \phi \| = 1, \ and \ \overline{span \mathrm{D}} = H.$$

The coherence of a dictionary is defined as

$$\mu := \sup_{\phi,\psi \in D, \phi \neq \psi} |\langle \phi, \psi \rangle| \tag{1}$$

Dictionaries with coherence μ are called μ -coherent.

ORTHOGONAL MATCHING PURSUIT (OMP) Set $f_0 := f \in H$, and $OMP_0(f, \mathrm{D}) := 0$. For each $m \geq 0$, we can find a $g_{m+1} \in \mathrm{D}$ such that

$$|\langle f_m, g_{m+1} \rangle| = \sup_{g \in \mathrm{D}} |\langle f_m, g \rangle| ,$$

and define

$$OMP_{m+1}(f, \mathrm{D}) := Proj_{span(g_1, \cdots, g_{m+1})}(f),$$

$$f_{m+1} := f - OMP_{m+1}(f, \mathrm{D}).$$

The best m -term approximation for a function $f \in H,$ is defined as

Received October 1, 2012; Revised November 29, 2012; Accepted December 5, 2012

$$\sigma_m(f) := \sigma_m(f, \mathcal{D}) := \inf_{c_i \in \mathbb{R}, \phi_i \in \mathcal{D}, 1 \le i \le m} \| f - \sum_{i=1}^{m} c_i \phi_i \|.$$

Suppose that dictionary $\mathcal{D}$ is μ-coherent and $m < \frac{1}{2}(\mu^{-1}+1)$. It is well known (see [3,4]) that if $f \in H$ is m-sparse, that is $\sigma_m(f, \mathcal{D}) = 0$, then

$$f = OMP_m(f, \mathcal{D}). \tag{2}$$

Moreover, Temlyakov and Zheltov showed that if $m > \frac{1}{2}(\mu^{-1}+1)$, then equality (2) does not hold for all μ-coherent dictionaries $\mathcal{D}$ and all m-sparse f :

$$\exists \mathcal{D}, m \ge \frac{1}{2}(\mu(\mathcal{D})^{-1}+1), f \in H : \| f - OMP_m(f, \mathcal{D}) \| > 0, \tag{3}$$
$$\sigma_m(f, \mathcal{D}) = 0.$$

Following Temlyakov, we recall results of *Lebesgue type inequalities* which connecting the error of Greedy approximation and the best m-term approximation. These inequalities do hold not for all $m \in \mathbb{N}$, but only for $m \le C(\mu)$; they provide an estimate for the quality of approximation of A(m) iteration of OMP by the best m-term approximation:

$$\| f - OMP_{A(m)}(f, \mathcal{D}) \| \le B(m)\sigma_m(f, \mathcal{D}), m \le C(\mu) \tag{4}$$

with some $A(m) \in \mathbb{N}$, $B(m), C(\mu) \in \mathbb{R}$.

The first Lebesgue type inequality for Greedy Algorithms was obtained by Gilbert et al. in [3] They established (4) for an optimal $A(m) := m,$ an order-optimal

$$C(\mu) = \frac{1}{8\sqrt{2}\mu} - 1,$$

and fast growing

$$B(m) := 8m^{1/2}.$$

Donoho et al. [2] obtained inequality (4) with optimal (up to a constant factor) B(m)=24, but not optimal A(m):=$\lfloor m \log m \rfloor$ and $C(\mu) = \frac{1}{20\mu^{2/3}}$.

Temlyakov and Zheltov [10] proved inequality (4) with $A(m) := m\left\lfloor 2^{\sqrt{\log m}} \right\rfloor,$ B(m):=3 and $C(\mu),$ which guarantees inequality $m2^{\sqrt{2\log m}} \le \frac{1}{26\mu}$ In other words, they proved

Theorem1. *For every $\mu-$coherent dictionary $\mathcal{D}$ and any function $f \in H$,*

$$\| f - OMP_{m\left\lfloor 2^{\sqrt{\log m}} \right\rfloor}(f, \mathcal{D}) \| \le 3\sigma_m(f, \mathcal{D}), if\ m2^{\sqrt{2\log m}} \le \frac{1}{26\mu}.$$

Recently Eugene Livshitz [6] prove (4) with $A(m) := 2m, B(m) := 2.7, C(\mu) = \frac{1}{20\mu}$. In other words, he provedhich guarantees inequality $m2^{\sqrt{2\log m}} \leq \frac{1}{26\mu}$. In other words, they proved

Theorem2. *For every* $\mu -$*coherent dictionary* D *and any function* $f \in H$,

$$\| f - OMP_{2m}(f, \mathrm{D}) \| = \| f_{2m} \| \leq 2.7\sigma_m(f, \mathrm{D})$$

for all $1 \leq m \leq \frac{1}{20\mu}$.

The aim of this paper is to prove (4) with

$$A(m) := 2m, \qquad B(m) := 2.47(1+\varepsilon), \forall\, \varepsilon > 0, \qquad C(\mu) = \frac{1}{20\mu}, \tag{5}$$

and thereby improves above result.

Theorem3. *For every* $\mu -$*coherent dictionary* D *and any function* $f \in H$, $\forall\, \varepsilon > 0$, we have

$$\| f_{2m} \| = \| f - OMP_{2m}(f, \mathrm{D}) \| \leq 2.47(1+\varepsilon)\sigma_m(f, \mathrm{D})$$

for all

$$1 \leq m \leq \frac{1}{20\mu}.$$

2. Preliminary lemmas

By conditions of Theorem 3 , we have

$$\mu \leq m\mu \leq 1/20. \tag{6}$$

We use several standard lemmas to prove Theorem 3.

Lemma 1. *For any* $n, 1 \leq n \leq 2m$, *and*

$$h = \sum_{i=1}^{n} c_i \phi_i,\ c_i \in \square,\ \phi_i \in \mathrm{D},$$

we have

$$\max_{1\leq i\leq n} |\langle h, \phi_i \rangle| \leq \max_{1\leq i\leq n} |c_i|(1+2m\mu), \tag{7}$$

$$\max_{1\leq i\leq n} |\langle h, \phi_i \rangle| \geq \max_{1\leq i\leq n} |c_i|[1-(2m-1)\mu], \tag{8}$$

$$\max_{1\le i\le n} |c_i| \le K_1 \max_{1\le i\le n} |\langle h, \phi_i \rangle|, \tag{9}$$

where

$$K_1 := \frac{1}{1-(2m-1)\mu} \le \frac{10}{9+10\mu} \le \frac{10}{9} \tag{10}$$

proof . For any $1 \le i \le n$, using (1), we have

$$\langle h, \phi_i \rangle = \langle c_i \phi_i, \phi_i \rangle + \langle \sum_{1\le j\le n, i\ne j} c_j \phi_j, \phi_i \rangle$$

$$\le c_i + (n-1)(\max_{1\le i\le n} |c_i|)\mu$$

$$\le c_i + (\max_{1\le i\le n} |c_i|)2m\mu$$

$$\le (\max_{1\le i\le n} |c_i|)(1+2m\mu).$$

Then we get inequality (7).
Similarly

$$\langle h, \phi_i \rangle \ge c_i - (n-1)(\max_{1\le i\le n} |c_i|)\mu$$

This imply inequality (8).
Inequality (9) follows from (8).
As a consequence of Lemma 1, we derive the following lemma.

Lemma 2. *Let* $n \le 2m, h \in H, \phi_i \in \mathrm{D}, 1 \le i \le n.$ Suppose that

$$Proj_{span(\phi_1,\cdots,\phi_n)}(h) = \sum_{i=1}^{n} c_i \phi_i.$$

Then

$$\max_{1\le i\le n} |c_i| \le K_1 \max_{1\le i\le n} |\langle h, \phi_i \rangle|.$$

proof .Set

$$h' = Proj_{span(\phi_1,\cdots,\phi_n)}(h) = \sum_{i=1}^{n} c_i \phi_i.$$

It is easy to see that

$$\langle h, \phi_i \rangle = \langle h', \phi_i \rangle, 1 \le i \le n.$$

Thus the lemma follows from inequality (9) for h' .

For $n \ge 1$ we define

$$d_n := \langle f_{n-1}, g_n \rangle. \tag{11}$$

Let number $x_{i,n}, n \ge 1, \ 1 \le i \le n$ satisfy the equality

$$f_n = f_{n-1} - \sum_{i=1}^{n} x_{i,n} g_i. \tag{12}$$

The following lemma tell us how the value of $x_{i,n}$ depends on the coherence of the dictionary.

Lemma 3. *For any* $n \le 2m$, *we can get the following estimates*:

$$| x_{i,n} | \le K_1 \mu | d_n |, \qquad 1 \le i \le n-1, \tag{13}$$

$$| x_{n,n} - d_n | \le K_1 \mu | d_n |. \tag{14}$$

proof .By the definition of OMP

$$f_l := f - OMP_l(f, \mathrm{D}) = f - Proj_{span(g_1,\cdots,g_l)}(f).$$

Then we can see

$$\langle f_l, g_i \rangle = 0, 1 \le i \le l. \tag{15}$$

Hence

$$\begin{aligned} f_n &= f_{n-1} - Proj_{span(g_1,\cdots,g_n)}(f_{n-1}) \\ &= f_{n-1} - d_n g_n - Proj_{span(g_1,\cdots,g_n)}(f_{n-1} - d_n g_n). \end{aligned} \tag{16}$$

Using (1) and (15) we have for $h := f_{n-1} - d_n g_n$,

$$| \langle h, g_i \rangle | \le | \langle f_{n-1}, g_i \rangle | + | d_n \langle g_i, g_n \rangle | \le \mu | d_n |, 1 \le i \le n-1,$$

$$| \langle h, g_n \rangle | = | \langle f_{n-1}, g_n \rangle | - | d_n \langle g_n, g_n \rangle | \overset{(11)}{=} | d_n | - | d_n | = 0.$$

Suppose that $x'_{i,n}, 1 \le i \le n,$ satisfy

$$Proj_{span(g_1,\cdots,g_n)}(h) = Proj_{span(g_1,\cdots,g_n)}(f_{n-1} - d_n g_n) = \sum_{i=1}^{n} x'_{i,n} g_i.$$

By Lemma 2

$$\max_{1 \le i \le n} | x'_{i,n} | \le K_1 \max_{1 \le i \le n} | \langle h, g_i \rangle | \le K_1 \mu | d_n |. \tag{17}$$

It follows from (16) and (12) that $f_{n-1} - d_n g_n - \sum_{i=1}^{n} x'_{i,n} g_i \overset{(16)}{=} f_n \overset{(12)}{=} f_{n-1} - \sum_{i=1}^{n} x_{i,n} g_i.$

Then we get the relation between $x_{i,n}$ *and* $x'_{i,n}, 1 \le i \le n,$

$$x_{i,n} = x'_{i,n}, 1 \le i \le n-1, \; x_{n,n} = d_n + x'_{n,n}.$$

This and (17) complete the proof.

The following lemma provides an estimate of $\{| d_n |\}$

Lemma 4. *For any* $1 \le l \le n \le 2m+1$, we have

$$|d_n| \le K_2 |d_l|,$$

where

$$K_2 := \exp(2m\mu \frac{10}{9}) \le \exp(1/9). \quad (18)$$

proof . Using Lemma 3, for $1 \le l \le n \le 2m$, we have

$$\begin{aligned} d_{n+1}| &= |\langle f_n, g_{n+1}\rangle| = |\langle f_{n-1} - \sum_{i=1}^{n} x_{i,n} g_i, g_{n+1}\rangle| \\ &\le |\langle f_{n-1}, g_{n+1}\rangle| + \sum_{i=1}^{n} |x_{i,n}\langle g_i, g_{n+1}\rangle| \\ &\overset{(1)}{\le} |d_n| + \mu(|x_{n,n}| + \sum_{i=1}^{n-1} |x_{i,n}|) \\ &\le |d_n|[1+\mu(1+2m\mu K_1)] \\ &\overset{(6),(10)}{\le} |d_n|[1+(1+\frac{2}{20}\frac{10}{9})\mu] \\ &= |d_n|(1+\frac{10}{9}\mu). \end{aligned}$$

Hence for any $n, 1 \le l \le n \le 2m+1$, we can get

$$|d_n| \le |d_l|(1+\frac{10}{9}\mu)^{n-l} \le |d_l|(1+\frac{2m\frac{10}{9}\mu}{2m})^{2m} \le |d_l| exp(2m\mu\frac{10}{9}) \le K_2 |d_l|.$$

3. Notations

By the definition of the best m-term approximation there exist $a_{j,0} \in \square$, $\psi_j \in \mathrm{D}$, $1 \le j \le m,$ and $\xi_0 \in H$ such that

$$\begin{aligned} f = f_0 &= \sum_{j=1}^{m} a_{j,0}\psi_j + \xi_0, \langle \xi_0, \psi_j\rangle = 0, 1 \le j \le m, \\ \|\xi_0\| &\le (1+\varepsilon)\sigma_m(f, \mathrm{D}) = (1+\varepsilon)\sigma_m(f), \forall\, \varepsilon > 0. \end{aligned} \quad (19)$$

Set

$$L := span(\psi_1, \cdots, \psi_m),\ P_L(.) := Proj_L(.),\ P_L^{\perp}(.) := Proj_L^{\perp}(.),$$
$$\xi_n := P_L^{\perp}(f_n),\ 0 \le n \le 2m.$$

Let the number $a_{j,n}, n \ge 0, 1 \le j \le m,$ satisfy equalities

$$f_n := P_L(f_n) + P_L^{\perp}(f_n) = \sum_{j=1}^{m} a_{j,n}\psi_j + \xi_n. \tag{20}$$

Define

$$T_1 := \{i \in \{1,\ldots,2m\} : g_i \in \{\psi_j\}_{j=1}^m\}, \qquad T_2 := \{1,\ldots,2m\} \setminus T_1.$$

Then, for $n \geq 1$, we let

$$T_2^n := T_2 \cap \{1,\ldots,n\},\ D := \sum_{n\in T_2} d_n^2. \tag{21}$$

4. Main lemmas

Lemma 5. $Let\ 1 \leq i < n \leq 2m, i, n \in T_2$. Then we have

$$|\langle P_L^{\perp}(g_n), g_i\rangle| \leq \frac{19}{18}\mu.$$

$proof$ .Let $P_L(g_n) = \sum_{j=1}^{m} c_j\psi_j$.

Since $n \in T_2$ and

$$g_n \neq \psi_j, |\langle g_n, \psi_j\rangle| \leq \mu, 1 \leq j \leq m,$$

it follows from Lemma 2 that

$$\max_{1\leq j\leq m} |c_j| \leq K_1\mu.$$

Therefore, we have

$$\begin{aligned}
|\langle P_L^{\perp}(g_n), g_i\rangle| &\leq |\langle g_n - P_L(g_n), g_i\rangle| \leq |\langle g_n, g_i\rangle| + |\langle P_L(g_n), g_i\rangle| \\
&\leq \mu + |\langle \sum_{i=1}^{n} c_j\psi_j, g_i\rangle| \leq \mu + (\max_{1\leq j\leq m}|c_j|)(\max_{1\leq j\leq m}|\langle\psi_j, g_i\rangle|) \\
&\leq \mu + (m\mu)K_1\mu \leq \frac{19}{18}\mu.
\end{aligned}$$

□

Lemma 6. $Let\ n \in T_1;$ then we have

$$\|\ \xi_n\|^2 \leq \|\ \xi_{n-1}\|^2 + 0.20D\mu.$$

$proof$. Let

$$t_n := \#T_2^n. \tag{22}$$

If $T_2^n = \varnothing$, then $\xi_n = \xi_{n-1} = \xi_0$ and no prove is needed, so we can assume that $t_n \geq 1$. By Lemma 4

$$|d_n| \le K_2 \min_{i\in T_2^n} |d_i|. \tag{23}$$

On the other hand, by definition (21) and (22), we have

$$(\min_{i\in T_2^n} |d_i|)^2 t_n \le \sum_{i\in T_2^n} d_i^2 \le \sum_{i\in T_2} d_i^2 = D,$$

so we get

$$\min_{i\in T_2^n} |d_i| \le (\frac{D}{t_n})^{1/2}.$$

Combining this with (23), we obtain

$$|d_n| \le K_2 (\frac{D}{t_n})^{1/2},$$
$$d_n^2 t_n \le K_2^2 D. \tag{24}$$

Define

$$h := \sum_{i\in T_2^n} x_{i,n} g_i \overset{n\in T_1}{=} \sum_{i\in T_2^{n-1}} x_{i,n} g_i. \tag{25}$$

According to the definition of ξ_n, we have

$$\xi_n = P_L^\perp(f_n) = P_L^\perp(f_{n-1} - \sum_{i=1}^{n} x_{i,n} g_i) = P_L^\perp(f_{n-1} - \sum_{i\in T_2^n} x_{i,n} g_i) = \xi_{n-1} - P_L^\perp(h).$$

So we get

$$\|\xi_n\|^2 = \|\xi_{n-1} - P_L^\perp(h)\|^2 \le \|\xi_{n-1}\|^2 + 2|\langle \xi_{n-1}, P_L^\perp(h)\rangle| + \|P_L^\perp(h)\|^2$$
$$\le \|\xi_{n-1}\|^2 + 2|\langle \xi_{n-1}, h\rangle| + \|h\|^2. \tag{26}$$

Thus to prove the lemma, we must estimate $|\langle \xi_{n-1}, h\rangle|$ and $\|h\|^2$. Using (15) and (25), we obtain

$$|\langle \xi_{n-1}, h\rangle| = |\langle f_{n-1} - \sum_{j=1}^{m} a_{j,n-1}\psi_j, h\rangle| \overset{(15)}{=} |\langle \sum_{j=1}^{m} a_{j,n-1}\psi_j, h\rangle|$$
$$\overset{(25)}{\le} \sum_{j=1}^{m} |\langle a_{j,n-1}\psi_j, \sum_{i\in T_2^{n-1}} x_{i,n} g_i\rangle|$$
$$\le \sum_{j=1}^{m} |a_{j,n-1}| \sum_{i\in T_2^{n-1}} |x_{i,n} \langle \psi_j, g_i\rangle|. \tag{27}$$

For any $1 \le l \le m$, we have

$$|\langle \sum_{j=1}^{m} a_{j,n-1}\psi_j, \psi_l\rangle| = |\langle \sum_{j=1}^{m} a_{j,n-1}\psi_j + \xi_{n-1}, \psi_l\rangle| \overset{(20)}{=} |\langle f_{n-1}, \psi_l\rangle| \le |d_n|.$$

By Lemma 1, we get

$$\max_{1\le i\le m}|a_{j,n-1}|\le K_1(\max_{1\le l\le m}|\langle\sum_{j=1}^{m}a_{j,n-1}\psi_j,\psi_l\rangle|\le K_1|d_n|. \tag{28}$$

Then we obtain the estimate

$$\sum_{j=1}^{m}|a_{j,n-1}|\le mK_1|d_n|.$$

It follows from (1) and Lemma 3 that for $1\le j\le m$,

$$\sum_{i\in T_2^{n-1}}|x_{i,n}\langle\psi_j,g_i\rangle|\le \mu\#T_2^{n-1}\max_{i\in T_2^{n-1}}|x_{i,n}|\overset{(13)}{\le}\mu\#T_2^{n}K_1\mu|d_n|\overset{(22)}{=}K_1\mu^2|d_n|t_n.$$

Thus, we can continue (27) as

$$|\langle\xi_{n-1},h\rangle|\le\sum_{j=1}^{m}|a_{j,n-1}|\sum_{i\in T_2^{n-1}}|x_{i,n}\langle\psi_j,g_i\rangle|\le m\mu^2K_1^2d_n^2t_n$$
$$\le m\mu^2K_1^2K_2^2D=m\mu K_1^2K_2^2D\mu\le\frac{5}{81}K_2^2D\mu$$

By condition (6), we can suppose that $0<\mu\le 1/40$, so

$$K_1\mu\le\frac{10\mu}{9+10\mu}=\frac{1}{\frac{9}{10\mu}+1}\le\frac{1}{37}. \tag{29}$$

According to Lemma 3, we can write

$$\|h\|^2\le\|\sum_{i\in T_2^{n-1}}x_{i,n}g_i\|^2\le(\max_{i\in T_2^{n-1}}x_{i,n}^2)[\#T_2^{n-1}+(\#T_2^{n-1})^2\mu]$$
$$\le K_1^2\mu^2d_n^2[\#T_2^{n-1}+(\#T_2^{n-1})^2\mu]$$
$$\le K_1^2\mu^2d_n^2(t_n+t_n^2\mu)\le K_1^2\mu^2K_2^2D(1+2m\mu)$$
$$=K_1(K_1\mu)K_2^2(1+2m\mu)D\mu\le\frac{10}{9}\frac{1}{37}1.1K_2^2D\mu=\frac{11}{333}K_2^2D\mu. \tag{30}$$

Now using the estimates for $|\langle\xi_{n-1},h\rangle|$ and $\|h\|^2$, we can continue inequality (26):

$$\|\xi_n\|^2\le\|\xi_{n-1}\|^2+2|\langle\xi_{n-1},h\rangle|+\|h\|^2$$
$$\le\|\xi_{n-1}\|^2+2\frac{5}{81}K_2^2D\mu+\frac{11}{333}K_2^2D\mu$$
$$\le\|\xi_{n-1}\|^2+0.20D\mu.$$

This estimate completes the proof of the lemma.
Now we proceed to the estimate of $\|\xi_n\|$ for $n\in T_2$.

Lemma 7. *Let* $n \in T_2$; then we have

$$\| \xi_n \|^2 \le \| \xi_{n-1} \|^2 - 0.76 d_n^2 .$$

proof . Just as in the proof of Lemma 6, we use the element

$$h := \sum_{i \in T_2^{n-1}} x_{i,n} g_i .$$

Set

$$\xi'_n := P_L^{\perp}(f_{n-1} - x_{n,n} g_n).$$

Then we can write

$$\begin{aligned}
\xi_n &= P_L^{\perp}(f_n) = P_L^{\perp}(f_{n-1} - \sum_{i=1}^{n} x_{i,n} g_i) \\
&= P_L^{\perp}(f_{n-1} - x_{n,n} g_n) - P_L^{\perp}(\sum_{i=1}^{n-1} x_{i,n} g_i) \\
&= \xi'_n - P_L^{\perp}(\sum_{i \in T_2^{n-1}} x_{i,n} g_i) = \xi'_n - P_L^{\perp}(h),
\end{aligned}$$

$$\| \xi_n \|^2 = \| \xi'_n \|^2 - 2\langle \xi'_n, P_L^{\perp}(h) \rangle + \| P_L^{\perp}(h) \|^2 \le \| \xi'_n \|^2 + 2 | \langle \xi'_n, h \rangle | + \| h \|^2 . \quad (31)$$

Therefore, to prove the lemma it suffices to obtain upper bounds for $\| \xi'_n \|^2$, $| \langle \xi'_n, h \rangle |$, $\| h \|^2$. To estimate $\| h \|^2$, we can use inequality (31) from Lemma 6

$$\begin{aligned}
\| h \|^2 &\le K_1^2 \mu^2 d_n^2 [\#T_2^{n-1} + (\#T_2^{n-1})^2 \mu] \\
&\le K_1^2 \mu^2 d_n^2 [2m + (2m)^2 \mu] = K_1 (K_1 \mu) 2m\mu (1 + 2m\mu) d_n^2 \\
&\le \frac{10}{9} \frac{1}{10} \frac{1}{37} 1.1 d_n^2 = \frac{11}{3330} d_n^2 \le 0.004 d_n^2 . \quad (32)
\end{aligned}$$

Then we proceed to the estimate of $\| \xi'_n \|^2$.
Using (20), (28) and the inclusion $n \in T_2$, we can write

$$\begin{aligned}
&| \langle \xi_{n-1}, g_n \rangle - d_n | = | \langle f_{n-1} - \sum_{j=1}^{m} a_{j,n-1} \psi_j, g_n \rangle - d_n | \\
&= | \langle f_{n-1}, g_n \rangle - \sum_{j=1}^{m} a_{j,n-1} \langle \psi_j, g_n \rangle - d_n | = | \sum_{j=1}^{m} a_{j,n-1} \langle \psi_j, g_n \rangle | \\
&\le (\max_{1 \le j \le m} | a_{j,n-1} |) m \max_{1 \le j \le m} | \langle \psi_j, g_n \rangle | \le K_1 | d_n | m\mu . \quad (33)
\end{aligned}$$

Then using Lemma 3 and (29),(6), we obtain the estimate

$$\begin{aligned}
2x_{n,n} \langle \xi_{n-1}, g_n \rangle &= 2[d_n + (x_{n,n} - d_n)][d_n + (\langle \xi_{n-1}, g_n \rangle - d_n)] \\
&\ge 2(| d_n | - K_1 \mu | d_n |)(| d_n | - K_1 | d_n | m\mu)
\end{aligned}$$

$$= 2|d_n|^2(1-K_1\mu)(1-K_1 m\mu)$$
$$\geq 2|d_n|^2(1-\frac{1}{37})(1-\frac{10}{9}\frac{1}{20}) = \frac{68}{37}d_n^2$$

and finally, obtain

$$\|\xi_n'\|^2 = \|P_L^{\perp}(f_{n-1}-x_{n,n}g_n)\|^2 = \|\xi_{n-1}-x_{n,n}P_L^{\perp}(g_n)\|^2$$
$$= \|\xi_{n-1}\|^2 - 2x_{n,n}\langle\xi_{n-1}, P_L^{\perp}(g_n)\rangle + x_{n,n}^2\|P_L^{\perp}(g_n)\|^2$$
$$\leq \|\xi_{n-1}\|^2 - 2x_{n,n}\langle\xi_{n-1}, g_n\rangle + x_{n,n}^2 \leq \|\xi_{n-1}\|^2 - \frac{68}{37}d_n^2 + x_{n,n}^2$$
$$\leq \|\xi_{n-1}\|^2 - \frac{68}{37}d_n^2 + (|d_n| + K_1\mu|d_n|)^2$$
$$= \|\xi_{n-1}\|^2 + d_n^2[(1+K_1\mu)^2 - \frac{68}{37}] \leq \|\xi_{n-1}\|^2 + d_n^2[(1+\frac{1}{37})^2 - \frac{68}{37}]$$
$$= \|\xi_{n-1}\|^2 - \frac{1072}{1369}d_n^2 \leq \|\xi_{n-1}\|^2 - 0.78d_n^2. \quad (34)$$

It remains to estimate $|\langle\xi_n', h\rangle|$. Equalities (15) imply that

$$\langle f_{n-1}, h\rangle = 0.$$

We have

$$|\langle\xi_n', h\rangle| = |\langle P_L^{\perp}(f_{n-1}) - x_{n,n}P_L^{\perp}(g_n), h\rangle| = |\langle\xi_{n-1} - x_{n,n}P_L^{\perp}(g_n), h\rangle|$$
$$= |\langle f_{n-1} - \sum_{j=1}^{m} a_{j,n-1}\psi_j - x_{n,n}P_L^{\perp}(g_n), h\rangle|$$
$$\leq \sum_{j=1}^{m}|\langle a_{j,n-1}\psi_j, h\rangle| + |x_{n,n}\langle P_L^{\perp}(g_n), h\rangle|$$
$$\leq \sum_{j=1}^{m}|a_{j,n-1}|\sum_{i\in T_2^{n-1}}|\langle\psi_j, x_{i,n}g_i\rangle| + \sum_{i\in T_2^{n-1}}|x_{n,n}x_{i,n}\langle P_L^{\perp}(g_n), g_i\rangle|$$
$$=: A + B. \quad (35)$$

Let us estimate the summands A and B separately.

$$A \leq \sum_{j=1}^{m}|a_{j,n-1}|\sum_{i\in T_2^{n-1}}|x_{i,n}||\langle\psi_j, g_i\rangle|$$
$$\leq \max_{1\leq j\leq m}|a_{j,n-1}|\max_{i\in T_2^{n-1}}|x_{i,n}|\sum_{j=1}^{m}\sum_{i\in T_2^{n-1}}\mu$$
$$\leq K_1|d_n|K_1\mu|d_n|m(n-1)\mu \leq |d_n|^2\mu^2 m(2m-1)K_1^2$$
$$\leq [2K_1^2\mu^2 m^2 - K_1(K_1\mu)m\mu]|d_n|^2 \leq [2(\frac{10}{9})^2(\frac{1}{20})^2 - \frac{10}{9}\frac{1}{37}\frac{1}{20}]d_n^2$$
$$= \frac{42}{8991}d_n^2 \leq 0.005d_n^2.$$

Using Lemma 5 and Lemma 3, we find that

$$B \le |x_{n,n}| \sum_{i\in T_2^{n-1}} |x_{i,n} \langle P_L^{\perp}(g_n), g_i\rangle| \le |x_{n,n}| \sum_{i\in T_2^{n-1}} |x_{i,n}| \frac{19}{18}\mu$$
$$\le (|d_n| + K_1\mu |d_n|)(nK_1\mu |d_n|)\frac{19}{18}\mu \le d_n^2(1+K_1\mu)2m\mu K_1\mu\frac{19}{18}$$
$$\le (1+\frac{1}{37})\frac{1}{37}\frac{19}{18}\frac{1}{10}d_n^2 = \frac{361}{123210}d_n^2 \le 0.003d_n^2.$$

Substituting estimate for A and B into (35), we obtain

$$|\langle \xi_n', h\rangle| \le A + B \le 0.008d_n^2. \quad (36)$$

Using inequalities (33), (34) and (36), we can continue estimate (31) and complete the proof:

$$\|\xi_n\|^2 \le \|\xi'\|^2 + 2|\langle \xi_n', h\rangle| + \|h\|^2$$
$$\le \|\xi_{n-1}\|^2 - 0.78d_n^2 + 2(0.008d_n^2) + 0.004d_n^2$$
$$\le \|\xi_{n-1}\|^2 - 0.76d_n^2.$$

Lemma 8. The following estimates hold

$$D^{1/2} \le \sqrt{\frac{4}{3}}(1+\varepsilon)\sigma_m(f), \forall\, \varepsilon > 0, \quad (37)$$

$$\|\xi_{2m}\| \le \|\xi_0\|. \quad (38)$$

proof. Using Lemma 6 and Lemma 7, we obtain

$$((1+\varepsilon)\sigma_m(f))^2 \ge \|\xi_0\|^2 \ge \|\xi_0\|^2 - \|\xi_{2m}\|^2 = \sum_{n=1}^{2m}(\|\xi_{n-1}\|^2 - \|\xi_n\|^2)$$
$$= \sum_{n\in T_1}(\|\xi_{n-1}\|^2 - \|\xi_n\|^2) + \sum_{n\in T_2}(\|\xi_{n-1}\|^2 - \|\xi_n\|^2)$$
$$\ge \#T_1(-0.20D\mu) + \sum_{n\in T_2} 0.76d_n^2$$
$$\ge m(-0.20D\mu) + 0.76D \ge 0.75D > 0.$$

Hence

$$D^{1/2} \le (1+\varepsilon)(0.75^{-1/2})\sigma_m(f) \le \sqrt{\frac{4}{3}}(1+\varepsilon)\sigma_m(f).$$

5. Proof of theorem 2

First we estimate $\|P_l(f_{2m})\| = \|\sum_{j=1}^{m} a_{j,2m}\psi_j\|$, using (28) and Lemma 4, we can write for any $l, 1 \le l \le 2m,$

$$\max_{1\le j\le m} | a_{j,2m} |\le K_1 | d_{2m+1} |\le K_1K_2 | d_l |. \quad (39)$$

Since $\#T_2 \ge m$, using definition (21), we obtain

$$\sum_{j=1}^{m} a_{j,2m}^2 \le m \max_{1\le j\le m} a_{j,2m}^2 \le \sum_{l\in T_2} (K_1K_2 | d_l |)^2 = (K_1K_2)^2 D.$$

Applying a well-known inequality (see, for example, lemma 2.1 from [4]) and substituting the values of K_1 and K_2 (see (10) and (18)), we will find the estimates

$$\| \sum_{j=1}^{m} a_{j,2m}\psi_j \|^2 \le (\sum_{j=1}^{m} a_{j,2m}^2)(1+m\mu) \le (K_1K_2)^2 D1.05 \le 1.62D. \quad (40)$$

Then using lemma 8 and (40), for $\forall\ \varepsilon > 0$, we obtain the.

$$\begin{aligned} &\| f_{2m} \| = \| \sum_{j=1}^{m} a_{j,2m}\psi_j + \xi_{2m} \| \le \| \sum_{j=1}^{m} a_{j,2m}\psi_j \| + \| \xi_0 \| \\ &\le (1.62D)^{1/2} + (1+\varepsilon)\sigma_m(f) \le (1.62)^{1/2}\sqrt{\frac{4}{3}}(1+\varepsilon)\sigma_m(f) + (1+\varepsilon)\sigma_m(f) \\ &\le 2.47(1+\varepsilon)\sigma_m(f). \end{aligned}$$

This complete the proof of Theorem 3.

This work is supported by National Nature Science Foundation of China under Grant 10971251, 11271199 and 11101220.

References

[1] D Donoho, M Elad, VN Temlyakov. Stable recovery of sparse over complete representation in the presence of noise. *IEEE Trans. Inform*. Theory 2006; 52(1): 6-18.
[2] D Donoho, M Elad, VN Temlyakov. On Lebesgue-type inequalities for greedy approximation. *Approx. Theroy*. 2007; 147(2): 185-195.
[3] AC Gilbert, M Muthukrishnan, J Strauss. Approximation of functions over redundant dictionaries using coherence, in: Proc. 14th Annu. *ACM-SIAM Symp. Discrete Algorithms.* 2003; 234-252.
[4] R Gilbonval, M Nielsen. On the strong uniqueness of highly sparse expansions from redundant dictionaries, in: *Proc. Int Conf. Independent Component Anal*. (ICA'04), (2004).
[5] ED Livshits. On greedy algorithms with bounded cumulative coherence, Math. Notes 2010; 87(5): 792-795.
[6] VN Temlyakov. Greedy Algorithms and m-term approximation with regard to redundant dictionaries. *Approx. Theory*. 1999; 98: 117-145.
[7] VN Temlyakov, P Zheltov. On performance of greedy algorithms. *Approx. Theory*. 2011; 163(9): 1134-1145.
[8] JA Tropp. Greedy is good: algorithmic results for sparse approximation. *IEEE Trans. Inform. Theory* 2004; 50(10): 2231-2242.
[9] Qinghua Wu, Hanmin Liu, Yuxin Sun, Fang Xie, Jin Zhang, Xuesong Yan. Research of Function Optimization Algorithm. TELKOMNIKA. 2012; 10(4): 858-863.
[10] ED Livshits. On the efficiency of the Orthogonal Matching Pursuit compressed sensing .Mat. Sb. 2012; 203: 33-44.
[11] Windra Swastika, Hideaki Haneishi. Compressed Sensing for Thoracic MRI with Partial Random Circulant Matrices. *TELKOMNIKA*. 2012; 10(1): 147-154.
[12] ED Livshitz. On the optimality of Orthogonal Greedy Algorithm for-coherent dictionaries. *Approx. Theory 2012*; 164: 668-681.