New Coherence and RIP Analysis for Weak Orthogonal Matching Pursuit

Mingrui Yang, *Member, IEEE,* and Frank de Hoog

Abstract

In this paper we define a new coherence index, named the *global 2-coherence*, of a given dictionary and study its relationship with the traditional mutual coherence and the restricted isometry constant. By exploring this relationship, we obtain more general results on sparse signal reconstruction using greedy algorithms in the compressive sensing (CS) framework. In particular, we obtain an improved bound over the best known results on the restricted isometry constant for successful recovery of sparse signals using orthogonal matching pursuit (OMP).

Index Terms

Compressive sensing, mutual coherence, global 2-coherence, restricted isometry property, weak orthogonal matching pursuit (WOMP), orthogonal matching pursuit (OMP)

I. INTRODUCTION

Compressive sensing (CS) [1], [2] is a newly developed and fast growing field of research. It provides a new sampling scheme that breaks the traditional Shannon-Nyquist sampling rate [3] given that the signal of interest is sparse in a certain basis or tight frame. More specifically, for a vector $a \in \mathbb{R}^d$, let $\|a\|_0$ denote the ℓ_0 "norm" of a, which counts the number of nonzero entries of a. We say a is *k-sparse* if $\|a\|_0 \leq k$. CS has established conditions for finding the unique sparse solution of the following ℓ_0 minimization problem

$$\min_a \|a\|_0 \text{ subject to } f = \Phi a, \tag{1}$$

where $\Phi \in \mathbb{R}^{n \times d}$ ($n \ll d$) and $f \in \mathbb{R}^n$. To ensure that the k-sparse solution is unique, we need the following restricted isometry property introduced by Candes and Tao in [4].

Definition I.1 (Restricted Isometry Property (RIP))**.** *A matrix Φ satisfies the* restricted isometry property *of order k with the* restricted isometry constant (RIC) *δ_k if $\delta_k \in (0,1)$ is the smallest constant such that*

$$(1 - \delta_k)\|a\|_2^2 \leq \|\Phi a\|_2^2 \leq (1 + \delta_k)\|a\|_2^2 \tag{2}$$

holds for all k-sparse signal a.

Mingrui Yang is with CSIRO Computational Informatics, Pullenvale, QLD 4069 Australia (email: mingrui.yang@csiro.au).

Frank de Hoog is with CSIRO Computational Informatics, Acton, ACT 2061 Australia (email: frank.dehoog@csiro.au).

It has been shown in [4] that if $\delta_{2k} < 1$, then the ℓ_0 minimization problem (1) has a unique k-sparse solution. However, solving an ℓ_0 minimization problem is in general NP-hard. One solution to this problem is to relax the ℓ_0 "norm" to the ℓ_1 norm. Candes has shown in [5] if $\delta_{2k} < \sqrt{2} - 1$, then ℓ_1 minimization is equivalent to ℓ_0 minimization. Another alternative is to use heuristic greedy algorithms to approximate the solution of the ℓ_0 minimization problem . Orthogonal matching pursuit (OMP) is one of the simplest and most popular algorithms of this type. For the analysis of greedy algorithms, the metric chosen for the sensing matrix are usually coherence indices rather than the RIC.

For simplicity, from now on, we always assume that the columns of the matrix (dictionary) Φ are normalized such that for any column $\phi \in \Phi$, $\|\phi\|_2 = 1$.

Definition I.2. *The* mutual coherence $M(\Phi)$ *of a matrix* Φ *is defined by*

$$M(\Phi) := \max_{\substack{\phi_i, \phi_j \in \Phi \\ i \neq j}} |\langle \phi_i, \phi_j \rangle|, \tag{3}$$

where $\langle \cdot, \cdot \rangle$ *represents the usual inner product.*

It has been shown that if $(2k - 1)M < 1$, then OMP can recover every k-sparse signal exactly in k iterations [6]. Recently, researchers have started to investigate the performance of OMP using RIP. Davenport and Wakin [7] have proved that $\delta_{k+1} < \frac{1}{3\sqrt{k}}$ is sufficient for OMP to recover any k-sparse signal in k iterations. Mo and Shen [8] improve the bound to $\delta_{k+1} < \frac{1}{1+\sqrt{k}}$. They also give an example that OMP fails to recover a k-sparse signal in k steps when $\delta_{k+1} = \frac{1}{\sqrt{k}}$. This leaves a question if their bound can be further improved.

It is then natural to examine the relationship between the mutual coherence M and the RIC δ_k, since the bound for M is already sharp. However, approaching this directly was not fruitful and this motivated us to define a new coherence index, namely the *global 2-coherence*, and establish a bridge connecting the mutual coherence, the global 2-coherence, and the RIC. Then by using this connection, we analyze the performance of weak orthogonal matching pursuit (WOMP), a weak version of OMP. In particular, we extend the results given in [9] to show that $\delta_k + \sqrt{k}\delta_{k+1} < 1$ is sufficient for OMP to recover any k-sparse signal in k iterations, which provides an improved bound over the best known result given in [8] and confirms that it is not yet optimal. As mentioned above, the results presented in this paper is an extension of [9], where we introduced a new algorithm to CS, called orthogonal matching pursuit with thresholding (OMPT), and showed its reconstruction stability and robustness.

II. GLOBAL 2-COHERENCE

We first define a new coherence index, the *global 2-coherence*, $\nu_k(\Phi)$ for a given dictionary Φ. Then based on this new coherence index, we establish the connections among the coherence indices and the RIC δ_k.

Definition II.1. *Denote by* $[d]$ *the index set* $\{1, 2, \ldots, d\}$. *The global 2-coherence of a dictionary* $\Phi \in \mathbb{R}^{n \times d}$ *is defined as*

$$\nu_k(\Phi) := \max_{i \in [d]} \max_{\substack{\Lambda \subseteq [d] \setminus \{i\} \\ |\Lambda| \leq k}} \left(\sum_{j \in \Lambda} \langle \phi_i, \phi_j \rangle^2 \right)^{1/2}, \tag{4}$$

where ϕ_i, ϕ_j are columns from the dictionary Φ.

The global 2-coherence $\nu_k(\Phi)$ defined above is a generalization of the mutual coherence defined in Definition I.2 and the coherence indices defined in [10], [11]. In particular, when $k = 1$, ν_1 is exactly the mutual coherence.

The following lemma describes the relations among the mutual coherence M, the 2-coherence ν_k, and the restricted isometry constant δ_k.

Lemma II.2. *For $k > 1$, we have*

$$M \leq \nu_{k-1} \leq \delta_k \leq \sqrt{k-1}\nu_{k-1} \leq (k-1)M. \tag{5}$$

The next lemma is needed to proceed to our main results.

Lemma II.3. *Let $\Lambda \subset [d]$ with $|\Lambda| = k$. Let $f = \Phi a + w$ with $\mathrm{supp}(a) = \Lambda$ and $\|w\|_2 \leq \epsilon$. In addition, assume that there exits $\Omega \subseteq \Lambda$ with $|\Omega| = m$, such that*

$$\langle \Phi a, \phi_i \rangle = 0, \ for \ i \in \Lambda \setminus \Omega.$$

Then

$$\max_{i \in [d] \setminus \Lambda} |\langle f, \phi_i \rangle| \leq \nu_k \|a\|_2 + \epsilon,$$

$$\max_{i \in \Lambda} |\langle f, \phi_i \rangle| \geq \frac{\sqrt{1-\delta_k}}{\sqrt{m}} \|\Phi a\|_2 - \epsilon.$$

III. MAIN RESULTS

We first begin with a well known greedy algorithm, the weak orthogonal matching pursuit (WOMP), which was defined in [12]. Here we present a simple version in Algorithm 1 where the weak parameter ρ is a constant for each iteration. Notice that when $\rho = 1$, WOMP becomes standard OMP.

Let us consider the case where a sparse signal is contaminated by a perturbation. Specifically, let $\Lambda \subset [d]$ with $|\Lambda| = k$. We consider a signal $f = \Phi a + w$, where $a \in \mathbb{R}^d$ with $\mathrm{supp}(a) = \Lambda$ and $\|w\|_2 \leq \epsilon$.

Theorem III.1. *Denote by a_{min} the nonzero entry of a with the least magnitude, and $\hat{a}_{womp}$ the recovered representation of f in Φ by WOMP after k iterations. If*

$$\sqrt{k}\nu_k < \rho(1 - \delta_k) \tag{6}$$

and the noise level obeys

$$\epsilon < \frac{\rho(1 - \delta_k) - \sqrt{k}\nu_k}{1 + \rho} |a_{min}|, \tag{7}$$

then

a) $\hat{a}_{womp}$ has the correct sparsity pattern

$$\mathrm{supp}(\hat{a}_{womp}) = \mathrm{supp}(a);$$

Algorithm 1 Weak Orthogonal Matching Pursuit (WOMP)

1: **Input:** weak parameter $\rho \in (0, 1]$, dictionary Φ, signal f, and the noise level ϵ.

2: **Initialization:** $r_0 := f$, $x_0 := 0$, $\Lambda_0 := \emptyset$, $s := 0$.

3: **while** $\|r_s\|_2 > \epsilon$ **do**

4:     Find an index i such that

$$|\langle r_s, \phi_i \rangle| \geq \rho \cdot \max_{\phi} |\langle r_s, \phi \rangle|$$

    where ϕ is any column of Φ;

5:     Update the support:

$$\Lambda_{s+1} = \Lambda_s \cup \{i\};$$

6:     Update the estimate:

$$x_{s+1} = \arg\min_{z} \|f - \Phi_{\Lambda_{s+1}} z\|_2;$$

7:     Update the residual:

$$r_{s+1} = f - \Phi_{\Lambda_{s+1}} x_{s+1};$$

8:     $s = s + 1$;

9: **end while**

10: **Output:** If the algorithm is stopped after k iterations, then the output estimate $\hat{a}$ of a is $\hat{a}_{\Lambda_k} = x_k$ and $\hat{a}_{\Lambda_k^C} = 0$.

---

b) $\hat{a}_{womp}$ approximates the ideal noiseless representation

$$\|\hat{a}_{womp} - a\|_2^2 \leq \frac{\epsilon^2}{1 - \delta_k}. \tag{8}$$

From Lemma II.2, it follows that

Corollary III.2. *Let $f = \Phi a$ with $\|a\|_0 = k$. If one of the following conditions is satisfied,*

a) $\sqrt{k}\delta_{k+1} < \rho(1 - \delta_k)$,

b) $\sqrt{k}\nu_k < \rho(1 - \nu_{k-1}\sqrt{k-1})$,

c) $kM < \rho(1 - (k-1)M)$,

then, a is the unique sparsest representation of f and moreover, WOMP recovers a exactly in k iterations.

The performance of WOMP decreases as ρ decreases. Now if we set $\rho = 1$ in WOMP, then we obtain immediately the following corollary for OMP.

Corollary III.3. *Let $f = \Phi\alpha$ with $\|a\|_0 = k$. If*

$$\delta_k + \sqrt{k}\delta_{k+1} < 1, \tag{9}$$

then, a is the unique sparsest representation of f and moreover, OMP recovers a exactly in k iterations.

Remark III.4. *The condition in* (9) *gives an improved bound on the restricted isometry constant compared to the bound obtained in [8] for OMP for successful recovery after k iterations, where the bound was $\delta_{k+1} < \frac{1}{\sqrt{k+1}}$.*

IV. CONCLUSION

In this paper, we have introduced a new generalized coherence index, the *global 2-coherence*, and established two connections among the mutual coherence, the global 2-coherence, and the restricted isometry constant. Based on these relations, we analyzed the performance of WOMP as well as OMP for their recovery ability of sparse representations in both ideal noiseless and noisy cases. In particular, for the noiseless case, we showed an improved bound over the best known results on the restricted isometry constant for successful recovery using OMP.

APPENDIX

Proof of Lemma II.2: It is easy to show that ν_k increases with k while $\frac{\nu_k}{\sqrt{k}}$ decreases with k. Therefore, the first and the last relations follow immediately.

We now prove the second inequality.

$$\nu_{k-1}(\Phi) = \max_{i\in[d]} \max_{\substack{\Lambda\subseteq[d]\setminus\{i\} \\ |\Lambda|\leq k-1}} \left(\sum_{j\in\Lambda} \langle\phi_i,\phi_j\rangle^2\right)^{\frac{1}{2}}$$

$$= \max_{\substack{\Lambda\subseteq[d] \\ |\Lambda|\leq k}} \max_{i\in\Lambda} \left(\sum_{j\in\Lambda\setminus\{i\}} \langle\phi_i,\phi_j\rangle^2\right)^{\frac{1}{2}}$$

$$= \max_{\substack{\Lambda\subseteq[d] \\ |\Lambda|\leq k}} \|\Phi_\Lambda^T\Phi_\Lambda - I\|_{\infty,2},$$

where $\Phi_\Lambda \in \mathbb{R}^{n\times|\Lambda|}$ is a submatrix of Φ with columns indexed in Λ.

On the other hand, according to Proposition 2.5 in [13], one has

$$\delta_k = \max_{\substack{\Lambda\subseteq[d] \\ |\Lambda|\leq k}} \|\Phi_\Lambda^T\Phi_\Lambda - I\|_{2,2}$$

$$\geq \max_{\substack{\Lambda\subseteq[d] \\ |\Lambda|\leq k}} \|\Phi_\Lambda^T\Phi_\Lambda - I\|_{\infty,2}$$

$$= \nu_{k-1}(\Phi),$$

which completes the proof for the second inequality.

Next we prove the third inequality. Consider the Gram matrix $G = \Phi_\Lambda^T \Phi_\Lambda$, where its entries $g_{ij} = \langle \phi_i, \phi_j \rangle$. Clearly its diagonal entries $g_{ii} = 1$. Then by the Gershgorin Circle Theorem, each eigenvalue λ of G is in at least one of the disks $\{z : |z - 1| \leq R_i\}$, where $R_i = \sum_{\substack{j \in \Lambda \\ j \neq i}} |g_{ij}|$. Equivalently, we have

$$1 - R_i \leq \lambda \leq 1 + R_i$$

for some i. Therefore,

$$
\begin{aligned}
\delta_k &\leq \max_i R_i = \max_i \sum_{\substack{j \in \Lambda \\ j \neq i}} |g_{ij}| \\
&\leq \max_i \sqrt{k-1} \Big(\sum_{\substack{j \in \Lambda \\ j \neq i}} |g_{ij}|^2 \Big)^{\frac{1}{2}} \\
&\leq \sqrt{k-1}\, \nu_{k-1}.
\end{aligned}
$$

∎

Proof of Lemma II.3: For $i \in [d] \setminus \Lambda$, we have

$$
\begin{aligned}
|\langle f, \phi_i \rangle| &= |\langle \Phi a + w, \phi_i \rangle| \\
&\leq |\langle \Phi a, \phi_i \rangle| + |\langle w, \phi_i \rangle| \\
&\leq \nu_k \|a\|_2 + \|w\|_2 \|\phi_i\|_2 \\
&\leq \nu_k \|a\|_2 + \epsilon.
\end{aligned}
$$

Taking maximum on both sides completes the proof of the first inequality.

Now for $i \in \Lambda$,

$$
\begin{aligned}
\max_{i \in \Lambda} |\langle f, \phi_i \rangle| &= \max_{i \in \Lambda} |\langle \Phi a + w, \phi_i \rangle| \\
&\geq \max_{i \in \Lambda} |\langle \Phi a, \phi_i \rangle| - \max_{i \in \Lambda} |\langle w, \phi_i \rangle| \\
&\geq \frac{\sqrt{1 - \delta_k}}{\sqrt{m}} \|\Phi a\|_2 - \max_{i \in \Lambda} \|w\|_2 \|\phi_i\|_2 \\
&\geq \frac{\sqrt{1 - \delta_k}}{\sqrt{m}} \|\Phi a\|_2 - \epsilon.
\end{aligned}
$$

This completes the proof of the second inequality. ∎

Proof of Theorem III.1: First, we show that WOMP recovers the correct support of a.

We start with the first iteration. Note that $r_0 = f$. We need to show

$$\max_{i \in [d] \setminus \Lambda} |\langle f, \phi_i \rangle| < \rho \max_{i \in \Lambda} |\langle f, \phi_i \rangle|. \tag{10}$$

By Lemma II.3, we have

$$\max_{i \in [d] \setminus \Lambda} |\langle f, \phi_i \rangle| \leq \nu_k \|a\|_2 + \epsilon, \tag{11}$$

and

$$\max_{i \in \Lambda} |\langle f, \phi_i \rangle| \geq \frac{1 - \delta_k}{\sqrt{k}} \|\Phi a\|_2 - \epsilon$$
$$\geq \frac{1 - \delta_k}{\sqrt{k}} \|a\|_2 - \epsilon. \tag{12}$$

Now since $\|a\|_2 \geq \sqrt{k}|a_{min}|$, by imposing conditions (6) and (7), we get

$$\nu_k \|a\|_2 + \epsilon < \rho \left(\frac{1 - \delta_k}{\sqrt{k}} \|a\|_2 - \epsilon \right),$$

and relation (10) follows from the two bounds (11) and (12). Hence, WOMP only selects one atom from $\{\phi_i\}_{i \in \Lambda}$ in the first iteration.

Now we argue that by repeatedly applying the above procedure, we are able to correctly recover the support of a. In fact, we have for the s-th iteration

$$r_s = f - P_{\Lambda_s}(f)$$
$$= \Phi a + w - (P_{\Lambda_s}(\Phi a) + P_{\Lambda_s}(w))$$
$$= (I - P_{\Lambda_s})\Phi a + (I - P_{\Lambda_s})w$$
$$= \Phi a_s + w_s$$

where

$$\Phi a_s = (I - P_{\Lambda_s})\Phi a$$

and

$$w_s = (I - P_{\Lambda_s})w.$$

Therefore, $\langle \Phi a_s, \phi_i \rangle = 0$ for $i \in \Lambda_s$. Note that $(k - s)$ components of a_s are the same as that of a. Then the result follows from the inequality $\|a_s\|_2 \geq \sqrt{k - s}|a_{min}|$ and Lemma II.3 for s-th iteration. In addition, the orthogonal projection step guarantees that the procedure will not repeat the atoms already chosen in previous iterations. Therefore, the correct support of the noiseless representation a can be recovered exactly after k iterations.

Next, by following the idea of the proof of Theorem 5.1 in [6] and using the relation $\sigma_{min} \geq 1 - \delta_k$, where σ_{min} denotes the smallest singular value of Φ, we are able to prove the error bound (8).

∎

REFERENCES

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

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006. [Online]. Available: http://dx.doi.org/10.1109/TIT.2006.871582

[3] C. Shannon, "Communication in the presence of noise," *Proceedings of the Institute of Radio Engineers (IRE)*, vol. 37, pp. 10–21, 1949.

[4] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. on Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[5] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9-10, pp. 589 – 592, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1631073X08000964

[6] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 6–18, 2006. [Online]. Available: http://dx.doi.org/10.1109/TIT.2005.860430

[7] M. Davenport and M. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4395–4401, 2010.

[8] Q. Mo and Y. Shen, "A remark on the restricted isometry property in orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 58, no. 6, pp. 3654–3656, 2012.

[9] M. Yang and F. de Hoog, "Orthogonal matching pursuit with thresholding and its application in compressed sensing," *in preparation*.

[10] J. Tropp, "Greed is good: algorithmic results for sparse approximation," *Information Theory, IEEE Transactions on*, vol. 50, no. 10, pp. 2231–2242, 2004.

[11] Y. Eldar and H. Rauhut, "Average case analysis of multichannel sparse recovery using convex relaxation," *Information Theory, IEEE Transactions on*, vol. 56, no. 1, pp. 505–519, 2010.

[12] V. Temlyakov, "Weak greedy algorithms," *Advances in Computational Mathematics*, vol. 12, no. 2-3, pp. 213–227, 2000.

[13] H. Rauhut, "Compressive sensing and structured random matrices," in *Theoretical Foundations and Numerical Methods for Sparse Recovery*, ser. Radon Series Comp. Appl. Math., M. Fornasier, Ed.   deGruyter, 2010, vol. 9, pp. 1–92.