A ROBUST ACS CONJECTURE

CHRISTOPHER J. HILLAR AND DARREN RHEA

1. Introduction

Let $S_{p,k}$ denote the set of all k-sparse vectors in $\mathbb{R}^p$ (k-sparse means at most k nonzero components).

Definition 1. *We say that $A \in \mathbb{R}^{m\times p}$ has $(2k,\delta)$-lower-RIP when*

$$\text{for all } a_1, a_2 \in S_{p,k},\ \|A(a_1 - a_2)\| \geq \sqrt{1-\delta}\|a_1 - a_2\|. \tag{1}$$

Conjecture 1. *Suppose $\varepsilon, \delta \in (0,1)$. Suppose $A, B \in \mathbb{R}^{m\times p}$ where A has has $(2k,\delta)$-lower-RIP. Suppose there is a function $f : S_{p,k} \to S_{p,k}$ satisfying the almost recovery condition*

$$\text{for all } a \in S_{p,k} \text{ with } ||a|| \leq 1,\ \ ||Aa - Bf(a)|| \leq \varepsilon. \tag{2}$$

Then there exists a permutation matrix $P \in \mathbb{R}^{p\times p}$ and a diagonal matrix $D \in \mathbb{R}^{p\times p}$ such that

$$\text{for all } k\text{-sparse } a \text{ with } ||a|| \leq 1,\ \ ||f(a) - PDa|| \leq \frac{2}{\sqrt{1-\delta}} \cdot \varepsilon. \tag{3}$$

There are some natural norms relevant for comparing dictionaries. For instance, here is the definition of what you might call the k-restricted $\frac{\text{Euclid}}{\text{Euclid}}$ norm:

Definition 2. *The set of sparse $a \in S_{p,k} \subseteq \mathbb{R}^p$ such that $\|a\|_{Euclid} = 1$ is compact. Therefore for matrices $\underset{m\times p}{M}$ we can define the k-restricted-Euclidean-over-Euclidean sorta-matrix norm via*

$$\|M\|_{restricted} := \max_{\substack{a\in S_{p,k} \\ \|a\|_{Euclid}=1}} \|Ma\|_{Euclid}. \tag{4}$$

This is a "vector" norm because

- $\|M\|_{restricted} = 0$ *if and only if* $M = 0$.
- $\|cM\|_{restricted} = |c| \cdot \|M\|_{restricted}$.
- $\|M + N\|_{restricted} \leq \|M\|_{restricted} + \|N\|_{restricted}$.
- *the claim that* $\|MN\|_{restricted} \leq \|M\|_{restricted} \cdot \|N\|_{restricted}$ *is not generally true, thus sorta-matrix norm.*

It is not at all clear that the Euclidean norm should be the "denominator" norm. Maybe the L^1 norm would be more appropriate for the denominator, although for small k the difference wouldn't be much.

Lemma 1. *Suppose $k \in \mathbb{Z}_{\geq 1}$. Suppose $\delta \in (0,1)$. Suppose that $\underset{m\times p}{A}$ satisfies $(2k,\delta)$-lower-RIP (Think of this A as the actually correct sparse dictionary.) Suppose that $\underset{m\times p}{B}$ and $\underset{p\times p}{P}$ is a permutation matrix and $\underset{p\times p}{D}$ is a diagonal matrix and and $\underset{p\times p}{J}$ is a diagonal matrix whose diagonal elements are ± 1 such that*

$$\|J - D\|_{spectral} < \varepsilon \tag{5}$$

and

$$\|D^{-1}\|_{spect} < \frac{1}{1-\varepsilon}. \tag{6}$$

Suppose $\underset{m\times p}{C} := BPD$ has $\|A - C\|_{restricted} \leq \varepsilon$ under the Euclidean over Euclidean k-restricted sorta-norm. (Think B is estimated dictionary under noise.) Suppose that vector $y = Aa$ with $a \in S_{p,k}$ (Think y's true explanation is coefficients a over dictionary A.) Suppose that $\|Bb - y\| \leq \eta$ for some $b \in S_{p,k}$. (It was attempted to express y k-sparsely with respect to the inferred dictionary B, and it was accomplished within $\|Bb - y\| = \|BPDc - y\| = \|Cc - y\| \leq \eta$ for $b, c \in S_{p,k}$, where $c = D^{-1}P^T b$.) Then $\|J^{-1}P^T b - a\| \leq \frac{\varepsilon\|b\|}{1-\varepsilon}(\frac{1}{\sqrt{1-\delta}} + 1) + \frac{\eta}{\sqrt{1-\delta}}$.

Proof. By substitution $y = Aa$ and $\|Cc - y\| \leq \eta$ yield $\|Cc - Aa\| \leq \eta$. By the definition of the Euclidean/Euclidean sorta-norm and $c \in S_{p,k}$ we get $\|(A - C)c\| \leq \varepsilon\|c\|$. By the triangle inequality $\|A(c - a)\| \leq \varepsilon\|c\| + \eta$. By $2k$-RIP and $c - a \in S_{p,2k}$, $\|c - a\| \leq \frac{\varepsilon\|c\|+\eta}{\sqrt{1-\delta}}$.

Unfortunately, we never learn c, only b, so we need to estimate $\|J^{-1}P^T b - a\| = \|J^{-1}Dc - a\| = \|c - a + (J^{-1}D - I)c\| \leq \|c - a\| + \|(J^{-1}D - I)c\| \leq \|c - a\| + \|J^{-1}D - I\| \cdot \|c\| \leq \|c - a\| + \varepsilon \cdot \|c\|$ since $\|J^{-1}D - I\| = \|J^{-1}(D - J)\| = \|D - J\| < \varepsilon$. Therefore $\|J^{-1}P^T b - a\| \leq \|c - a\| + \varepsilon \cdot \|c\| \leq \frac{\varepsilon\|c\|+\eta}{\sqrt{1-\delta}} + \varepsilon \cdot \|c\|$.

Since $\|D^{-1}\|_{\text{spectral}} \leq \frac{1}{1-\varepsilon}$, $\|c\| \leq \frac{1}{1-\varepsilon}\|b\|$, so $\|J^{-1}P^T b - a\| \leq \frac{\varepsilon\|b\|}{1-\varepsilon}(\frac{1}{\sqrt{1-\delta}} + 1) + \frac{\eta}{\sqrt{1-\delta}}$. □

Theorem 1. *Suppose that $k = 1$. Suppose $\delta \in (0,1)$. Suppose that A satisfies $(2,\delta)$-lower-RIP. Suppose $0 < \varepsilon < \sqrt{\frac{1-\delta}{2}}$. Suppose $f : S_{p,1} \to S_{p,1}$. Suppose the almost recovery condition*

$$\text{for all standard basis vectors } a, \ \|Aa - Bf(a)\|_{Euclid} \leq \varepsilon. \tag{7}$$

Then there exist a diagonal matrix $\underset{p\times p}{D}$ and a permutation matrix $\underset{p\times p}{P}$ s.t.

$$\|A - BPD\|_{restricted} \leq \varepsilon. \tag{8}$$

If in addition you suppose that real matrices $\underset{m\times p}{A}$ and $\underset{m\times p}{B}$ have Euclidean length 1 columns, we get that there is a diagonal matrix $\underset{p\times p}{J}$ whose diagonal *entries are* ± 1 *such that*

$$\|J - D\|_{spectral} < \varepsilon \tag{9}$$

and

$$\|D^{-1}\|_{spectral} < \frac{1}{1-\varepsilon}. \tag{10}$$

Proof. Since $k = 1$, f taking in a 1-sparse vector gives out a 1-sparse vector, so for each $i = 1, 2, 3, \ldots, p$ we can define c_i and π_i s.t.

$$f(e_i) = c_i e_{\pi_i}.$$

Each $c_i \neq 0$ because $c_i = 0$ would cause a contradiction because the almost recovery condition would tell us that $\|Ae_i\| = \|Ae_i - Bf(e_i)\| \leq \varepsilon$ whereas the RIP condition tells us that $\|Ae_i\| \geq \sqrt{1-\delta} > \varepsilon$.

We claim that $\pi : [p] \to [p]$ is injective. To see this, suppose instead that $\pi(i) = \pi(j)$ for $i \neq j$. By almost recovery,

$$\|Ae_i - Bf(e_i)\| \leq \varepsilon,$$

and

$$Bf(e_i) = B(c_i e_{\pi_i}) = \frac{c_i}{c_j} B(c_j e_{\pi_j}) = \frac{c_i}{c_j} Bf(e_j)$$

And thus

$$\|Ae_i - \frac{c_i}{c_j} Bf(e_j)\| \leq \varepsilon.$$

Also by almost recovery

$$\|Ae_j - Bf(e_j)\| \leq \varepsilon$$

and thus

$$\|\frac{c_i}{c_j} Ae_j - \frac{c_i}{c_j} Bf(e_j)\| \leq \varepsilon \frac{|c_i|}{|c_j|}$$

Putting these together by triangle inequality gives

$$\|Ae_i - \frac{c_i}{c_j} Ae_j\| \leq \varepsilon(1 + \frac{|c_i|}{|c_j|}). \tag{11}$$

Meanwhile, the lower-RIP condition on $x = e_i - \frac{c_i}{c_j} e_j$ gives

$$\|A(e_i - \frac{c_i}{c_j} e_j)\| \geq \sqrt{1-\delta}\sqrt{1 + \frac{c_i^2}{c_j^2}} > \varepsilon\sqrt{2}\sqrt{1 + \frac{c_i^2}{c_j^2}}.$$

Because $\forall x \in \mathbb{R}$, $1 + x \leq \sqrt{2}\sqrt{1+x^2}$, this is a contradiction via $x = \frac{|c_i|}{|c_j|}$ to (11). Thus π is injective, and thus bijective.

Let P be the permutation matrix whose i-th column is $e_{\pi(i)}$. Let D be the $p \times p$ diagonal matrix with $c_1, c_2, \ldots, c_p$ down the diagonal. We know that $\|Ae_i - B(c_i e_{\pi(i)})\| \leq \varepsilon$, i.e. that the i-th column Ae_i of A is very close to the i-th column of BPD, $BPDe_i = B(c_i e_{\pi(i)})$, and thus

$$\|A - BPD\|_{\text{restricted}} \leq \varepsilon.$$

Since the columns Ae_i of A are length one, and the columns $Be_{\pi(i)}$ of B are length one, by the triangle inequality we get that $\varepsilon \geq \|Ae_i - B(c_i e_{\pi(i)})\| \geq \big|\|Ae_i\| - |c_i|\|Be_{\pi(i)}\|\big| = |1 - |c_i||$, i.e. $1 - \varepsilon < |c_i| < 1 + \varepsilon$. Thus there is a diagonal matrix $\underset{p \times p}{J}$ whose diagonal entries are ± 1 such that

$$\|J - D\|_{\text{spectral}} < \varepsilon \tag{12}$$

and

$$\|D^{-1}\|_{\text{spect}} < \frac{1}{1-\varepsilon} \tag{13}$$

□

2. New Thoughts

Suppose $k = 2$. Suppose $\underset{4\times 4}{A}$ and $\underset{4\times 4}{B}$ satisfy $(2k, \delta)$-RIP, i.e. they are both nearly isometries, and they have length one columns. We define the notion of A's $\{i_1, i_2\}$ ellipse to be the set

$$A\{i_1, i_2\} := \{A(a_1 e_{i_1} + a_2 e_{i_2}) : a_1^2 + a_2^2 = 1\} = \{Aa : \|a\| = 1, \text{supp}(a) \subseteq \{i_1, i_2\}\}.$$

Similarly we define the notion of B's $\{r_1, r_2\}$ ellipse to be the set

$$B\{r_1, r_2\} := \{B(b_1 e_{r_1} + b_2 e_{r_2}) : b_1^2 + b_2^2 = 1\} = \{Bb : \|b\| = 1, \text{supp}(b) \subseteq \{r_1, r_2\}\}.$$

For any ellipse $A\{i_1, i_2\}$ there must be a corresponding $B\{r_1, r_2\}$ that minimizes the max distance

$$\max_{p \in A\{i_1, i_2\}} d(p, B\{r_1, r_2\}) \tag{14}$$

where the Euclidean distance from a point to a compact set is defined in the usual way. In this manner, the category of signals that A explains as being $\{i_1, i_2\}$-sparse is best explained by B as $\{r_1, r_2\}$-sparse, and this suggests a mapping $\tau(\{i_1, i_2\}) := \{r_1, r_2\}$. We must show that this map is well-defined, and has certain properties.

3. First question

If $B\{r_1, r_2\}$ and $B\{j_1, j_2\}$ are closer than $\sqrt{1-\delta}$ to each other, does that imply that $\{r_1, r_2\} = \{j_1, j_2\}$? Yes.

Proof. Suppose instead that $\{r_1, r_2\} \neq \{j_1, j_2\}$. WLOG we can assume that $r_1 \notin \{j_1, j_2\}$. Since they are closer than $\sqrt{1-\delta}$ to each other, there must be a point $B(b_1 e_{j_1} + b_2 e_{j_2})$ from $B\{j_1, j_2\}$ which is closer than $\sqrt{1-\delta}$ to Be_{r_1}. But then $\sqrt{1-\delta} > \|B(b_1 e_{j_1} + b_2 e_{j_2} - e_{r_1})\| \geq \sqrt{1-\delta}\|b_1 e_{j_1} + b_2 e_{j_2} - e_{r_1}\| \geq \sqrt{1-\delta}$, which is a contradiction. □

4. Coloring Theorems

Of course we don't get to pick $\epsilon > 0$ and $\delta \in (0,1)$, they are constants given to us. Define functions of η

$$\beta_{\text{lower}} := \frac{\sqrt{1-\delta}-\varepsilon}{\sqrt{1+\delta}} - \eta$$

and

$$\beta_{\text{upper}} := \frac{\sqrt{1+\delta}+\varepsilon}{\sqrt{1-\delta}} + \eta.$$

and then pick η and under constraints

$$\eta\sqrt{2} > 2\varepsilon\sqrt{1+\delta},$$
$$\sqrt{1-\delta}\cdot\beta_{\text{lower}} > \sqrt{1+\delta}\cdot\eta\cdot\sqrt{3}+\varepsilon\sqrt{3},$$

to minimize the quantity

$$\frac{2\varepsilon+\sqrt{1+\delta}\,(\beta_{\text{upper}}-\beta_{\text{lower}}+2\eta)}{\sqrt{1-\delta}}.$$

We call η the *threshold of insignificant involvement,* the notion being that in a given linear combination of basis vectors, if a basis vector's coefficient is less than η in absolute value, then that vector is deemed "not really involved" in the linear combination.

Fix $i_1, i_2 \in [p]$. Suppose $a_1, a_2, a_3 \in \operatorname{span}\{e_{i_1}, e_{i_2}\}$ with $\forall n\, \|a_n\| = 1$. Then by the assumptions, there certainly exist $j_1, j_2, j_3, k_1, k_2, k_3 \in [p]$ and $c_1, c_2, c_3, d_1, d_2, d_3 \in \mathbb{R}$ such that

$$\|Aa_1 - B(c_1 e_{j_1} + d_1 e_{k_1})\| \leq \varepsilon$$
$$\|Aa_2 - B(c_2 e_{j_2} + d_2 e_{k_2})\| \leq \varepsilon$$
$$\|Aa_3 - B(c_3 e_{j_3} + d_3 e_{k_3})\| \leq \varepsilon$$

Suppose that Aa_1, Aa_2 and Aa_3 are interpreted by B to be very close to 1-sparse and with different major support indices in the sense that the $j_1, j_2, j_3 \in [p]$ are distinct with the coefficients $|d_1|, |d_2|, |d_3|$ close to zero, say $\forall n = 1, 2, 3$

$$|d_n| \leq \eta.$$

We do not ask that $k_1, k_2, k_3 \in [p]$ be necessarily distinct from each other.

It will follow from this assumption that the coefficients $|c_1|, |c_2|, |c_3|$ must be close to one, more particularly it follows that $\forall n = 1, 2, 3$

$$\beta_{\text{lower}} \leq |c_n| \leq \beta_{\text{upper}},$$

where

$$\beta_{\text{lower}} := \frac{\sqrt{1-\delta}-\varepsilon}{\sqrt{1+\delta}} - \eta$$

and

$$\beta_{\text{upper}} := \frac{\sqrt{1+\delta}+\varepsilon}{\sqrt{1-\delta}} + \eta$$

via the following lemma:

Lemma 2 (Beta Bounds)**.** *Suppose that* $\|a\| = 1$,

$$\|Aa - B(ce_j + de_k)\| \leq \varepsilon$$

and $|d| \leq \eta$. *Then*

$$\beta_{lower} \leq |c| \leq \beta_{upper},$$

where

$$\beta_{lower} := \frac{\sqrt{1-\delta} - \varepsilon}{\sqrt{1+\delta}} - \eta$$

and

$$\beta_{upper} := \frac{\sqrt{1+\delta} + \varepsilon}{\sqrt{1-\delta}} + \eta.$$

Proof.

$$\begin{aligned} |c| + |d| = \|ce_j\| + \|de_k\| \geq \|ce_j + de_k\| &\geq \frac{\|B(ce_j + de_k)\|}{\sqrt{1+\delta}} \\ &\geq \frac{\|Aa\| - \varepsilon}{\sqrt{1+\delta}} \\ &\geq \frac{\sqrt{1-\delta} - \varepsilon}{\sqrt{1+\delta}} \end{aligned}$$

Thus

$$|c| \geq \frac{\sqrt{1-\delta} - \varepsilon}{\sqrt{1+\delta}} - \eta =: \beta_{\text{lower}}.$$

On the other side

$$\begin{aligned} |c| - |d| = \|ce_j\| - \|de_k\| \leq \|ce_j + de_k\| &\leq \frac{\|B(ce_j + de_k)\|}{\sqrt{1-\delta}} \\ &\leq \frac{\|Aa\| + \varepsilon}{\sqrt{1-\delta}} \\ &\leq \frac{\sqrt{1+\delta} + \varepsilon}{\sqrt{1-\delta}} \end{aligned}$$

Thus

$$|c| \leq \frac{\sqrt{1+\delta} + \varepsilon}{\sqrt{1-\delta}} + \eta =: \beta_{\text{upper}}.$$

□

We will say that a_1 is B 1-sparseish with support index j_1. Similarly will say that a_2 is B 1-sparseish with support index j_2, and a_3 is B 1-sparseish with support index j_3. Suppose further that it is true that

$$\sqrt{1-\delta} \cdot \beta_{\text{lower}} > \sqrt{1+\delta} \cdot \eta \cdot \sqrt{3} + \varepsilon\sqrt{3}.$$

Claim: this is a contradiction, i.e. you cannot have on the same circle $\{i_1, i_2\}$ three B 1-sparseish a's with distinct support indices.

Proof. To see this, by $a_1, a_2, a_3 \in \text{span}\{e_{i_1}, e_{i_2}\}$ we can pick $g_1, g_2, g_3 \in \mathbb{R}$ such that $g_1 a_1 + g_2 a_2 + g_3 a_3 = 0$ and $g_1^2 + g_2^2 + g_3^2 = 1$. Thus by triangle inequality

$$\begin{aligned}\|A(g_1 a_1 + \ldots + g_3 a_3) - B[g_1(c_1 e_{j_1} + d_1 e_{k_1}) + \ldots + g_3(c_3 e_{j_3} + d_3 e_{k_3})]\| \\ \leq \varepsilon(|g_1| + |g_2| + |g_3|)\end{aligned}$$

i.e.

$$\|B[g_1(c_1 e_{j_1} + d_1 e_{k_1}) + \ldots + g_3(c_3 e_{j_3} + d_3 e_{k_3})]\| \leq \varepsilon(|g_1| + |g_2| + |g_3|)$$

and triangle inequality

$$\|B(g_1 c_1 e_{j_1} + \ldots + g_3 c_3 e_{j_3})\| - \|B(g_1 d_1 e_{k_1} + \ldots + g_3 d_3 e_{k_3})\| \leq \varepsilon(|g_1| + |g_2| + |g_3|)$$

move to the other side

$$\|B(g_1 c_1 e_{j_1} + \ldots + g_3 c_3 e_{j_3})\| \leq \|B(g_1 d_1 e_{k_1} + \ldots + g_3 d_3 e_{k_3})\| + \varepsilon(|g_1| + |g_2| + |g_3|)$$

B is $(4, \delta)$-RIP so

$$\sqrt{1-\delta}\|g_1 c_1 e_{j_1} + \ldots + g_3 c_3 e_{j_3}\| \leq \sqrt{1+\delta}\|g_1 d_1 e_{k_1} + \ldots + g_3 d_3 e_{k_3}\| + \varepsilon(|g_1| + |g_2| + |g_3|)$$

distinctness of $j_1, j_2, j_3 \in [p]$ gives

$$\sqrt{1-\delta}\sqrt{g_1^2 c_1^2 + \ldots + g_3^2 c_3^2} \leq \sqrt{1+\delta}\|g_1 d_1 e_{k_1} + \ldots + g_3 d_3 e_{k_3}\| + \varepsilon(|g_1| + |g_2| + |g_3|)$$

triangle inequality

$$\sqrt{1-\delta}\sqrt{g_1^2 c_1^2 + \ldots + g_3^2 c_3^2} \leq \sqrt{1+\delta}(|g_1 d_1| + \ldots + |g_3 d_3|) + \varepsilon(|g_1| + |g_2| + |g_3|)$$

$|c|_{\text{min}} := \min\{|c_1|, |c_2|, |c_3|\} \geq \beta_{\text{lower}}$, $|d|_{\text{max}} := \max\{|d_1|, |d_2|, |d_3|\} \leq \eta$ gives

$$\sqrt{1-\delta} \cdot \beta_{\text{lower}} \cdot \sqrt{g_1^2 + \ldots + g_3^2} \leq \sqrt{1+\delta} \cdot \eta \cdot (|g_1| + \ldots + |g_3|) + \varepsilon(|g_1| + |g_2| + |g_3|)$$

$g_1^2 + g_2^2 + g_3^2 = 1$ and Cauchy-Schwarz

$$\sqrt{1-\delta}\beta_{\text{lower}} \leq \sqrt{1+\delta} \cdot \eta \cdot \sqrt{3} + \varepsilon\sqrt{3}.$$

This is a contradiction. Thus this circle $A\{i_1, i_2\}$ could have at most two distinct support indices j_1 and j_2 which have B 1-sparseish vectors, but not three. □

Lemma 3 (Same 1-Sparseish Support Index Implies Close)**.** *Suppose both* $\|a_1\| = 1$ *and* $\|a_2\| = 1$ *are " B 1-sparseish with the same support index* $j \in [p]$*" in that*

$$\|Aa_1 - B(c_1 e_j + d_1 e_{k_1})\| \leq \varepsilon$$

and

$$\|Aa_2 - B(c_2 e_j + d_2 e_{k_2})\| \leq \varepsilon,$$

where $|d_1|, |d_2| \leq \eta$*. Suppose also that the sign of* c_1 *is the same as the sign of* c_2*. Then* a_1 *and* a_2 *must this close to each other:*

$$\|a_1 - a_2\| \leq \frac{2\varepsilon + \sqrt{1+\delta}\,(\beta_{upper} - \beta_{lower} + 2\eta)}{\sqrt{1-\delta}}.$$

Proof. By triangle inequality

$$\|A(a_1 - a_2)\| \leq 2\varepsilon + \|B[(c_1 - c_2)e_j + d_1 e_{k_1} - d_2 e_{k_2}]\|$$

by $(4, \delta)$-RIP on A and B

$$\sqrt{1-\delta}\|(a_1 - a_2)\| \leq 2\varepsilon + \sqrt{1+\delta}\|(c_1 - c_2)e_j + d_1 e_{k_1} - d_2 e_{k_2}\|$$

More triangle inequality

$$\sqrt{1-\delta}\|(a_1 - a_2)\| \leq 2\varepsilon + \sqrt{1+\delta}\,(|c_1 - c_2| + |d_1| + |d_2|)$$

c_1 and c_2 have the same sign and the beta bounds give

$$\sqrt{1-\delta}\|(a_1 - a_2)\| \leq 2\varepsilon + \sqrt{1+\delta}\,(\beta_{\text{upper}} - \beta_{\text{lower}} + \eta + \eta)$$

$$\|a_1 - a_2\| \leq \frac{2\varepsilon + \sqrt{1+\delta}\,(\beta_{\text{upper}} - \beta_{\text{lower}} + 2\eta)}{\sqrt{1-\delta}}$$

□

Thus a given circle $A\{i_1, i_2\}$ can have at most four small regions on it where the B explanation is 1-sparseish, namely for at most two support indices j_1 and j_2, and signs on them: positive coefficient times e_{j_1}-ish, positive coefficient times e_{j_2}-ish, negative coefficient times e_{j_1}-ish, and negative coefficient times e_{j_2}-ish. This is not enough to cover the whole circle $\{i_1, i_2\}$ (imagine a circle with four small sections missing), so there must a long segment of points on the circle where the B explanation is NOT 1-sparseish, i.e. B explains those points as a large coefficiented linear combination of some e_{j_1}, e_{j_2}. But then either this segment has a single color=support set $=\{j_1, j_2\}$ that B-explains them all, or there is a point a that has two different color explanations $\{j_1, j_2\}$ and $\{k_1, k_2\}$. But the second possibility is a contradiction since the coefficients on this long segment must be large since every point on it is NOT B 1-sparseish:

$$\|Aa - B(c_1 e_{j_1} + c_2 e_{j_2})\| \leq \varepsilon$$

with $|c_1|, |c_2| > \eta$ both big and

$$\|Aa - B(d_1 e_{k_1} + d_2 e_{k_2})\| \leq \varepsilon$$

with $|d_1|, |d_2| > \eta$ both big leads to

$$\|B(c_1 e_{j_1} + c_2 e_{j_2}) - B(d_1 e_{k_1} + d_2 e_{k_2})\| \leq 2\varepsilon$$

which by $(4, \delta)$-RIP leads to

$$\|c_1 e_{j_1} + c_2 e_{j_2} - d_1 e_{k_1} - d_2 e_{k_2}\| \leq 2\varepsilon\sqrt{1+\delta}$$

and at least two of $e_{k_1}, e_{k_2}, e_{j_1}, e_{j_2}$ have no one to cancel with, so

$$\eta\sqrt{2} \leq 2\varepsilon\sqrt{1+\delta}$$

By the way that η was chosen, this is a contradiction. Thus there is a long almost quarter segment of any circle $\{i_1, i_2\}$ which is "monochromatic" in its B-explanation.

Let m_1 and m_2 be two same $\{j_1, j_2\}$-B-colored unit vectors on the $\{i_1, i_2\}$ circle at maximally uncorrelated angle θ, hopefully almost $\approx \pi/2$ radians apart. and form a matrix $M = [m1|m2]$. Let $\|a\| = 1$ be any vector on the $\{i_1, i_2\}$ circle. Certainly there are coefficients c_1, c_2 such that $a = c_1 m_1 + c_2 m_2 = Mc$. Since

$$\|Am_1 - B(d_1 e_{j_1} + d_2 e_{j_2})\| \leq \varepsilon$$

and

$$\|Am_2 - B(g_1 e_{j_1} + g_2 e_{j_2})\| \leq \varepsilon$$

by linear combination and triangle inequality

$$\begin{aligned} \|A(c_1 m_1 + c_2 m_2) - B((c_1 d_1 + c_2 g_1)e_{j_1} + (c_1 d_2 + c_2 g_2)e_{j_2})\| &\leq (|c_1| + |c_2|)\varepsilon \\ &\leq \sqrt{2} \cdot \sqrt{c_1^2 + c_2^2} \cdot \varepsilon \\ &\leq \frac{\sqrt{2} \cdot \varepsilon}{1 - |\cos(\theta)|} \end{aligned}$$

because

$$\begin{aligned} \max_{\|c_1 m_1 + c_2 m_2\| = 1} \left(c_1^2 + c_2^2\right) &= \max_{\|Mc\|=1} \|c\|^2 = \max_{c^T M^T M c = 1} \|c\|^2 \\ &= \frac{1}{\min\limits_{\|\hat{c}\|=1} \hat{c}^T M^T M \hat{c}} = \frac{1}{(1 - |\cos(\theta)|)^2} \end{aligned}$$

since

$$M^T M = \begin{bmatrix} 1 & \langle m1, m2 \rangle \\ \langle m1, m2 \rangle & 1 \end{bmatrix} = \begin{bmatrix} 1 & \cos(\theta) \\ \cos(\theta) & 1 \end{bmatrix}$$

Whose eigenvalues are $1 \pm |\cos(\theta)| = 1 \pm |\langle m_1, m_2 \rangle|$.

Thus we see that the entire $\{i_1, i_2\}$ circle can be B-explained as being $\{j_1, j_2\}$-B-colored if you increase the tolerance to $\frac{\sqrt{2} \cdot \varepsilon}{1 - |\cos(\theta)|}$.

5. Wedge Products

Consider $\wedge^2 \mathbb{R}^n = \wedge^2(\mathbb{R}^n)$, i.e. the span of all the symbols $\{e_i \wedge e_j \,|\, i, j \in [n]\}$ modded out by the usual truths like $x_i \wedge y_j = -y_j \wedge x_i$, left and right distributive, constant pullout, etc., familiar from differential forms.

The inner product on $\wedge^2 \mathbb{R}^p$ is defined by

$$\langle u \wedge v, x \wedge y \rangle := \det \begin{bmatrix} \langle u, x \rangle & \langle u, y \rangle \\ \langle v, x \rangle & \langle v, y \rangle \end{bmatrix} = \langle u, x \rangle \langle v, y \rangle - \langle u, y \rangle \langle v, x \rangle. \tag{15}$$

so in particular

$$\|u \wedge v\|^2 = \langle u \wedge v, u \wedge v \rangle = \|u\|^2 \|v\|^2 - \langle u, v \rangle^2. \tag{16}$$

Given $A : \mathbb{R}^p \to \mathbb{R}^m$, we define $\wedge^2 A : \wedge^2 \mathbb{R}^p \to \wedge^2 \mathbb{R}^m$ via defining it on the basis elements via

$$\wedge^2 A(e_i \wedge e_j) := (Ae_i) \wedge (Ae_j).$$

Suppose that A satisfies the $(2k, \delta)$-RIP. We claim that $\wedge^2 A$ satisfies $(k, 4\delta)$-RIP.

Lemma 4. *Suppose that A satisfies the $(2k, \delta)$-RIP. Then for any $u, v \in S_{p,k}$*

$$\langle u, v\rangle - \frac{1}{2}\delta(\|u\|^2 + \|v\|^2) \le \langle Au, Av\rangle \le \langle u, v\rangle + \frac{1}{2}\delta(\|u\|^2 + \|v\|^2). \tag{17}$$

Proof. $u + v, u - v \in S_{p,2k}$ so

$$\begin{aligned} 4(1-\delta)\langle u, v\rangle - 2\delta(\|u-v\|^2) &= 4\langle u, v\rangle - 2\delta(\|u\|^2 + \|v\|^2) = \\ (1-\delta)\|u+v\|^2 - (1+\delta)\|u-v\|^2 &\le \|A(u+v)\|^2 - \|A(u-v)\|^2 = 4\langle Au, Av\rangle \\ &\le (1+\delta)\|u+v\|^2 - (1-\delta)\|u-v\|^2 \\ &= 4\langle u, v\rangle + 2\delta(\|u\|^2 + \|v\|^2) = \\ &= 4(1+\delta)\langle u, v\rangle + 2\delta(\|u-v\|^2) \end{aligned}$$

Now divide through by 4. □

Corollary 1. *Suppose that A satisfies the $(2, \delta)$-RIP. Then for any $i \neq j \in [p]$, $|\langle Ae_i, Ae_j\rangle| \le \delta$.*

Lemma 5. *Suppose that A satisfies the $(2, \delta)$-RIP. For any $i \neq j \in [p]$, $1 - 3\delta + \delta^2 \le \|(\wedge^2 A)(e_i \wedge e_j)\|^2$.*

Proof.

$$\begin{aligned} \|(\wedge^2 A)(e_i \wedge e_j)\|^2 &:= \|Ae_i \wedge Ae_j\|^2 = \|Ae_i\|^2 \cdot \|Ae_j\|^2 - \langle Ae_i, Ae_j\rangle^2 \\ &\ge \|Ae_i\|^2 \cdot \|Ae_j\|^2 - \delta \\ &\ge (1-\delta)^2 - \delta = 1 - 3\delta + \delta^2. \end{aligned}$$

Similarly

$$\begin{aligned} \|(\wedge^2 A)(e_i \wedge e_j)\|^2 &:= \|Ae_i \wedge Ae_j\|^2 = \|Ae_i\|^2 \cdot \|Ae_j\|^2 - \langle Ae_i, Ae_j\rangle^2 \\ &\le \|Ae_i\|^2 \cdot \|Ae_j\|^2 \\ &\le (1+\delta)^2 = 1 + 2\delta + \delta^2. \end{aligned}$$

□

Theorem 2. *Suppose that A satisfies the $(2, \delta)$-RIP. For any $\{i, j\} \neq \{k, l\} \subseteq [p]$,*

$$(1 - 4\delta)\|c(e_i \wedge e_j) + d(e_k \wedge e_l)\|^2 \le \|(\wedge^2 A)(c \cdot e_i \wedge e_j + d \cdot e_k \wedge e_l)\|^2,$$

which is to say that $\wedge^2 A$ satisfies $(2, 4\delta)$-lower-RIP. Also $\wedge^2 A$ satisfies $(2, 5\delta)$-upper-RIP.

Proof.

$$
\begin{aligned}
&\|(\wedge^2 A)(c \cdot e_i \wedge e_j + d \cdot e_k \wedge e_l)\|^2 := \|c \cdot Ae_i \wedge Ae_j + d \cdot Ae_k \wedge Ae_l)\|^2 \\
&= c^2\|Ae_i \wedge Ae_j\|^2 + d^2\|Ae_k \wedge Ae_l\|^2 + 2cd\langle Ae_i \wedge Ae_j, Ae_k \wedge Ae_l\rangle \\
&= c^2\|Ae_i \wedge Ae_j\|^2 + d^2\|Ae_k \wedge Ae_l\|^2 + 2cd(\langle Ae_i, Ae_k\rangle\langle Ae_j, Ae_l\rangle - \langle Ae_i, Ae_l\rangle\langle Ae_j, Ae_k\rangle) \\
&\geq (1 - 3\delta + \delta^2)(c^2 + d^2) - 4|c| \cdot |d| \cdot \delta^2 \\
&\geq (1 - 3\delta + \delta^2)(c^2 + d^2) - 2(c^2 + d^2)\delta^2 \\
&= (1 - 3\delta - \delta^2)(c^2 + d^2) \\
&\geq (1 - 4\delta)(c^2 + d^2) = (1 - 4\delta)\|c \cdot e_i \wedge e_j + d \cdot e_k \wedge e_l\|^2.
\end{aligned}
$$

Similarly

$$
\begin{aligned}
&\|(\wedge^2 A)(c \cdot e_i \wedge e_j + d \cdot e_k \wedge e_l)\|^2 := \|c \cdot Ae_i \wedge Ae_j + d \cdot Ae_k \wedge Ae_l)\|^2 \\
&= c^2\|Ae_i \wedge Ae_j\|^2 + d^2\|Ae_k \wedge Ae_l\|^2 + 2cd\langle Ae_i \wedge Ae_j, Ae_k \wedge Ae_l\rangle \\
&= c^2\|Ae_i \wedge Ae_j\|^2 + d^2\|Ae_k \wedge Ae_l\|^2 + 2cd(\langle Ae_i, Ae_k\rangle\langle Ae_j, Ae_l\rangle - \langle Ae_i, Ae_l\rangle\langle Ae_j, Ae_k\rangle) \\
&\leq (1 + 2\delta + \delta^2)(c^2 + d^2) + 4|c| \cdot |d| \cdot \delta^2 \\
&\leq (1 + 2\delta + \delta^2)(c^2 + d^2) + 2(c^2 + d^2)\delta^2 \\
&= (1 + 4\delta + \delta^2)(c^2 + d^2) \\
&\leq (1 + 5\delta)(c^2 + d^2) = (1 + 5\delta)\|c \cdot e_i \wedge e_j + d \cdot e_k \wedge e_l\|^2.
\end{aligned}
$$

□

So RIP for A seems to force weaker constanted RIP for $\wedge^2 A$. Who cares? This machinery needs a purpose:

Theorem 3. *Suppose that we satisfy the various hypotheses of the coloring theorems, so that for any $\{i_1, i_2\}$ the circle $A\{i_1, i_2\}$ possesses a B-explanation by corresponding support indices $\{j_1, j_2\}$ up to error E, as concluded by the coloring theorem. Then in particular there exist $a, b \in \mathbb{R}$ such that*

$$\|Ae_{i_1} - B(ae_{j_1} + be_{j_2})\| \leq E,$$

and similarly there exist $c, d \in \mathbb{R}$ such that

$$\|Ae_{i_2} - B(ce_{j_1} + de_{j_2})\| \leq E.$$

Therefore conceptually

$$
\begin{aligned}
&(\wedge^2 A)(e_{i_1} \wedge e_{i_2}) = Ae_{i_1} \wedge Ae_{i_2} \approx [B(ae_{j_1} + be_{j_2})] \wedge [B(ce_{j_1} + de_{j_2})] \\
&= [aBe_{j_1} + bBe_{j_2}] \wedge [cBe_{j_1} + dBe_{j_2}] \\
&= (ad - bc)Be_{j_1} \wedge Be_{j_2} = (ad - bc)(\wedge^2 B)(e_{j_1} \wedge e_{j_2})
\end{aligned}
$$

And thus

$$\|(\wedge^2 A)(e_{i_1} \wedge e_{i_2}) - (\wedge^2 B)\left((ad - bc)e_{j_1} \wedge e_{j_2}\right)\| \leq error.$$

This is to say that $\wedge^2 A$ and $\wedge^2 B$ satisfy the approximation hypothesis 7 of theorem 1. Therefore, since we know that A and B having $(2, \delta)$-RIP gives $\wedge^2 A$ and $\wedge^2 B$ $(2, 4\delta)$-lower-RIP, we know that theorem 1 can be applied. Thus the matrices representing $\wedge^2 A$ and $\wedge^2 B$ are different only up to permutation of the columns and scaling.

We could also recurse because $\wedge^2 A$ and $\wedge^2 B$ satisfy an RIP and a approximation hypothesis so $\wedge^2(\wedge^2 A) = \wedge^4 A$ and $\wedge^2(\wedge^2 B) = \wedge^4 A$ also satisfy an RIP and an approximation hypothesis. Even this needs a purpose.

Mathematical Sciences Research Institute, 17 Gauss Way, Berkeley, CA 94720

Redwood Center for Theoretical Neuroscience, University of California, Berkeley, CA 94720
E-mail address: chillar@msri.org

Department of Mathematics, University of California, Berkeley, CA 94720