Errata to the Paper "On the Restricted Isometry of the Columnwise Khatri-Rao Product"

Saurabh Khanna, *Member, IEEE* and Chandra R. Murthy, *Senior Member, IEEE*

Dept. of ECE, Indian Institute of Science, Bangalore, India

{saurabh, cmurthy}@iisc.ac.in

***Abstract*—In [1], Proposition 15 is incorrect. Due to this error, the statements of Theorems 2 and 3 in [1] claiming $m \geq O(\sqrt{k}\log^{3/2} n)$ as sufficient for k^{th} order restricted isometry property (RIP) of the columnwise Khatri-Rao product of two $m \times n$ sized random matrices containing independent subgaussian entries may not hold true. This errata corrects the claims of Theorems 2 and 3 in [1] to show that a higher sample complexity requirement, $m \geq O(k \log n)$, is the new sufficient condition. The k-RIP compliance of the columnwise Khatri-Rao product for m scaling sublinearly with k remains an open question.**

The deterministic bounds for the k^{th}-order restricted isometric constants of a generic columnwise Khatri-Rao product presented in [1] remain unchanged.

I. Error in Proposition 15

Proposition 15 in [1] makes an erroneous claim that a non-negative random variable $\mathbf{z}$ with a subgaussian tail probability $\left(\mathbb{P}\left(\mathbf{z} - \mathbb{E}\mathbf{z} > t\right) \leq \exp\left(-t^2/2\nu^2\right)\right)$ satisfies $\mathbb{E}\mathbf{z} \leq \sqrt{2\pi}\nu$. As a consequence, the proofs of Lemmas 4 and 6 in [1] which rely on Proposition 15 are invalid, and the probabilistic bounds for the restricted isometry constants (RICs) of the columnwise Khatri-Rao product between random subgaussian matrices in Theorems 2 and 3 may not hold.

In Section II, we state and prove a corrected, weaker version of Theorem 2 in [1], which discusses a probabilistic bound for the k-RIC of the columnwise Khatri-Rao product between two independent random subgaussian matrices. In Section III, we replace Theorem 3 in [1] with its weaker version which provides a probabilistic k-RIC bound for the Khatri-Rao product of a random subgaussian matrix with itself. The proof of Theorem 3 is omitted due to lack of space, but it follows along similar lines as the proof of Theorem 2. The detailed proof can be found in [2]. All through this note, the notation is the same as in [1].

II. Correction To Theorem 2

We begin with a corollary of the Hanson-Wright inequality [1, Theorem 13] about the tail probability of the weighted inner product between two subgaussian vectors.

Corollary 1. *Let $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) \in \mathbb{R}^n$ and $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) \in \mathbb{R}^n$ be independent random vectors with independent subgaussian components satisfying $\mathbb{E}\mathbf{u}_i = \mathbb{E}\mathbf{v}_i = 0$ and $||\mathbf{u}_i||_{\psi_2} \leq K$, $||\mathbf{v}_i||_{\psi_2} \leq K$. Let $\mathbf{D}$ be an $n \times n$ matrix. Then, for every $t \geq 0$,*

$$\mathbb{P}\left\{\left|\mathbf{u}^T\mathbf{D}\mathbf{v}\right| > t\right\} \leq 2\exp\left[-c\min\left(\frac{t^2}{K^4\,||\mathbf{D}||_{HS}^2}, \frac{t}{K^2\,|||\mathbf{D}|||_2}\right)\right]$$

where c is a universal positive constant.

Proof. The desired tail bound is obtained by using the Hanson-Wright inequality [1, Theorem 13] with $\mathbf{x} = \left[\mathbf{u}^T\mathbf{v}^T\right]^T$ and $\mathbf{A} = [\mathbf{0}_{n\times n} \mid \mathbf{D}; \mathbf{0}_{n\times n} \mid \mathbf{0}_{n\times n}]$. □

Given a pair of random input matrices with i.i.d. subgaussian entries, the following corrected version of Theorem 2 in [1] provides an upper bound for the k-RIC of their columnwise Khatri-Rao product.

Theorem 2. *Suppose $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ matrices with real i.i.d. subgaussian entries, such that $\mathbb{E}\mathbf{A}_{ij} = 0$, $\mathbb{E}\mathbf{A}_{ij}^2 = 1$, and $||\mathbf{A}_{ij}||_{\psi_2} \leq K$, and similarly for $\mathbf{B}$. Then, for any $\delta > 0$, the k^{th} order restricted isometry constant δ_k of $\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{B}}{\sqrt{m}}$ satisfies $\delta_k \leq \delta$ with probability at least $1 - 10\,n^{-2(\gamma-1)}$ for any $\gamma > 1$, provided that*

$$m \geq 4c\gamma K_o^4\left(\frac{k\log n}{\delta}\right).$$

Here, $K_o = \max(K, 1)$ and c is a universal positive constant.

Proof. We begin with a variational definition of the k-RIC:

$$\delta_k\left(\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{B}}{\sqrt{m}}\right) = \sup_{\substack{\mathbf{z}\in\mathbb{R}^n,\\ ||\mathbf{z}||_2=1, ||\mathbf{z}||_0\leq k}} \left|\left|\left|\left(\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{B}}{\sqrt{m}}\right)\mathbf{z}\right|\right|_2^2 - 1\right|. \tag{1}$$

In order to find a probabilistic upper bound for δ_k, we seek to find a constant $\delta \in (0,1)$ such that $\mathbb{P}(\delta_k\left(\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{B}}{\sqrt{m}}\right) \geq \delta)$ is arbitrarily close to zero. We therefore consider the tail event

$$\mathcal{E} \triangleq \left\{\sup_{\substack{\mathbf{z}\in\mathbb{R}^n,\\ ||\mathbf{z}||_2=1, ||\mathbf{z}||_0\leq k}} \left|\left|\left|\left(\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{B}}{\sqrt{m}}\right)\mathbf{z}\right|\right|_2^2 - 1\right| \geq \delta\right\}, \tag{2}$$

and show that for m sufficiently large, $\mathbb{P}(\mathcal{E})$ can be driven arbitrarily close to zero. In other words, the constant δ serves as a probabilistic upper bound for $\delta_k\left(\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{B}}{\sqrt{m}}\right)$. Let $\mathcal{U}_k$ denote the set of all k or less sparse unit norm vectors in $\mathbb{R}^n$. Then, using Proposition 12 in [1], the tail event in (2) can be rewritten as

$$\begin{aligned}\mathbb{P}(\mathcal{E}) &= \mathbb{P}\left(\sup_{\mathbf{z}\in\mathcal{U}_k}\left|\mathbf{z}^T\left(\mathbf{A}\odot\mathbf{B}\right)^T\left(\mathbf{A}\odot\mathbf{B}\right)\mathbf{z} - m^2\right| \geq \delta m^2\right) \\ &= \mathbb{P}\left(\sup_{\mathbf{z}\in\mathcal{U}_k}\left|\mathbf{z}^T\left(\mathbf{A}^T\mathbf{A}\circ\mathbf{B}^T\mathbf{B}\right)\mathbf{z} - m^2\right| \geq \delta m^2\right) \\ &= \mathbb{P}\left(\sup_{\mathbf{z}\in\mathcal{U}_k}\left|\sum_{i=1}^n\sum_{j=1}^n z_i z_j\left(\mathbf{a}_i^T\mathbf{a}_j\right)\left(\mathbf{b}_i^T\mathbf{b}_j\right) - m^2\right| \geq \delta m^2\right), \end{aligned} \tag{3}$$

where $\mathbf{a}_i$ and $\mathbf{b}_i$ denote the ith column of $\mathbf{A}$ and $\mathbf{B}$, respectively. Further, by applying the triangle inequality and the union bound, the above tail probability splits as

$$\mathbb{P}(\mathcal{E}) \leq \mathbb{P}\left(\sup_{\mathbf{z}\in\mathcal{U}_k}\left|\sum_{i=1}^n z_i^2 \left|\left|\mathbf{a}_i\right|\right|_2^2 \left|\left|\mathbf{b}_i\right|\right|_2^2 - m^2\right| \geq \alpha\delta m^2\right)$$
$$+\mathbb{P}\left(\sup_{\mathbf{z}\in\mathcal{U}_k}\left|\sum_{i=1}^n \sum_{j=1,j\neq i}^n z_i z_j \mathbf{a}_i^T\mathbf{a}_j\mathbf{b}_i^T\mathbf{b}_j\right| \geq (1-\alpha)\delta m^2\right). \quad (4)$$

In the above, $\alpha \in (0,1)$ is a variational union bound parameter which can be optimized at a later stage. We now proceed to find separate upper bounds for each of the two probability terms in (4).

From [1, (32)], the first tail probability term in (4) is bounded as

$$\mathbb{P}\left(\sup_{\mathbf{z}\in\mathcal{U}_k}\left|\sum_{i=1}^n z_i^2 \left|\left|\mathbf{a}_i\right|\right|_2^2 \left|\left|\mathbf{b}_i\right|\right|_2^2 - m^2\right| \geq \alpha\delta m^2\right)$$
$$\leq 8ne^{-cm\frac{\alpha^2\delta^2}{4K_o^4}(1-\alpha\delta/4)^2}$$
$$= 8n^{-\left(\frac{cm\alpha^2\delta^2(1-\alpha\delta/4)^2}{4K_o^4\log n}-1\right)}. \quad (5)$$

In order to bound the second tail probability term in (4), we note that

$$\sup_{\mathbf{z}\in\mathcal{U}_k}\left|\sum_{i=1}^n \sum_{j=1,j\neq i}^n z_i z_j \mathbf{a}_i^T\mathbf{a}_j\mathbf{b}_i^T\mathbf{b}_j\right|$$
$$\leq \sup_{\mathbf{z}\in\mathcal{U}_k}\sum_{i=1}^n \sum_{j=1,j\neq i}^n |z_i z_j| \left|\mathbf{a}_i^T\mathbf{a}_j\right| \left|\mathbf{b}_i^T\mathbf{b}_j\right|$$
$$\leq \sup_{\mathbf{z}\in\mathcal{U}_k}\left(\sum_{i=1}^n \sum_{j=1,j\neq i}^n |z_i z_j|\right)\left(\max_{\substack{i,j\in\text{supp}(\mathbf{u}),\\ i\neq j}}\left|\mathbf{a}_i^T\mathbf{a}_j\right| \left|\mathbf{b}_i^T\mathbf{b}_j\right|\right)$$
$$\leq k\left(\max_{\substack{i,j\in[n],\\ i\neq j}}\left|\mathbf{a}_i^T\mathbf{a}_j\right| \left|\mathbf{b}_i^T\mathbf{b}_j\right|\right), \quad (6)$$

where the second step is an application of Hölders inequality. The last step follows from $||\mathbf{z}||_1 \leq \sqrt{k}$ for $\mathbf{z} \in \mathcal{U}_k$. Using (6), and by applying the union bound over $\binom{n}{2}$ possible distinct (i,j) pairs, the second probability term in (4) can be bounded as

$$\mathbb{P}\left(\sup_{\mathbf{z}\in\mathcal{U}_k}\left|\sum_{i=1}^n \sum_{j=1,j\neq i}^n z_i z_j \mathbf{a}_i^T\mathbf{a}_j\mathbf{b}_i^T\mathbf{b}_j\right| \geq (1-\alpha)\delta m^2\right)$$
$$\leq \frac{n^2}{2}\mathbb{P}\left(\left|\mathbf{a}_1^T\mathbf{a}_2\right| \left|\mathbf{b}_1^T\mathbf{b}_2\right| \geq \frac{(1-\alpha)\delta m^2}{k}\right)$$
$$\leq n^2\mathbb{P}\left(\left|\mathbf{a}_1^T\mathbf{a}_2\right| \geq \frac{\sqrt{(1-\alpha)\delta}m}{\sqrt{k}}\right)$$
$$\leq 2n^2 e^{-\frac{c(1-\alpha)\delta m}{K_o^4 k}} = 2n^{-\left(\frac{c(1-\alpha)\delta m}{K_o^4 k\log n}-2\right)}. \quad (7)$$

The last inequality in the above is obtained by using the tail bound for $|\mathbf{a}_1^T\mathbf{a}_2|$ from Corollary 1. Finally, by combining (4), (5) and (7), and setting $\alpha = 1/2$, we obtain the following simplified tail bound,

$$\mathbb{P}(\mathcal{E}) \leq 8n^{-\left(\frac{cm\delta^2(1-\delta/8)^2}{16K_o^4\log n}-1\right)} + 2n^{-\left(\frac{c\delta m}{2K_o^4 k\log n}-2\right)}. \quad (8)$$

From (8), for $m > \max\left(\frac{4\gamma K_o^4 k\log n}{c\delta}, \frac{32\gamma K_o^4\log n}{c\delta^2(1-\delta/8)^2}\right)$ and any $\gamma > 1$, we have $\mathbb{P}(\mathcal{E}) < 10\,n^{-2(\gamma-1)}$. Note that, in terms of k and n, the first term in the inequality for m scales as $k\log n$; it dominates the second term, which scales as $\log n$. This ends our proof. □

III. Correction To Theorem 3

Theorem 3. *Let $\mathbf{A}$ be an $m \times n$ matrix with real i.i.d. subgaussian entries, such that $\mathbb{E}\mathbf{A}_{ij} = 0$, $\mathbb{E}\mathbf{A}_{ij}^2 = 1$, and $||\mathbf{A}_{ij}||_{\psi_2} \leq K$. Then, for any $\delta > 0$ the k^{th} order restricted isometry constant δ_k of the column-normalized self Khatri-Rao product $\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{A}}{\sqrt{m}}$ satisfies $\delta_k \leq \delta$ with probability at least $1 - 5n^{-2(\gamma-1)}$ for any $\gamma \geq 1$, provided*

$$m \geq 4c'\gamma K_o^4\left(\frac{k\log n}{\delta}\right).$$

Here, $K_o = \max(K,1)$ and $c' > 0$ is a universal constant.

Proof. A detailed proof is given in [2]. □

IV. Remarks

Remark 1: According to Theorem 2, for fixed k and n, $\delta_k\left(\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{B}}{\sqrt{m}}\right) \leq O\left(\frac{1}{m}\right)$ with high probability, which is an improvement over $O(\frac{1}{\sqrt{m}})$ decay rate [3] for individual k-RICs of the input subgaussian matrices $\frac{\mathbf{A}}{\sqrt{m}}$ and $\frac{\mathbf{B}}{\sqrt{m}}$. Therefore, we conclude that the Khatri-Rao product exhibits stronger restricted isometry property, with smaller k-RICs compared to the k-RICs for the input matrices.

Remark 2: For $\mathbf{A}, \mathbf{B}$ as constructed in Theorem 2, a straightforward application of [4, Lemma 2] and the eigenvalue interlacing theorem [5] gives the following relation.

$$\delta_k\left(\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{B}}{\sqrt{m}}\right) \leq \delta_k\left(\frac{\mathbf{A}}{\sqrt{m}} \otimes \frac{\mathbf{B}}{\sqrt{m}}\right) \leq O\left(\frac{1}{\sqrt{m}}\right), \quad (9)$$

for n, k fixed. In comparison, Theorem 2 suggests a tighter upper bound $\delta_k\left(\frac{\mathbf{A}}{\sqrt{m}} \odot \frac{\mathbf{B}}{\sqrt{m}}\right) \leq O\left(\frac{1}{m}\right)$. We conjecture that an even faster $O\left(\frac{1}{m^2}\right)$ decay rate prevails, but it appears that a different approach would be required to establish this result.

Acknowledgment

We thank Piya Pal at the University of California, San Diego, USA, for pointing us to the error in [1].

References

[1] S. Khanna and C. R. Murthy, "On the restricted isometry of the columnwise Khatri-Rao product," vol. 66, no. 5, pp. 1170–1183, March 2018.
[2] ——, "On the restricted isometry of the columnwise Khatri-Rao product," *CoRR*, vol. abs/1709.05789, 2017. [Online]. Available: http://arxiv.org/abs/1709.05789
[3] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Birkhäuser Basel, 2013.
[4] M. F. Duarte and R. G. Baraniuk, "Kronecker compressive sensing," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 494–504, Feb. 2012.
[5] R. A. Horn and C. R. Johnson, Eds., *Matrix Analysis*. Cambridge University Press, 1986.