On the Spectral Norm of the Pseudo-Inverse for Non-Standard Gaussian Matrices*

Arvind Rathnashyam
RPI CS and Math, rathna@rpi.edu

November 2, 2023

Abstract

In this paper we explore upper bounds on the spectral norm for Gaussian Matrices with columns standard from Central Correlated Multivariate Normal Distributions. As a special case we pick up results when the Covariance Matrix is the Identity. We utilize a lemma from [Chi17, CWS09] and extend the analysis from [CD05]. These bounds find applications in the generalization of the randomized SVD given in [BT22] and wireless network science.

1 Introduction

The study of the expectation of the norms of the pseudoinverse of standard normal gaussian matrices first appeared in [HMT11] when analyzing the error bounds for the Randomized SVD algorithm. The bounds developed in [HMT11] used theory developed in analyzing the condition numbers of standard normal matrices in [CD05]. In a generalization of the Randomized SVD, the need for bounds on the expectation of the spectral norm for correlated Gaussian matrices appears in [BT22].

2 Relevant Work in Standard Uncorrelated Matrices

In this section we will briefly discuss bounds developed for the inequalities of standard normal matrices.

Proposition 1. *(HMT Proposition 10.2). Draw a $k \times (k+p)$ standard Gaussian matrix $\mathbf{G}$ with $k \geq 2$ and $p \geq 2$. Then*

$$\mathbb{E}\left\|\mathbf{G}^{\dagger}\right\| \leq \frac{e\sqrt{k+p}}{p} \tag{1}$$

From our search in the literature, there is no bound on Equation (1) when the columns are not sampled from a multiple of the identity.

3 Theory

We will first introduce the necessary lemmas needed to prove our main results.

3.1 Necessary Lemmas

Lemma 2. *[WLRT08, Lemma 3.6]. Let $m, n \in \mathbb{N}$ s.t. $n \geq m$. Suppose $\mathbf{A} \in \mathbb{R}^{n\times m}$, then if $(\mathbf{A}^\top\mathbf{A})$ is invertible*

$$\left\|\left(\mathbf{A}^\top\mathbf{A}\right)^{-1}\mathbf{A}^\top\right\| = \frac{1}{\sigma_m(\mathbf{A})} \tag{2}$$

*Preliminary Work

Lemma 3. *[Chi17, Lemma 1]. Draw a $m \times n$ matrix $\mathbf{G}$ s.t. the columns of $\mathbf{G}$ are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{C})$ where the eigenvalues of $\mathbf{C}$ are represented as $\sigma_1 > \sigma_2 > \cdots > \sigma_m$. Let $\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{C})$. The eigenvalue distribution is given as*

$$f(x_1, \ldots, x_n) = K_{\mathbf{C}} \left|\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})\right| \cdot \prod_{i=1}^{m-1} \prod_{j=i+1}^{m} (x_i - x_j) \prod_{i=1}^{n} x_i^{n-m} \tag{3}$$

where $\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma}) = \left\{e^{-x_i/\sigma_j}\right\}_{i,j=1}^{m} = \begin{bmatrix} e^{-\frac{x_1}{\sigma_1}} & \ldots & e^{-\frac{x_1}{\sigma_m}} \\ \vdots & \ddots & \vdots \\ e^{-\frac{x_m}{\sigma_1}} & \ldots & e^{-\frac{x_m}{\sigma_m}} \end{bmatrix}$ *and*

$$K_{\mathbf{C}}^{-1} = \prod_{i=1}^{m-1} \prod_{j=i+1}^{m} (\sigma_i - \sigma_j) \prod_{i=1}^{m} \sigma_i^{n-m+1} (n-i)! \tag{4}$$

With these lemmas we will go to proving the main results.

3.2 Main Results

Theorem 4. *Draw a $m \times m$ matrix $\mathbf{G}$ s.t. the columns of $\mathbf{G}$ are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{C})$ where the eigenvalues of $\mathbf{C}$ are represented as $\sigma_1 > \sigma_2 > \cdots > \sigma_m$. Then*

$$\mathbb{E} \left\|\mathbf{G}^{\dagger}\right\| = \sqrt{\pi \operatorname{Tr}\left(\mathbf{C}^{-1}\right)} \tag{5}$$

Proof. We will first note

$$\left\|\mathbf{G}^{\dagger}\right\| \overset{\text{lem. } 2}{=} \frac{1}{\sigma_{\min}(\mathbf{G})} = \frac{1}{\sqrt{\lambda_{\min}(\mathbf{G}\mathbf{G}^{\top})}} \tag{6}$$

For $\mathbf{W}$ sampled from $\mathcal{W}_m(m, \mathbf{C})$. We will now derive the distribution for minimum eigenvalue of $\mathbf{W}$ similar to [NZYY08].

$$f_{\lambda_{\min}}(x_m) = \int_{x_2}^{\infty} \cdots \int_{x_{m-1}}^{\infty} K_{\mathbf{C}} \left|\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})\right| \cdot \prod_{i<j}^{m} (x_i - x_j) \prod_{i=1}^{m} x_j^{m-m} \prod_{i=1}^{m-1} dx_i \tag{7}$$

$$= K_{\mathbf{C}} \int_{x_2}^{\infty} \cdots \int_{x_{m-1}}^{\infty} \left|\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})\right| \cdot \prod_{i=1}^{m-2} \prod_{j=i+1}^{m-1} (x_i - x_j) \prod_{i=1}^{m-1} (x_i - x_m) \prod_{i=1}^{m-1} dx_i \tag{8}$$

$$\overset{\zeta_1}{=} \exp\left(-\sum_{i=1}^{m} \frac{x_m}{\sigma_i}\right) \left(\int_{y_2}^{\infty} \cdots \int_{y_{m-1}}^{\infty} \sum_{i=1}^{m} (-1)^{i+m} K_{\mathbf{C}} \left|\mathbf{E}_i\left(\mathbf{x} - \mathbf{x}_m, \boldsymbol{\sigma}\right)\right| \prod_{i=1}^{m-2} \prod_{j=i+1}^{m-1} (y_i - y_j) \prod_{i-1}^{m-1} dy_i\right) \tag{9}$$

$$\overset{\zeta_2}{=} \Xi \exp\left(-\sum_{i=1}^{m} \frac{x_m}{\sigma_i}\right) \tag{10}$$

(ζ_1) follows due to the properties of the determinant. (ζ_2) follows as the intergral expresion in Equation (9) no longer integrates over x_m and thus integrates to some constant we define as Ξ. Since the PDF must integrate to 1, we thus have,

$$f_{\lambda_{\min}}(x) = \left(\sum_{k=1}^{m} \frac{1}{\sigma_k}\right) \exp\left(-x \sum_{k=1}^{m} \frac{1}{\sigma_k}\right) \tag{11}$$

The Expected Value of the minimum eigenvalue of $\mathbf{W}$ follows from a simple integration.

$$\mathbb{E}\lambda_{\min}\left(\mathbf{W}\right) = \int_0^{\infty} \left(\sum_{k=1}^{m} \frac{1}{\sigma_k}\right) x \exp\left(-x \sum_{k=1}^{m} \sigma_k^{-1}\right) dx = \left(\sum_{k=1}^{m} \frac{1}{\sigma_k}\right)^{-1} = \operatorname{Tr}\left(\mathbf{C}^{-1}\right) \tag{12}$$

The CDF is given as

$$\mathbb{P}\left\{\lambda_{\min}\left(\mathbf{W}\right) < t\right\} = \left(\sum_{k=1}^{m}\frac{1}{\sigma_k}\right)^{-1}\int_0^t \exp\left(-t\sum_{k=1}^{m}\sigma_k^{-1}\right)dt = 1 - \exp\left(-t\sum_{k=1}^{m}\sigma_k^{-1}\right) \tag{13}$$

We can then calculate the expectation of $\left\|\mathbf{G}^{\dagger}\right\|$.

$$\mathbb{E}\left\|\mathbf{G}^{\dagger}\right\| = \left(\sum_{k=1}^{m}\frac{1}{\sigma_k}\right)\int_0^{\infty}\frac{1}{\sqrt{x}}\exp\left(-x\sum_{k=1}^{m}\frac{1}{\sigma_k}\right)dx = \sqrt{\pi\sum_{k=1}^{m}\frac{1}{\sigma_k}} = \sqrt{\pi\,\mathrm{Tr}\left(\mathbf{C}^{-1}\right)} \tag{14}$$

This completes the proof. ■

In our next theorem, we will consider the matrix is rectangle and all the singular values of the covariance matrix are distinct.

Corollary 5. *Draw a $m \times m$ matrix $\mathbf{G}$ s.t. the columns of $\mathbf{G}$ are sampled from from $\mathcal{N}_m(\mathbf{0},\mathbf{I})$. Then it follows,*

$$\mathbb{E}\left\|\mathbf{G}^{\dagger}\right\| = \sqrt{m\pi} \tag{15}$$

Proof. Consider the following covariance matrix.

$$\mathbf{C} = \begin{bmatrix} 1+\xi & & & \\ & 1+2\xi & & \\ & & \ddots & \\ & & & 1+m\xi \end{bmatrix} \tag{16}$$

From Theorem 4, we have

$$\mathbb{E}\left\|\mathbf{G}^{\dagger}\right\| = \lim_{\xi\to 0}\sqrt{\pi\sum_{k=1}^{m}\frac{1}{1+k\xi}} = \sqrt{m\pi} \tag{17}$$

This completes the proof. ■

Theorem 6. *Draw a $m\times n$ matrix $\mathbf{G}$ s.t. the columns of $\mathbf{G}$ are sampled from $\mathcal{N}_m(\mathbf{0},\mathbf{C})$ where the eigenvalues of $\mathbf{C}$ are represented as $\sigma_1 > \sigma_2 > \cdots > \sigma_m > 0$. Let $\mathbf{W} \sim \mathcal{W}_m(n,\mathbf{C})$. Then it follows*

$$\mathbb{E}\left\|\mathbf{G}^{\dagger}\right\| \le \sum_{k=1}^{m}(-1)^{k+m}\frac{\left(\prod_{i=1}^{m}\sigma_i\right)\sqrt{\frac{2\pi}{n-m+\frac{1}{2}}}\left(\frac{n-m+\frac{1}{2}}{e}\right)^{n-m+\frac{1}{2}}}{\prod_{i>k}^{m}(\sigma_i-\sigma_k)\prod_{i<k}^{m}(\sigma_k-\sigma_i)\cdot\sigma_k^{3/2}(n-m)!} \tag{18}$$

Proof. Let $K_{\mathbf{C}}$ and $\mathbf{E}(\mathbf{x},\boldsymbol{\sigma})$ be defined as in Lemma 3.

$$f_{\lambda_{\min}}(x_m) = \int_{x_2}^{\infty}\cdots\int_{x_{m-1}}^{\infty}K_{\mathbf{C}}\left|\mathbf{E}(\mathbf{x},\boldsymbol{\sigma})\right|\cdot\prod_{i<j}^{m}(x_i-x_j)\prod_{i=1}^{m}x_j^{n-m}\prod_{i=1}^{m-1}dx_i \tag{19}$$

$$= K_{\mathbf{C}}x_m^{n-m}\int_{x_2}^{\infty}\cdots\int_{x_{m-1}}^{\infty}\left|\mathbf{E}\left(\mathbf{x},\boldsymbol{\sigma}\right)\right|\cdot\prod_{i<j}^{m-1}\left(x_i-x_j\right)\prod_{i=1}^{m-1}\left(x_i-x_m\right)\prod_{i=1}^{m-1}x_i^{n-m}\prod_{i=1}^{m-1}dx_i \tag{20}$$

$$\le K_{\mathbf{C}}x_m^{n-m}\int_{x_2}^{\infty}\cdots\int_{x_{m-1}}^{\infty}\left|\mathbf{E}\left(\mathbf{x},\boldsymbol{\sigma}\right)\right|\cdot\prod_{i<j}^{m-1}\left(x_i-x_j\right)\prod_{i=1}^{m-1}x_i^{n-m+1}\prod_{i=1}^{m-1}dx_i \tag{21}$$

$$\overset{\zeta_1}{=} K_{\mathbf{C}}x_m^{n-m}\sum_{i=1}^{m}\left((-1)^{i+m}\exp\left(-\frac{x_m}{\sigma_i}\right)\int_{x_2}^{\infty}\cdots\int_{x_{m-1}}^{\infty}\left|\mathbf{E}_{m,i}(\mathbf{x},\boldsymbol{\sigma})\right|\cdot\prod_{i<j}^{m-1}\left(x_i-x_j\right)\prod_{i=1}^{m-1}x_i^{n-m+1}\prod_{i=1}^{m-1}dx_i\right) \tag{22}$$

$$= x_m^{n-m} \underbrace{K_{\mathbf{C}} \sum_{i=1}^{m} (-1)^{i+m} \exp\left(-\frac{x_m}{\sigma_i}\right) K_{\mathbf{C},i}^{-1}}_{\Xi} \tag{23}$$

(ζ_1) follows from a Laplace expansion of the determinant. In Equation (23), we define

$$K_{\mathbf{C},k}^{-1} \triangleq \prod_{i=1, i\neq k}^{m-1} \prod_{j=i+1, j\neq k}^{m} (\sigma_i - \sigma_j) \prod_{i=1, i\neq k}^{m} \sigma_i^{n-m+2} (n-k)! \tag{24}$$

We will now upper bound Ξ.

$$\Xi \triangleq K_{\mathbf{C}} \sum_{i=1}^{m} (-1)^{i+m} \exp\left(-\frac{x_m}{\sigma_i}\right) K_{\mathbf{C},i}^{-1} \tag{25}$$

$$= \sum_{k=1}^{m} (-1)^{k+m} \exp\left(-\frac{x_m}{\sigma_k}\right) \frac{\prod_{i=1, i\neq k}^{m-1} \prod_{j=i+1, j\neq k}^{m} (\sigma_i - \sigma_j) \prod_{i=1, i\neq k}^{m} \sigma_i^{n-m+2} (n-k)!}{\prod_{i=1}^{m-1} \prod_{j=i+1}^{m} (\sigma_i - \sigma_j) \prod_{i=1}^{m} \sigma_i^{n-m+1} (n-k)!} \tag{26}$$

$$= \sum_{k=1}^{m} (-1)^{k+m} \exp\left(-\frac{x_m}{\sigma_k}\right) \left(\frac{\prod_{i=1}^{m} \sigma_i}{\sigma_k}\right) \left(\prod_{i>k} (\sigma_i - \sigma_k) \prod_{i<k} (\sigma_k - \sigma_i) \cdot \sigma_k^{n-m+1} (n-m)!\right)^{-1} \tag{27}$$

We thus have

$$f_{\lambda_{\min}}(x_m) \leq K x_m^{n-m} \sum_{i=1}^{m} (-1)^{k+m} \exp\left(-\frac{x_m}{\sigma_i}\right) \left(\frac{\prod_{i=1}^{m} \sigma_i}{\sigma_k}\right) = \mathcal{O}\left(x_m^{n-m} \sum_{i=1}^{m} \exp\left(-\frac{x_m}{\sigma_i}\right) \left(\frac{\prod_{i=1}^{m} \sigma_i}{\sigma_k}\right)\right) \tag{28}$$

Now we will integrate over $f_{\lambda_{\min}}(x_m)$.

$$\mathbb{E}\left\|\mathbf{G}^{\dagger}\right\| = \int_0^{\infty} (-1)^{k+m} \mathcal{O}\left(x^{n-m-\frac{1}{2}} \sum_{i=1}^{m} \exp\left(-\frac{x}{\sigma_i}\right)\right) dx \tag{29}$$

$$= \sum_{i=1}^{m} (-1)^{k+m} \int_0^{\infty} \mathcal{O}\left(x^{n-m-\frac{1}{2}} \exp\left(-\frac{x}{\sigma_i}\right)\right) dx \tag{30}$$

$$= \sum_{i=1}^{m} (-1)^{k+m} \mathcal{O}\left(\sigma_i^{n-m+\frac{1}{2}} \Gamma\left(n-m+\frac{1}{2}\right)\right) \tag{31}$$

Now we will plug in the constant.

$$\mathbb{E}\left\|\mathbf{G}^{\dagger}\right\| \leq \sum_{k=1}^{m} (-1)^{k+m} \frac{\left(\prod_{i=1}^{m} \sigma_i\right) \sigma_k^{n-m+\frac{1}{2}} \Gamma\left(n-m+\frac{1}{2}\right)}{\prod_{i>k}^{m} (\sigma_i - \sigma_k) \prod_{i<k}^{m} (\sigma_k - \sigma_i) \cdot \sigma_k^{n-m+2} (n-m)!} \tag{32}$$

$$\overset{\zeta_1}{\approx} \sum_{k=1}^{m} (-1)^{k+m} \frac{\left(\prod_{i=1}^{m} \sigma_i\right) \sqrt{\frac{2\pi}{n-m+\frac{1}{2}}} \left(\frac{n-m+\frac{1}{2}}{e}\right)^{n-m+\frac{1}{2}}}{\prod_{i>k}^{m} (\sigma_i - \sigma_k) \prod_{i<k}^{m} (\sigma_k - \sigma_i) \cdot \sigma_k^{3/2} (n-m)!} \tag{33}$$

(ζ_1) folows from an application of Stirling's Approximation of the Gamma Function [Rob55]. Rewriting the denominator gives us the desired result. ■

3.3 Approximations

Theorem 6 gives insight into the complexity of the expected spectral norm of a non-centered Gaussian matrix. The bound is difficult to compute numerically, therefore we will instead give approximations to the bound when there is a known singular value gap.

Theorem 7. [1] *Draw a $m \times n$ matrix $\mathbf{G}$ s.t. the columns of $\mathbf{G}$ are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{C})$ where the eigenvalues of $\mathbf{C}$ are represented as $\sigma_1 > \sigma_2 > \cdots > \sigma_m > 0$ with polynomial or faster spectral decay:*

$$\sigma_j(\mathbf{C}) = \mathcal{O}\left(j^{-\beta}\right) \quad (1 \le j \le m) \tag{34}$$

for some $\beta \in \mathbb{Z}_{++}$. Then it follows

$$\mathbb{E}\left\|\mathbf{G}^\dagger\right\| \le \frac{\Gamma\left(n-m+\frac{1}{2}\right)}{(n-m)!}\left(\sum_{k=1,\text{odd}}^{m} \sigma_k^{-3/2} + (\sigma_k - \sigma_{k+1})\right) \tag{35}$$

Proof. We will start from the results of Theorem 6.

$$\begin{aligned}
\mathbb{E}\left\|\mathbf{G}^\dagger\right\| &\overset{\text{thm. 6}}{\le} \sum_{k=1}^{m}(-1)^{k+m}\frac{\left(\prod_{i=1}^m \sigma_i\right)\Gamma\left(n-m+\frac{1}{2}\right)}{\prod_{i>k}^m(\sigma_i-\sigma_k)\prod_{i<k}^m(\sigma_k-\sigma_i)\cdot\sigma_k^{3/2}(n-m)!} && (36)\\
&= \left(\frac{\Gamma\left(n-m+\frac{1}{2}\right)\prod_{i=1}^m\sigma_i}{(n-m)!}\right)\underbrace{\sum_{i=1}^{m}(-1)^{k+m}\left(\prod_{i>k}^{m}(\sigma_i-\sigma_k)\prod_{i<k}^{m}(\sigma_k-\sigma_i)\cdot\sigma_k^{3/2}\right)^{-1}}_{\text{A}} && (37)
\end{aligned}$$

We have seperated the constant and the term dependent on the singular values of $\mathbf{C}$. Let us assume WLOG m is odd, then

$$\begin{aligned}
\text{A} &= \sum_{k=1,\text{odd}}^{m-1}\left|\left(\prod_{i>k}^{m}(\sigma_i-\sigma_{k+1})\prod_{i<k+1}^{m}(\sigma_{k+1}-\sigma_i)\cdot\sigma_{k+1}^{3/2}\right)^{-1} - \left(\prod_{i>k}^{m}(\sigma_i-\sigma_k)\prod_{i<k}^{m}(\sigma_k-\sigma_i)\cdot\sigma_k^{3/2}\right)^{-1}\right| && (38)\\
&= \sum_{k=1,\text{odd}}^{m-1}\left|\left(\prod_{i=1,i\ne k}^{m}|\sigma_{k+1}-\sigma_i|\cdot\sigma_{k+1}^{3/2}\right)^{-1} - \left(\prod_{i=1,i\ne k}^{m}|\sigma_k-\sigma_i|\cdot\sigma_k^{3/2}\right)^{-1}\right| && (39)\\
&= \sum_{k=1,\text{odd}}^{m-1}\left|\left(\prod_{i=1,i\ne k+1}^{m}|\sigma_{k+1}-\sigma_i|^{-1}\cdot\sigma_{k+1}^{-3/2}\right) - \left(\prod_{i=1,i\ne k}^{m}|\sigma_k-\sigma_i|^{-1}\cdot\sigma_k^{-3/2}\right)\right| && (40)\\
&\overset{\zeta_1}{\le} \sum_{k=1,\text{odd}}^{m-1}\sigma_k^{-3/2}\underbrace{\left|\prod_{i=1,i\ne k}^{m}|\sigma_k-\sigma_i|^{-1} - \prod_{i=1,i\ne k+1}^{m}|\sigma_{k+1}-\sigma_i|^{-1}\right|}_{\text{B}} + \underbrace{\left|\prod_{i=1,i\ne k+1}^{m}|\sigma_{k+1}-\sigma_i|^{-1}\right|}_{\text{C}}\left|\sigma_k^{-3/2}-\sigma_{k+1}^{-3/2}\right| && (41)
\end{aligned}$$

where (ζ_1) follows from the following set of inequalities for any $\alpha, \beta, \gamma, \delta \in \mathbb{R}$,

$$|\alpha\beta - \gamma\delta| = |\alpha\beta - \alpha\delta + \alpha\delta - \gamma\delta| = |\alpha(\beta-\delta) + \delta(\alpha-\gamma)| \le |\alpha|\,|\beta-\delta| + |\delta|\,|\alpha-\gamma| \tag{42}$$

We will first bound C. Let $\mathbf{C}_k \in \mathbb{R}^{m-1\times m-1} = \operatorname{diag}(\sigma_1, \ldots, \sigma_{k-1}, \sigma_{k+1}, \ldots, \sigma_m)$. Then using properties of the determinant, we have

$$\begin{aligned}
\text{B} &= \left|\prod_{i=1,i\ne k}^{m}|\sigma_k-\sigma_i|^{-1} - \prod_{i=1,i\ne k+1}^{m}|\sigma_{k+1}-\sigma_i|^{-1}\right| && (43)\\
&= \left|\left(\prod_{i=1,i\ne k}^{m}|\sigma_k-\sigma_i|\right)^{-1} - \left(\prod_{i=1,i\ne k+1}^{m}|\sigma_{k+1}-\sigma_i|\right)^{-1}\right| && (44)
\end{aligned}$$

[1] In Progress.

$$= \left|\det\left(\mathbf{C}_k - \sigma_k \mathbf{I}\right)^{-1} - \det\left(\mathbf{C}_{k+1} - \sigma_{k+1}\mathbf{I}\right)^{-1}\right| \tag{45}$$

$$= \left|\frac{\det\left(\mathbf{C}_{k+1} - \sigma_{k+1}\mathbf{I}\right) - \det\left(\mathbf{C}_k - \sigma_k \mathbf{I}\right)}{\det\left(\mathbf{C}_k - \sigma_k \mathbf{I}\right)\det\left(\mathbf{C}_{k+1} - \sigma_{k+1}\mathbf{I}\right)}\right| \tag{46}$$

$$= \left|\frac{\det\left(\mathbf{C}_{k+1} - \sigma_{k+1}\mathbf{I}\right) - \det\left(\mathbf{C}_k + \sigma_{k+1}\mathbf{I} - \sigma_{k+1}\mathbf{I} - \sigma_k \mathbf{I}\right)}{\det\left(\mathbf{C}^{k+1}\mathbf{C}_k - \sigma_k \mathbf{C}_k - \sigma_k \mathbf{C}_{k+1} + \sigma_k \sigma_{k+1}\mathbf{I}\right)}\right| \tag{47}$$

$$\leq \frac{\det\left(\mathbf{C}_{k+1} - \sigma_{k+1}\mathbf{I}\right) - \det\left(\mathbf{C}_k - \sigma_{k+1}\mathbf{I}\right) - \det\left(\left(\sigma_{k+1} - \sigma_k\right)\mathbf{I}\right)}{\det\left(\mathbf{C}_k\left(\mathbf{C}_{k+1} - \sigma_k \mathbf{I}\right)\right) + \det\left(\sigma_k\left(\mathbf{C}_{k+1} - \sigma_{k+1}\mathbf{I}\right)\right)} \tag{48}$$

$$= \frac{\det\left(\mathbf{C}_{k+1} - \sigma_{k+1}\mathbf{I}\right) - \left|\sigma_{k+1} - \sigma_k\right|^m}{\sigma_k^m \det\left(\mathbf{C}_{k+1} - \sigma_{k+1}\mathbf{I}\right)} \tag{49}$$

$$= \sigma_k^{-m} - \frac{\left(\sigma_{k+1} - \sigma_k\right)^m}{\sigma_k^m \det\left(\mathbf{C}_{k+1} - \sigma_{k+1}\mathbf{I}\right)} \tag{50}$$

$$= \mathcal{O}\left(k^{m\beta}\right)\left(1 - \frac{\left(\sigma_{k+1} - \sigma_k\right)^m}{\prod_{i=1, i\neq k+1}^{m}\left|\sigma_{k+1} - \sigma_i\right|}\right) \tag{51}$$

$$\leq \mathcal{O}\left(k^{m\beta}\right)\left(1 - \left(\sigma_k - \sigma_{k+1}\right)\left(\frac{\sigma_k - \sigma_{k+1}}{1 - \sigma_{k+1}}\right)^{m-1}\right) \leq \mathcal{O}\left(k^{m\beta}\right) = \sigma_k^{-m} \tag{52}$$

In Equation (51), we see clearly the second term in the parenthesis is less than one as each term in the product of the denominator is greater than the term in the product. We then have as follows

$$\text{Ⓐ} \leq \sum_{k=1,\text{odd}}^{m-1} 2\sigma_k^{-m-3/2} \leq \sum_{k=1}^{m} \sigma_k^{-m-3/2} \tag{53}$$

■

An interesting note we can point here is if we are to use $\beta = 0$ in the result of Theorem 7, we can give bounds on the identity covariance. From here the bound is given as

$$\mathbb{E}\left\|\mathbf{G}^{\dagger}\right\| \leq \frac{\Gamma\left(n - m + \frac{1}{2}\right)m}{(n-m)!} \tag{54}$$

This bound is slightly weaker than the HMT bound asymptotitically.

Theorem 8. [2] *Draw a $m \times n$ matrix $\mathbf{G}$ s.t. the columns of $\mathbf{G}$ are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{C})$ where the eigenvalues of $\mathbf{C}$ are represented as $\sigma_1 > \sigma_2 > \cdots > \sigma_m > 0$ with polynomial or faster spectral decay:*

$$\sigma_{m-j+1}(\mathbf{C}) = \mathcal{O}\left(j^{\beta}\right) \quad (1 \leq j \leq m) \tag{55}$$

for some $\beta \in \mathbb{Z}_{++}$. Then it follows

$$\mathbb{E}\left\|\mathbf{G}^{\dagger}\right\| \leq \Xi \tag{56}$$

Proof.

■

4 Numerical Experiments

We consider diagonal covariance matrices with different singular value decay.

$$\mathbf{\Sigma}_1 = \sum_{k=1}^{n} k^{\ell}\mathbf{e}_k\mathbf{e}_k^{\top} \quad \ell \in \{0, 1, 2\} \tag{57}$$

$$\mathbf{\Sigma}_2 = \sum_{k=1}^{n} k^{-\ell}\mathbf{e}_k\mathbf{e}_k^{\top} \quad \ell \in \{0, 1, 2\} \tag{58}$$

In Figure 1, we verify the results given in Theorem 4. In Figure 2, we verify the results given in Theorem 6. Our results are clearly good in Figure 1. In the rectangular case, we find are results are good when there is

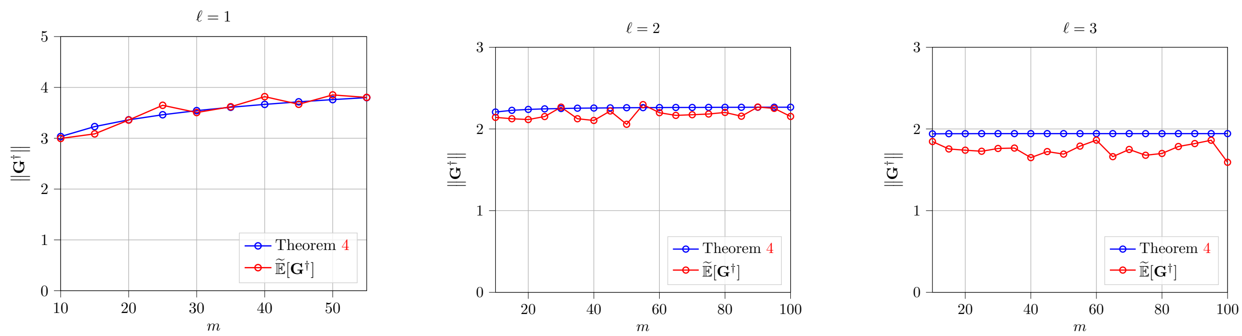

Figure 1: Comparing the expected norm upper bound on $\left\|\mathbf{G}^\dagger\right\|$ where $\mathbf{G} \in \mathbb{R}^{m\times m}$ and the columns of $\mathbf{G}$ are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{K})$ where $\mathbf{K}$ is given by Equation (57). We then calculate the median norm of $\mathbf{G}^\dagger$ over 1000 trials. The expected norm is calculated with Theorem 4.

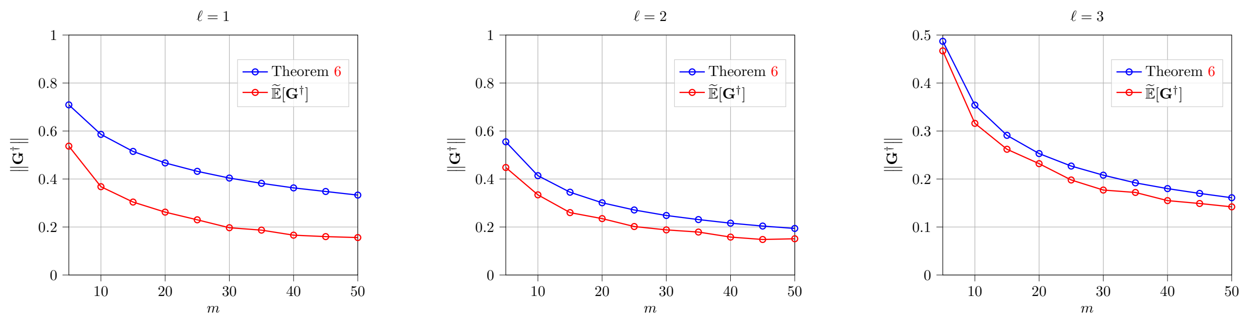

Figure 2: Comparing the expected norm upper bound on $\left\|\mathbf{G}^\dagger\right\|$ where $\mathbf{G} \in \mathbb{R}^{2m\times m}$ and the columns of $\mathbf{G}$ are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{K})$ where $\mathbf{K}$ is given by Equation (57). We then calculate the median norm of $\mathbf{G}^\dagger$ over 100 trials. The expected norm is calculated with Theorem 6.

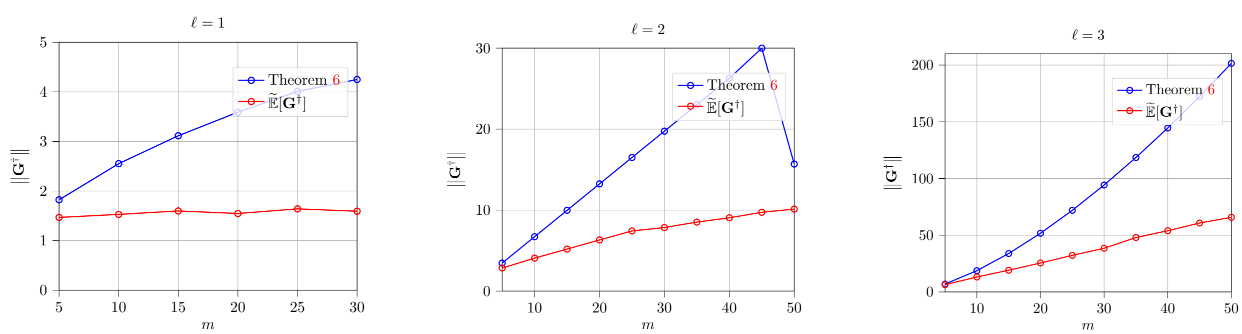

Figure 3: Comparing the expected norm upper bound on $\left\|\mathbf{G}^\dagger\right\|$ where $\mathbf{G} \in \mathbb{R}^{2m\times m}$ and the columns of $\mathbf{G}$ are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{K})$ where $\mathbf{K}$ is given by Equation (58). We then calculate the median norm of $\mathbf{G}^\dagger$ over 100 trials. The expected norm is calculated with Theorem 6.

significant singular decay. When there is insufficient singular value decay and the values are small, e.g. the set up in Corollary 5 we find the upper bound becomes close to infinite due to the denominator and upper bound becomes meaniningless. [3]

[2] In Progress.

[3] Code for experiments will be released on Github.

5 Conclusions

In this paper, we derive novel upper bounds for the spectral norm of Gaussian matrices with columns sampled from a central correlated multivariate normal distribution with various distributions of the singular values of the covariance matrix. On the side, we pick up the expected spectral norm of the Pseudoinverse of the Standard Gaussian matrix. The results for the square case are clean and strong, the rectangular case is quite strong but not as clean. To that end, for the case when the spectrum has polynomial decay, we then derive a weaker yet more interpretable upper bound. We have shown in experiments are results are good and we believe they can find use in applications.

Acknowledgements

Thank you to Christopher Wang (Cornell PhD Student) for the idea of strengthening the bounds for the Generalized Randomized SVD by extending Random Matrix Theory results in the HMT Paper to Non-Standard Gaussian Matrices. One such extension inspired this work.

References

[BT22] Nicolas Boullé and Alex Townsend. A generalization of the randomized singular value decomposition. In *International Conference on Learning Representations*, 2022.

[CD05] Zizhong Chen and Jack J. Dongarra. Condition numbers of gaussian random matrices. *SIAM Journal on Matrix Analysis and Applications*, 27(3):603–620, 2005.

[Chi17] Marco Chiani. On the probability that all eigenvalues of gaussian, wishart, and double wishart random matrices lie within an interval, 2017.

[CWS09] Marco Chiani, Moe Z Win, and Hyundong Shin. Mimo networks: The effects of interference. *IEEE Transactions on information theory*, 56(1):336–349, 2009.

[HMT11] N. Halko, P. G. Martinsson, and J. A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53(2):217–288, 2011.

[NZYY08] Fangfang Niu, Haochuan Zhang, Hongwen Yang, and Dacheng Yang. Distribution of the smallest eigenvalue of complex central semi-correlated wishart matrices. In *2008 IEEE International Symposium on Information Theory*, pages 1788–1792, 2008.

[Rob55] Herbert Robbins. A remark on stirling's formula. *The American mathematical monthly*, 62(1):26–29, 1955.

[WLRT08] Franco Woolfe, Edo Liberty, Vladimir Rokhlin, and Mark Tygert. A fast randomized algorithm for the approximation of matrices. *Applied and Computational Harmonic Analysis*, 25(3):335–366, 2008.