

Upper Bounds on the Spectral Norm of the Pseudo-Inverse of Non-Standard Gaussian Matrices

Arvind Rathnashyam
RPI CS and Math
rathna@rpi.edu

October 18, 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. 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

$$\mathbb{E} \|\mathbf{G}^\dagger\| \leq \frac{e\sqrt{k+p}}{p} \quad (1)$$

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. [Jam64, Eq. (58,59)]. If $\lambda_1 \geq \dots \geq \lambda_m$ are the eigenvalues of \mathbf{W} s.t. $\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{C})$ s.t. $n > m - 1$, then the joint PDF of eigenvalues is

$$f(\lambda_1, \dots, \lambda_m) = K_{m,n} (\det \mathbf{C})^{-n/2} \exp\left(-\frac{1}{2} \text{Tr}(\mathbf{C}^{-1} \mathbf{W})\right) \prod_{i=1}^m \lambda_i^{(n-m-1)/2} \prod_{i < j} (\lambda_i - \lambda_j) \quad (2)$$

where

$$K_{m,n} = \frac{\pi^{m^2/2}}{\Gamma_m(\frac{1}{2}m) \Gamma_m(\frac{1}{2}n)} \quad (3)$$

Lemma 3. [WLR08, 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

$$\left\| (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \right\| = \frac{1}{\sigma_m(\mathbf{A})} \quad (4)$$

Lemma 4. [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 $\sigma_1 > \sigma_2 > \dots > \sigma_m$. Let $\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{C})$. The eigenvalue distribution is given as

$$f(x_1, \dots, x_n) = K_{\mathbf{C}} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i=1}^{m-1} \prod_{j=i+1}^m (x_i - x_j) \prod_{i=1}^n x_i^{n-m} \quad (5)$$

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

$$K_{\mathbf{C}}^{-1} = \prod_{i=1}^{m-1} \prod_{j=i+1}^m (\sigma_i - \sigma_j) \prod_{i=1}^m \sigma_i^{n-m+1} (n-i)! \quad (6)$$

Theorem 5. Consider a sequence of independent random matrices, $\mathbf{X}_k \in \mathbb{R}^{n \times m}$, such that

$$\mathbb{E}[\mathbf{X}_k] = \mathbf{0} \quad (7)$$

$$\|\mathbf{X}_k\| \leq R \quad \forall k \quad (8)$$

$$\nu \triangleq \max \left\{ \left\| \mathbb{E} \sum_k \mathbf{X}_k \mathbf{X}_k^\top \right\|, \left\| \mathbb{E} \sum_k \mathbf{X}_k^\top \mathbf{X}_k \right\| \right\} \quad (9)$$

Then we have for all $t \geq 0$,

$$\mathbb{P} \left\{ \left\| \sum_k \mathbf{X}_k \right\| \geq t \right\} \leq (m+) \exp \left(\frac{-t^2/2}{\nu + Rt/3} \right) \quad (10)$$

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

3.2 Main Results

Theorem 6. 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 $\sigma_1 > \sigma_2 > \dots > \sigma_m$. Then

$$\mathbb{E} \|\mathbf{G}^\dagger\| \geq \left(\sum_{k=1}^m \frac{1}{\sigma_k} \right)^{-1} \quad (11)$$

Proof. We will first note

$$\|\mathbf{G}^\dagger\| \stackrel{\text{lem. 3}}{=} \frac{1}{\sigma_{\min}(\mathbf{G})} = \frac{1}{\sqrt{\sigma_{\min}(\mathbf{G}\mathbf{G}^\top)}} = \frac{1}{\lambda_{\min}(\mathbf{G}\mathbf{G}^\top)} \quad (12)$$

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 [NZZY08].

$$f_{\lambda_{\min}}(x_m) = \int_{x_2}^{\infty} \dots \int_{x_{m-1}}^{\infty} K_{\mathbf{C}} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i < j}^m (x_i - x_j) \prod_{i=1}^m x_j^{m-m} \prod_{i=1}^{m-1} dx_i \quad (13)$$

$$= K_{\mathbf{C}} \int_{x_2}^{\infty} \cdots \int_{x_{m-1}}^{\infty} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i=1}^{m-2} \prod_{j=i+1}^{m-1} (x_i - x_j) \prod_{i=1}^{m-1} (x_i - x_m) \prod_{i=1}^{m-1} dx_i \quad (14)$$

$$\stackrel{\zeta_1}{=} \exp \left(- \sum_{i=1}^m \frac{x_m}{\sigma_i} \right) \left(\int_{y_2}^{\infty} \cdots \int_{y_{m-1}}^{\infty} \sum_{i=1}^m (-1)^{i+m} K_{\mathbf{C}} |\mathbf{E}_i(\mathbf{x} - \mathbf{x}_m, \boldsymbol{\sigma})| \prod_{i=1}^{m-2} \prod_{j=i+1}^{m-1} (y_i - y_j) \prod_{i=1}^{m-1} dy_i \right) \quad (15)$$

$$\stackrel{\zeta_2}{=} \Xi \exp \left(- \sum_{i=1}^m \frac{x_m}{\sigma_i} \right) \quad (16)$$

(ζ_1) follows due to the properties of the determinant. (ζ_2) follows as the intergral expression in Equation (15) 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,

$$f_{\lambda_{\min}}(x) = \left(\sum_{k=1}^m \frac{1}{\sigma_k} \right) \exp \left(-x \sum_{k=1}^m \frac{1}{\sigma_k} \right) \quad (17)$$

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

$$\mathbb{E} \lambda_{\min}(\mathbf{W}) = \int_0^{\infty} \left(\sum_{k=1}^m \frac{1}{\sigma_k} \right) x \exp \left(-x \sum_{k=1}^m \sigma_k^{-1} \right) dx = \left(\sum_{k=1}^m \frac{1}{\sigma_k} \right)^{-1} \quad (18)$$

The CDF is given as

$$\mathbb{P} \{ \lambda_{\min}(\mathbf{W}) < t \} = \left(\sum_{k=1}^m \frac{1}{\sigma_k} \right)^{-1} \int_0^t \exp \left(-t \sum_{k=1}^m \sigma_k^{-1} \right) dt = 1 - \exp \left(-t \sum_{k=1}^m \sigma_k^{-1} \right) \quad (19)$$

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

$$\mathbb{E} \|\mathbf{G}^\dagger\| = \int_0^{\infty} \mathbb{P} \{ \|\mathbf{G}^\dagger\| > t \} dt = \int_0^{\infty} \mathbb{P} \left\{ \frac{1}{\lambda_{\min}(\mathbf{G}\mathbf{G}^\top)} > t \right\} dt \quad (20)$$

$$= \int_0^{\infty} \mathbb{P} \left\{ \lambda_{\min}(\mathbf{W}) < \frac{1}{t} \right\} dt \quad (21)$$

$$= \int_0^{\infty} 1 - \exp \left(-\frac{1}{t} \sum_{k=1}^m \sigma_k^{-1} \right) dt \quad (22)$$

$$= \infty \quad (23)$$

It therefore follows in expectation the pseudo inverse does not exist when $m = n$. The best we can give is a lower bound.

$$\mathbb{E} \|\mathbf{G}^\dagger\| = \mathbb{E} \left[\frac{1}{\lambda_{\min}(\mathbf{W})} \right] \stackrel{\text{Jensen}}{\geq} \frac{1}{\mathbb{E} \lambda_{\min}(\mathbf{W})} = \left(\sum_{k=1}^m \frac{1}{\sigma_k} \right) = \text{Tr}(\mathbf{C}^{-1}) \quad (24)$$

■

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

Theorem 7. 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

$$\mathbb{E} \|\mathbf{G}^\dagger\| \leq \sum_{i=1}^m |\mathbf{C}^2| \left(\prod_{i>k} (\sigma_i - \sigma_k) \prod_{i<k} (\sigma_k - \sigma_i) \cdot \sigma_i^2 \right)^{-1} \quad (25)$$

Proof. Let $K_{\mathbf{C}}$ and $\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})$ be defined as in Lemma 4.

$$f_{\lambda_{\min}}(x_m) = \int_{x_2}^{\infty} \cdots \int_{x_{m-1}}^{\infty} K_{\mathbf{C}} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i<j}^m (x_i - x_j) \prod_{i=1}^m x_j^{n-m} \prod_{i=1}^{m-1} dx_i \quad (26)$$

$$= K_{\mathbf{C}} x_m^{n-m} \int_{x_2}^{\infty} \cdots \int_{x_{m-1}}^{\infty} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i < j}^{m-1} (x_i - x_j) \prod_{i=1}^{m-1} (x_i - x_m) \prod_{i=1}^{m-1} x_i^{n-m} \prod_{i=1}^{m-1} dx_i \quad (27)$$

$$\leq K_{\mathbf{C}} x_m^{n-m} \int_{x_2}^{\infty} \cdots \int_{x_{m-1}}^{\infty} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i < j}^{m-1} (x_i - x_j) \prod_{i=1}^{m-1} x_i^{n-m+1} \prod_{i=1}^{m-1} dx_i \quad (28)$$

$$= K_{\mathbf{C}} x_m^{n-m} \sum_{i=1}^m \left((-1)^{i+m} \exp\left(-\frac{x_m}{\sigma_i}\right) \int_{x_2}^{\infty} \cdots \int_{x_{m-1}}^{\infty} |\mathbf{E}_{m,i}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i < j}^{m-1} (x_i - x_j) \prod_{i=1}^{m-1} x_i^{n-m+1} \prod_{i=1}^{m-1} dx_i \right) \quad (29)$$

$$= x_m^{n-m} K_{\mathbf{C}} \underbrace{\sum_{i=1}^m (-1)^{i+m} \exp\left(-\frac{x_m}{\sigma_i}\right) K_{\mathbf{C},i}^{-1}}_{\Xi} \quad (30)$$

We will now upper bound Ξ .

$$\Xi \triangleq K_{\mathbf{C}} \sum_{i=1}^m (-1)^{i+m} \exp\left(-\frac{x_m}{\sigma_i}\right) K_{\mathbf{C},i}^{-1} \quad (31)$$

$$= \sum_{k=1}^m (-1)^{k+m} \exp\left(-\frac{x_m}{\sigma_k}\right) \frac{\prod_{i=1}^{m-1} \prod_{j=i+1}^m \mathbb{I}\{i, j \neq k\} (\sigma_i - \sigma_j) \prod_{i=1}^m \mathbb{I}\{i \neq k\} \sigma_i^{n-(m-1)+1} (n-i)!}{\prod_{i=1}^{m-1} \prod_{j=i+1}^m (\sigma_i - \sigma_j) \prod_{i=1}^m \sigma_i^{n-m+1} (n-i)!} \quad (32)$$

$$= \sum_{k=1}^m (-1)^{k+m} \exp\left(-\frac{x_m}{\sigma_k}\right) \left(\frac{|\mathbf{C}^2|}{\sigma_k}\right) \left(\prod_{i>k}^m (\sigma_i - \sigma_k) \prod_{i<k}^m (\sigma_k - \sigma_i) \cdot \sigma_k^{n-m+1} (n-k)!\right)^{-1} \quad (33)$$

$$\leq \sum_{k=1}^m \exp\left(-\frac{x_m}{\sigma_k}\right) \left(\frac{|\mathbf{C}^2|}{\sigma_k}\right) \underbrace{\left(\prod_{i>k}^m (\sigma_i - \sigma_k) \prod_{i<k}^m (\sigma_k - \sigma_i) \cdot \sigma_k^{n-m+1} (n-k)!\right)^{-1}}_K \quad (34)$$

Now we will lower bound K for $i = k$. Define $\delta_k \triangleq \min\{(\sigma_k - \sigma_{k+1}), (\sigma_{k-1} - \sigma_k)\}$, then

$$K \triangleq \prod_{i>k}^m (\sigma_i - \sigma_k) \prod_{i<k}^m (\sigma_k - \sigma_i) \cdot \sigma_k^{n-m+1} (n-k)! \quad (35)$$

$$\geq \delta_k^m \sigma_k^{n-m+1} (n-k)! \quad (36)$$

We thus have

$$f_{\lambda_{\min}}(x_m) \leq K x_m^{n-m} \sum_{i=1}^m \exp\left(-\frac{x_m}{\sigma_i}\right) \left(\frac{|\mathbf{C}^2|}{\sigma_k}\right) = \mathcal{O}\left(x_m^{n-m} \sum_{i=1}^m \exp\left(-\frac{x_m}{\sigma_i}\right) \left(\frac{|\mathbf{C}^2|}{\sigma_k}\right)\right) \quad (37)$$

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

$$\mathbb{E} \|\mathbf{G}^\dagger\| = \int_0^\infty \mathcal{O}\left(x^{n-m-1} \sum_{i=1}^m \exp\left(-\frac{x}{\sigma_i}\right)\right) dx \quad (38)$$

$$= \sum_{i=1}^m \mathcal{O}\left(\sigma_i^{n-m} (n-m)!\right) \quad (39)$$

Now we will plug in the lower bound for K from Equation (36).

$$\mathbb{E} \|\mathbf{G}^\dagger\| \leq \sum_{i=1}^m \frac{|\mathbf{C}^2| \sigma_i^{n-m} (n-m)!}{\prod_{i>k}^m (\sigma_i - \sigma_k) \prod_{i<k}^m (\sigma_k - \sigma_i) \cdot \sigma_i^{n-m+2} (n-i)!} \quad (40)$$

$$= \sum_{i=1}^m \frac{|\mathbf{C}^2| (n-m)!}{\prod_{i>k}^m (\sigma_i - \sigma_k) \prod_{i<k}^m (\sigma_k - \sigma_i) \cdot \sigma_i^2 (n-i)!} \quad (41)$$

$$\leq \sum_{i=1}^m \frac{|\mathbf{C}^2|}{\prod_{i>k}^m (\sigma_i - \sigma_k) \prod_{i<k}^m (\sigma_k - \sigma_i) \cdot \sigma_i^2} \quad (42)$$

■

Alternative Proof. First, let us represent $\mathbf{W} = \sum_{i=1}^n \mathbf{x}\mathbf{x}^\top$ where $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \Sigma)$. Then we can lower $\sigma_{\min}(\mathbf{W})$,

$$\sigma_{\min}(\mathbf{W}) = \sigma_{\min} \left(\sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top \right) \quad (43)$$

$$= \sigma_{\min} \left(n\mathbb{E}[\mathbf{x}\mathbf{x}^\top] + \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top - \mathbb{E}[\mathbf{x}_i \mathbf{x}_i^\top] \right) \quad (44)$$

$$\geq n\sigma_{\min}(\Sigma) - \sigma_{\max} \left(\sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top - \Sigma \right) \quad (45)$$

$$= n\sigma_{\min}(\Sigma) - \sigma_{\max} \left(\sum_{i=1}^n \left(\Sigma^{1/2} \mathbf{v}_i \right) \left(\Sigma^{1/2} \mathbf{v}_i \right)^\top - \Sigma \right) \quad (46)$$

$$= n\sigma_{\min}(\Sigma) - \sigma_{\max} \left(\Sigma^{1/2} \left(\sum_{i=1}^n \mathbf{v}_i \mathbf{v}_i^\top - \mathbf{I} \right) \left(\Sigma^{1/2} \right)^\top \right) \quad (47)$$

$$\geq n\sigma_{\min}(\Sigma) - \underbrace{\sigma_{\max}(\Sigma) \sigma_{\max} \left(\sum_{i=1}^n \mathbf{v}_i \mathbf{v}_i^\top - \mathbf{I} \right)}_A \quad (48)$$

There has been significant theory in Random Matrix Theory and High Dimensional Probability analyzing Covariance Estimation, especially in the standard normal case, [T⁺15, Ver20, Rig15]. We will utilize the Matrix Bernstein inequality to upper bound A in probability and in expectation. To use Matrix Bernstein we need to upper bound $\mathbb{V}(\mathbf{v}_i \mathbf{v}_i^\top - \mathbf{I})$.

$$\mathbb{V}(\mathbf{v}\mathbf{v}^\top - \mathbf{I}) = \mathbb{E} \left[(\mathbf{v}\mathbf{v}^\top - \mathbf{I})^\top (\mathbf{v}\mathbf{v}^\top - \mathbf{I}) \right] = \mathbb{E} [\mathbf{v}\mathbf{v}^\top \mathbf{v}\mathbf{v}^\top - 2\mathbf{v}^\top + \mathbf{I}] \quad (49)$$

$$= \mathbb{E} [\mathbf{v}\mathbf{v}^\top \mathbf{v}\mathbf{v}^\top] - \mathbf{I} = \mathbb{E} [\|\mathbf{v}\|^2 \mathbf{v}\mathbf{v}^\top] - \mathbf{I} \quad (50)$$

$$\leq (mK) \mathbb{E} [\mathbf{v}\mathbf{v}^\top] - \mathbf{I} = (mK - 1) \mathbf{I} \quad (51)$$

Then from Theorem 5, we have

$$\mathbb{E}[A] \leq \sqrt{2(mK - 1)n \log(2m)} + \frac{1}{3}mK \log(2d) \quad (52)$$

and we have in probability,

$$\mathbb{P} \left\{ \sigma_{\max} \left(\sum_{i=1}^n \mathbf{v}_i \mathbf{v}_i^\top - \mathbf{I} \right) \geq t \right\} \leq 2m \exp \left(\frac{-t^2/2}{2(mK - 1)n \log(2m) + mK/3} \right) \quad (53)$$

Therefore we have in expectation,

$$\mathbb{E}\sigma_{\min}(\mathbf{G}) \geq n\sigma_{\min}(\Sigma) - \sigma_{\max}(\Sigma) \left(\sqrt{2(mK - 1)n \log(2m)} + \frac{1}{3}mK \log(2m) \right) \quad (54)$$

4 Numerical Experiments

We consider diagonal covariance matrices with different singular value decay.

$$\Sigma = \sum_{k=1}^n k^\ell \mathbf{e}_k \mathbf{e}_k^\top \quad \ell \in \{0, 1, 2\} \quad (55)$$

In Figure 1, we verify the results given in Theorem 6.

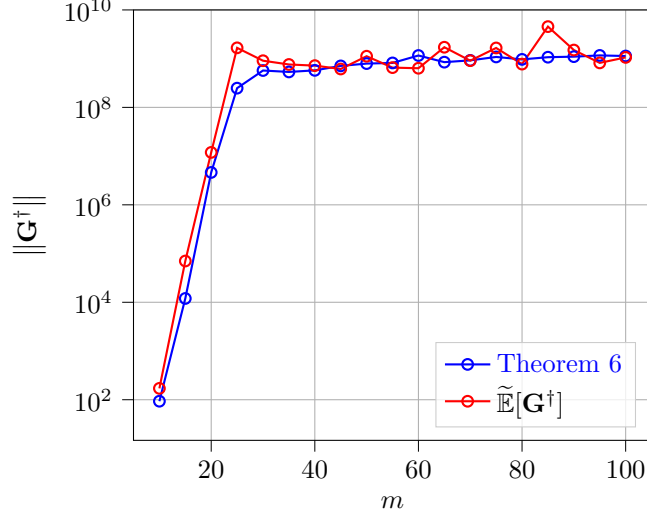


Figure 1: Comparing the expected norm upper bound on $\|\mathbf{G}^\dagger\|$ where $\mathbf{G} \in \mathbb{R}^{m \times m}$ and the columns of \mathbf{G} are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{K})$ with the average norm of \mathbf{G}^\dagger over 100 samples. The expected norm is calculated with Proposition 6.

In Figure 2, we verify the results given in Theorem 7.

Figure 2: Comparing the expected norm upper bound on $\|\mathbf{G}^\dagger\|$ where $\mathbf{G} \in \mathbb{R}^{m \times m}$ and the columns of \mathbf{G} are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{K})$ with the average norm of \mathbf{G}^\dagger over 100 samples. The expected norm is calculated with Proposition 6.

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.

References

- [BT22] Nicolas Boule 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.
- [Jam64] Alan T. James. Distributions of matrix variates and latent roots derived from normal samples. *The Annals of Mathematical Statistics*, 35(2):475–501, 1964.
- [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.
- [Rig15] Philippe Rigollet. 18. s997: High dimensional statistics. *Lecture Notes*, Cambridge, MA, USA: MIT Open-CourseWare, 2015.
- [Rob55] Herbert Robbins. A remark on stirling’s formula. *The American mathematical monthly*, 62(1):26–29, 1955.
- [T⁺15] Joel A Tropp et al. An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2):1–230, 2015.
- [Ver20] Roman Vershynin. High-dimensional probability. *University of California, Irvine*, 2020.
- [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.