A simple proof for the almost sure convergence of the largest singular value of a product of Gaussian matrices

Thiziri Nait Saada* and Alireza Naderi*

Mathematical Institute, University of Oxford

Abstract

Let $m \geq 1$ and consider the product of m independent $n \times n$ Gaussian matrices $\mathbf{W} = \mathbf{W}_1 \dots \mathbf{W}_m$, each $\mathbf{W}_i$ with i.i.d. normalised $\mathcal{N}(0, n^{-1/2})$ entries. It is shown in [PŻ11] that the empirical distribution of the squared singular values of $\mathbf{W}$ converges to a deterministic distribution compactly supported on $[0, u_m]$, where $u_m \coloneqq \frac{(m+1)^{m+1}}{m^m}$. This generalises the well-known case of $m = 1$, corresponding to the Marchenko-Pastur distribution for square matrices. Moreover, for $m = 1$, it was first shown by [Gem80] that the largest squared singular value almost surely converges to the right endpoint (aka "soft edge") of the support, i.e. $s_1^2(\mathbf{W}) \xrightarrow{a.s.} u_1$. Herein, we present a proof for the general case $s_1^2(\mathbf{W}) \xrightarrow{a.s.} u_m$ when $m \geq 1$. Although we do not claim novelty for our result, the proof is simple and does not require familiarity with modern techniques of free probability.

Theorem. *Let $m \geq 1$. Consider $\mathbf{W}_1, \dots, \mathbf{W}_m \in \mathbb{R}^{n \times n}$ be independent Gaussian matrices with i.i.d. $\mathcal{N}(0, n^{-1/2})$ entries. Then, almost surely,*

$$s_1^2(\mathbf{W}_1 \dots \mathbf{W}_m) \xrightarrow{n \to \infty} u_m. \tag{1}$$

Proof. We denote by $\mathbf{W}$ the product $\mathbf{W}_1 \dots \mathbf{W}_m$. That almost surely

$$\varliminf_{n \to \infty} s_1^2(\mathbf{W}) \geq u_m$$

is straightforward. Suppose otherwise that $s_1^2(\mathbf{W}) = \max_i s_i^2(\mathbf{W}) < u_m$ for infinitely many n and there will be a subsequence of empirical distributions whose supports are strictly contained in $[0, u_m]$, which is a contradiction. Now we must show that

$$\varlimsup_{n \to \infty} s_1^2(\mathbf{W}) \leq u_m$$

*Equal contribution.

almost surely. Fix $z > u_m$. Following Geman's strategy, we demonstrate $\mathbb{P}\big(\overline{\lim}_{n\to\infty} s_1^2(\mathbf{W}) \geq z\big) = 0$ using Borel-Cantelli lemma. For any $k \geq 1$,

$$\mathbb{P}\Big(s_1^2(\mathbf{W}) \geq z\Big) = \mathbb{P}\Big(\big(\frac{s_1^2(\mathbf{W})}{z}\big)^k \geq 1\Big) \leq \mathbb{E}\Big(\big(\frac{s_1^2(\mathbf{W})}{z}\big)^k\Big)$$

by Markov's inequality. Therefore, to exhibit an almost sure convergence, it suffices to show

$$\sum_n \mathbb{E}\Big(\big(\frac{s_1^2(\mathbf{W})}{z}\big)^k\Big) < \infty,$$

for some k. We will shortly see how we can carefully choose $k = k_n$ to make the above sum converge. Namely, the remainder of the proof is devoted to establishing

$$\sum_n \mathbb{E}\Big(\big(\frac{s_1^2(\mathbf{W})}{z}\big)^{k_n}\Big) < \infty, \tag{2}$$

where,

$$k_n = \lceil w \log n \rceil,$$

and

$$w > \frac{3}{\log(z/u_m)}.$$

We may simply bound a term of the series in (2) by the k_n-th moment of the empirical (non-limiting) distribution μ_n of the squared singular values of $\mathbf{W}$, i.e.

$$\begin{aligned}\mathbb{E}\Big(\big(s_1^2(\mathbf{W})\big)^{k_n}\Big) &\leq \mathbb{E}\Big(\sum_{i=1}^n \big(s_i^2(\mathbf{W})\big)^{k_n}\Big)\\ &= n\mathbb{E}\Big(\frac{1}{n}\sum_{i=1}^n \big(s_i^2(\mathbf{W})\big)^{k_n}\Big)\\ &= nG(m,n,k_n),\end{aligned}$$

where $G(m,n,k_n)$ is the k_n-th moment of μ_n. We borrow the computation of $G(m,n,k_n)$ from [AIK13, Eq. (58)], where it is worked out for the general case of the product of rectangular Gaussian matrices.[1] The calculation from [AIK13]

[1] In the case where all matrices are square, it is well known that $G(m,n,k)$ converges to the Fuss-Catalan number,

$$\mathrm{FC}_m(k) := \frac{1}{mk+1}\binom{mk+k}{k},$$

for any fixed k and m; see [PŻ11].

provides us with the following non-asymptotic formula, valid for any integer $k \geq 1$,

$$n^{mk+1} G(m, n, k) = \sum_{i=0}^{n-1} \frac{(-1)^{1+i} \prod_{j=0}^{n-1} (j - k - i)}{i!(n-1-i)!k} \times \Big(\frac{\Gamma(k+i+1)}{\Gamma(i+1)} \Big)^m,$$

which can be further simplified as,

$$\begin{aligned}
n^{mk+1} G(m, n, k) &= \frac{1}{k!} \sum_{i=n-k}^{n-1} (-1)^{n+1+i} \Big(\frac{(k+i)!}{i!} \Big)^{m+1} \binom{k-1}{k+i-n} \\
&= \frac{1}{k!} \sum_{j=0}^{k-1} (-1)^{k+1-j} \Big(\frac{(n+j)!}{(n+j-k)!} \Big)^{m+1} \binom{k-1}{j} \\
&= \frac{1}{k!} \sum_{j=0}^{k-1} (-1)^{k+1-j} \binom{k-1}{j} \\
&\qquad \times \big((n+j)(n+j-1) \dots (n+j-k+1) \big)^{m+1} \\
&= \frac{n^{k(m+1)}}{k!} \sum_{j=0}^{k-1} (-1)^{k+1-j} \binom{k-1}{j} \\
&\qquad \times \big((1 + \frac{j}{n})(1 - \frac{1}{n} + \frac{j}{n}) \dots (1 - \frac{k-1}{n} + \frac{j}{n}) \big)^{m+1}.
\end{aligned}$$

We now introduce β_r as the coefficient of x^r in the expansion of the polynomial $P(x) \coloneqq \prod_{i=0}^{k-1} (1 - \frac{i}{n} + x)^{m+1}$. Let $\mathcal{R}$ be the multiset of the $k(m+1)$ roots of P (counted with multiplicity), then each β_r can be explicitly written as,

$$\beta_r = \sum_{\substack{S \subseteq \mathcal{R} \\ |S| = k(m+1) - r}} \prod_{i \in S} (-i).$$

Provided $k \leq n$, all roots of P are negative with magnitude in $[1 - \frac{k-1}{n}, 1]$. Therefore, for all $0 \leq r \leq k(m+1)$, we have

$$\binom{k(m+1)}{r} (1 - \frac{k-1}{n})^{k(m+1)-r} \leq \beta_r \leq \binom{k(m+1)}{r},$$

which yields the asymptotic equivalence[2]

$$\beta_r \sim \binom{k(m+1)}{r}, \tag{3}$$

[2]In the following, we write $f(n) \sim g(n)$ whenever $\lim_{n \to \infty} f(n)/g(n) = 1$.

whenever $k^2 = o(n)$. Substituting back into the last formula, we get

$$\begin{aligned} n^{mk+1}G(m,n,k) &= \frac{n^{k(m+1)}}{k!}\sum_{j=0}^{k-1}(-1)^{k+1-j}\binom{k-1}{j}\sum_{r=0}^{k(m+1)}\beta_r\left(\frac{j}{n}\right)^r \\ &= \frac{n^{k(m+1)}}{k!}\sum_{r=0}^{k(m+1)}n^{-r}\beta_r\sum_{j=0}^{k-1}(-1)^{k+1-j}\binom{k-1}{j}j^r. \end{aligned}$$

The above alternating sums are known to be equal to Stirling numbers of the second kind $\left\{ {n \atop k} \right\}$, defined as the number of ways to partition a set of n objects into k non-empty subsets. They are given by

$$\left\{ {n \atop k} \right\} = \frac{1}{k!}\sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}i^n.$$

Thus, we can rewrite our original quantity as

$$n^{mk+1}G(m,n,k) = \frac{n^{k(m+1)}(k-1)!}{k!}\sum_{r=0}^{k(m+1)}n^{-r}\beta_r\left\{ {r \atop k-1} \right\}.$$

By definition, it is clear that $\left\{ {n \atop k} \right\} = 0$ for $k > n$, therefore,

$$n^{mk+1}G(m,n,k) = \frac{n^{k(m+1)}}{k}\sum_{r=k-1}^{k(m+1)}n^{-r}\beta_r\left\{ {r \atop k-1} \right\}. \tag{4}$$

We now move on to showing that for our specific choice of $k = k_n$, the latter sum is asymptotically dominated by its first term. To this end, we demonstrate that each term in the sum is dominated by the term immediately preceding it. It is sufficient to show, for any $k-1 \le r \le km+k-1$,

$$\frac{\beta_{r+1}\left\{ {r+1 \atop k-1} \right\}n^{-(r+1)}}{\beta_r\left\{ {r \atop k-1} \right\}n^{-r}} < Cn^{-\varepsilon}, \tag{5}$$

for some $\varepsilon > 0$ and C independent of n.

First, observe the recurrence relation satisfied by Stirling numbers of the second kind, which immediately gives,

$$\frac{\left\{ {r+1 \atop k-1} \right\}}{\left\{ {r \atop k-1} \right\}} = k-1+\frac{\left\{ {r \atop k-2} \right\}}{\left\{ {r \atop k-1} \right\}}. \tag{6}$$

Among other properties, Stirling numbers of the second kind are shown in [Lie68] to be logarithmically concave, meaning,

$$\left\{ {n \atop k} \right\}^2 \ge \left\{ {n \atop k+1} \right\}\left\{ {n \atop k-1} \right\},$$

for any $k = 2, \ldots, n-1$. This directly implies that,

$$\frac{\left\{ n \atop k-1 \right\}}{\left\{ n \atop k \right\}} \leq \frac{\left\{ n \atop k \right\}}{\left\{ n \atop k+1 \right\}} \leq \cdots \leq \frac{\left\{ n \atop n-1 \right\}}{\left\{ n \atop n \right\}} = \binom{n}{2}.$$

Back to our problem, we can thus bound the ratio in Eq. (6) as follows:

$$\frac{\left\{ r+1 \atop k-1 \right\}}{\left\{ r \atop k-1 \right\}} = k - 1 + \binom{r}{2} \leq \frac{1}{2} r(r+1).$$

To bound the ratio $\frac{\beta_{r+1}}{\beta_r}$, since $k_n = \lceil w \log(n) \rceil$ satisfies $k_n^2 = o(n)$, we can use Eq. (3) to write

$$\frac{\beta_{r+1}}{\beta_r} \sim \frac{\binom{k(m+1)}{r+1}}{\binom{k(m+1)}{r}} = \frac{k(m+1) - r}{r+1} < m + 1.$$

Altogether,

$$\frac{n^{-(r+1)} \beta_{r+1} \left\{ r+1 \atop k-1 \right\}}{n^{-r} \beta_r \left\{ r \atop k-1 \right\}} < n^{-1} \frac{m+1}{2} (r+1)^2.$$

Therefore, if we restrict the growth of r (and hence that of k_n) with n such that $r + 1 < n^{(1-\varepsilon)/2}$, we get

$$\frac{n^{-(r+1)} \beta_{r+1} \left\{ r+1 \atop k-1 \right\}}{n^{-r} \beta_r \left\{ r \atop k-1 \right\}} < \frac{m+1}{2} n^{-\varepsilon}, \tag{7}$$

which is sufficient to prove what needs to be shown. Our specific choice for $k_n = \lceil w \log(n) \rceil$ satisfies this constraint.

Having proven that each term in the sum in Eq. (4) is dominated by the preceding one, this sum can be effectively approximated by its first term, i.e.

$$\begin{aligned} n^{mk_n+1} G(m, n, k_n) &= \frac{n^{k_n(m+1)}}{k_n} \sum_{r=k_n-1}^{k_n(m+1)} n^{-r} \beta_r \left\{ r \atop k_n - 1 \right\} \\ &= \frac{n^{k_n(m+1)}}{k_n} n^{-(k_n-1)} \beta_{k_n-1} \big(1 + o(1)\big) \\ &= n^{mk_n+1} \frac{\beta_{k_n-1}}{k_n} \big(1 + o(1)\big), \end{aligned}$$

or simply

$$G(m, n, k_n) = \frac{\beta_{k_n-1}}{k_n} \big(1 + o(1)\big).$$

Given that $k_n \to \infty$ as n grows, we can use Stirling's approximation formula to get

$$
\begin{aligned}
\frac{\beta_{k_n-1}}{k_n} &\sim \frac{1}{k_n}\binom{k_n(m+1)}{k_n-1} = \frac{\big(k_n(m+1)\big)!}{k_n!(mk_n+1)!} \\
&\sim \frac{\sqrt{2\pi k_n(m+1)}}{\sqrt{2\pi k_n}\sqrt{2\pi(mk_n+1)}}\frac{\big(k_n(m+1)\big)^{k_n(m+1)}}{k_n^{k_n}(mk_n+1)^{k_n m+1}} \\
&\sim \sqrt{\frac{m+1}{2\pi m^3}}\frac{u_m^{k_n}}{k_n^{3/2}}. \qquad (8)
\end{aligned}
$$

Ultimately,

$$
\begin{aligned}
\mathbb{E}\Big(\big(s_1^2(\mathbf{W})\big)^{k_n}\Big) &\le nG(m,n,k_n) \\
&= n\frac{\beta_{k_n-1}}{k_n}\big(1+o(1)\big) \\
&= \sqrt{\frac{m+1}{2\pi m^3}}\frac{n}{k_n^{3/2}}u_m^{k_n}\big(1+o(1)\big).
\end{aligned}
$$

Substituting back in the Borel-Cantelli sum in (2), it remains to show that

$$
\sum_n \frac{n}{k_n^{3/2}}\Big(\frac{u_m}{z}\Big)^{k_n} < \infty
$$

for our choice of $k_n = \lceil \frac{3}{\log(z/u_m)} \log n \rceil$. Precisely, since $z/u_m > 1$,

$$
\begin{aligned}
\log\left[\frac{n}{k_n^{3/2}}\Big(\frac{u_m}{z}\Big)^{k_n}\right] &= \log n - \frac{3}{2}\log k_n + k_n \log\big(\frac{u_m}{z}\big) \\
&\le \log n + k_n \log\big(\frac{u_m}{z}\big) \\
&= \log n - k_n \log\big(\frac{z}{u_m}\big) \\
&\le \log n - \frac{3\log n}{\log(z/u_m)}\log\big(\frac{z}{u_m}\big) \\
&= -2\log n,
\end{aligned}
$$

hence the series converges and the proof is complete. □

Acknowledgement

The authors thank Jon Keating for providing feedback on the manuscript.

References

[AIK13] Gernot Akemann, Jesper R. Ipsen, and Mario Kieburg. "Products of rectangular random matrices: Singular values and progressive scattering". In: *Physical Review E* 88.5 (Nov. 2013). ISSN: 1550-2376. DOI: 10.1103/physreve.88.052118. URL: http://dx.doi.org/10.1103/PhysRevE.88.052118.

[Gem80] Stuart Geman. "A Limit Theorem for the Norm of Random Matrices". In: *The Annals of Probability* 8.2 (1980), pp. 252–261. DOI: 10.1214/aop/1176994775. URL: https://doi.org/10.1214/aop/1176994775.

[Lie68] Elliott H Lieb. "Concavity properties and a generating function for Stirling numbers". In: *Journal of Combinatorial Theory* 5.2 (1968), pp. 203–206.

[PŻ11] Karol A. Penson and Karol Życzkowski. "Product of Ginibre matrices: Fuss-Catalan and Raney distributions". In: *Physical Review E* 83.6 (June 2011). ISSN: 1550-2376. DOI: 10.1103/physreve.83.061118. URL: http://dx.doi.org/10.1103/PhysRevE.83.061118.